Lecture Notes in Mechanical Engineering

Woo Bang Lee
Byeongkuen Choi
Lin Ma
Joseph Mathew *Editors*

Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)

Springer

Lecture Notes in Mechanical Engineering

For further volumes:
http://www.springer.com/series/11236

About this Series

Lecture Notes in Mechanical Engineering (LNME) publishes the latest developments in Mechanical Engineering - quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNME. Also considered for publication are monographs, contributed volumes and lecture notes of exceptionally high quality and interest. Volumes published in LNME embrace all aspects, subfields and new challenges of mechanical engineering. Topics in the series include:

- Engineering Design
- Machinery and Machine Elements
- Mechanical Structures and Stress Analysis
- Automotive Engineering
- Engine Technology
- Aerospace Technology and Astronautics
- Nanotechnology and Microengineering
- Control, Robotics, Mechatronics
- MEMS
- Theoretical and Applied Mechanics
- Dynamical Systems, Control
- Fluid Mechanics
- Engineering Thermodynamics, Heat, and Mass Transfer
- Manufacturing
- Precision Engineering, Instrumentation, Measurement
- Materials Engineering
- Tribology and Surface Technology

Woo Bang Lee · Byeongkuen Choi
Lin Ma · Joseph Mathew
Editors

Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)

Springer

Editors
Woo Bang Lee
World Congress on Engineering
Asset Management
Seoul
Korea, Republic of (South Korea)

Byeongkuen Choi
World Congress on Engineering
Asset Management
Gyeongnam
Korea, Republic of (South Korea)

Lin Ma
Queensland University of Technology
Graceville, QLD
Australia

Joseph Mathew
Queensland University of Technology
Sunnybank Hills, QLD
Australia

ISSN 2195-4356 ISSN 2195-4364 (electronic)
ISBN 978-3-319-02461-5 ISBN 978-3-319-06966-1 (eBook)
DOI 10.1007/978-3-319-06966-1
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014946409

Printed on acid-free paper

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

Foreword

The *7th World Congress on Engineering Asset Management* (*WCEAM 2012*) was held at the Daejeon City Convention Centre in Daejeon City, Korea from 8 to 9 October 2012. Close to 200 delegates participated in the congress this year through the support of the Korean Standards Association and DNV.

This year's congress was jointly hosted by the Korea Engineering Asset Management Association (KEAMA) and the International Society of Engineering Asset Management (ISEAM) in recognition of the growing interest in asset management in Korea.

The conference opened with addresses by Mr. Chang Ryong Kim, Chairman of Korean Standards Association and Mr. Sang Duk Park, the Vice Mayor of Daejeon City. Dr. Woo Bang Lee, the congress chair and Prof. Joseph Mathew, Co-Chair also offered their words of welcome to the delegates.

The theme of WCEAM 2012 was "Towards Integration and Interoperability", while the format of the congress covered a wide range of topics.

Interoperability—or lack thereof—is a significant issue across a range of industry and services sectors. A complex challenge to overcome in practice, interoperability fundamentally comes down to managing information flow between the myriad operational systems found across any modern organisation.

Due to their complex nature and multi-stage lifecycle, asset enterprises such as government owned infrastructure organisations and large private asset owners find themselves particularly susceptible to the challenges of making different technical platforms interact. Engineering asset management has evolved as a specialised field dedicated to the optimisation of such organisations. It focuses on the strategic, whole-of-life management of infrastructure and large-scale assets to ensure optimum performance. Not surprisingly, interoperability is seen as a core issue.

The vast range of systems and technology platforms implemented across an asset's lifecycle—from design and construct, to operational maintenance, and de-commissioning—are not built to 'talk' to each other. In fact, they are mostly incompatible, collecting different types of data about an asset's performance that are rarely able to be compared side-by-side in order to ascertain the best way to achieve optimal results. For example, the technical platforms used to design and build large infrastructure assets are different from those used to manage and maintain them. Design and construct platforms cannot accommodate operation and maintenance functionality such as sensor placement and output. The consequence is massive inefficiencies as a result of replicating the information for operating systems.

Simply put, the need for interoperability is costing asset organisations billions of dollars globally.

With so much at stake, it seems pertinent that the 2012 World Congress on Engineering Asset Management (WCEAM) focused on the issue. Historically, Korea has been a world leader in the interoperability space, particularly in relation to integration of its 23 nuclear power stations. Dr. Lee Woo Bang—past Executive Vice President of the Korea Hydro and Nuclear Power Company and Visiting Professor of Pukyong National University, and the driving force behind its integration project, opened the conference keynote presentations highlighting his experiences with Korea Hydro and Nuclear Power Company, He has worked at the enterprise level of utility assets planning, acquisition, operation and maintenance for 35 years and has made significant advances in the area of interoperability in asset management.

Five selected plenary speakers of world repute shared in their learnings and experiences in key topics of interest at the congress.. They were:

1. Dr. Woo Bang Lee: Chairman, Korea Engineering Asset Management Association, Past Ex-Executive Vice President of Korea Hydro and Nuclear Power Company (KHNP) and Visiting Professor, Pukyong National University, Korea. Title—Systems Integration for Asset Management at KHNP.
2. Alan Johnston, President, MIMOSA and Dr. Nils Sandsmark—POSC Caesar Association General Manager. Title—Standards-based Interoperability in the Oil and Gas Industry.
3. Stephen Saladine, General Manager Production, Delta Electricity. Title—Achieving efficiencies in Engineering Asset Management at Delta Electricity, Australia.
4. Adj Prof. David Hood, National President, Engineers Australia, and Founding Chairman of AGIC and Co-Leader Sustainability and Organisational Performance, CIEAM. Title—Achieving Sustainability Outcomes from Infrastructure Asset Management.
5. Professor Kee Bong Yoon, Director, Energy Safety Research Center, Chung Ang University, Korea. Chairman, Government-Public Joint Committee for Safety Auditing of National Energy Facility. Title—Assets Maintenance Environments.

The paper sessions were vibrant and the socials and dinners were of an excellent standard. The highlight of the conference dinner was the award of the first ISEAM lifetime achievement award to Professor Bo Suk Yang, a foundation Fellow of ISEAM who passed away early in the year. Mr. Seung Uk Yang received the honour on behalf of his late father.

All papers published in the proceedings have been refereed by specialist members of an International Review Panel for academic and technical merit.

All this activity would not have been possible if not for the sponsors of WCEAM-AGIC 2012, who were DNV, Daejeon Convention Center (DCC), Korea Tourism Organization, Gyeongsang National University, PI.

Several asset management systems vendor companies mounted exhibition booths in the open area of the congress venue. They were: Korea Gas Technology Corporation, DNV Software Korea, Asset Technology Group, Vision IC and ON Technology.

Organisers: WCEAM 2012 was chaired by Dr. Lee Woo Bang of KEAMA and Prof. Joseph Mathew, CIEAM. The Chairs are indebted to the WCEAM 2012 Organizing Committee chaired by Associate Professor Byeongkuen Choi, Gyeongsang National University with contributions from Dr. Andy Chae of Korea Institute of Construction Technology (KICT), Dr. Lim Gang Min of ATG, Ms. Youngshin Hong of Korean Standards Association, Prof. Lin Ma of Queensland University of Technology, Australia, Dr. Terry Lin of Queensland University of Technology, Australia and Betty Goh of CIEAM, Australia.

The Chairs are also indepted to the team that assisted the Organising Committee in the compilation of the proceedings particularly Dr. Michael Cholette, Dr. Andrew Furda and Dr. Yu Yi, all CIEAM Researchers located at QUT.

The next congress in the series, the eighth, will be held in Hong Kong together with the 3rd International Conference on Utility Management and Safety (2013), from 30 October to 1 November 2013, and will be chaired by Ir Dr. King Wong of the Community and Construction Professionals' Development Centre with Ir Dr. Peter Tse, City University of Hong Kong chairing the Organising Community. For more information, please refer to www.wceam.com.

Woo Bang Lee
Joseph Mathew

Organizing Committee

International Review Panel

Amy Trappey, Taiwan
Andy Koronios, Australia
Andy Tan, Australia
Arun Kumar, Australia
Benoit Iung, France
Brett Kirk, Australia
Christos Emmanouilidis, Greece
David Hood, Australia
David Mba, United Kingdom
Dimitris Kiritsis, Switzerland
Don Sands, Australia
Maria Margarita Zelaya-Nunez, USA
Gerard Ledwich, Australia
Hari Thiruvanam, Australia
Helena.Kortelainen, Finland
Ian Barnard, Australia
Jay Lee, USA
Jayantha Liyanage, Norway
Jie Pan, Australia
Joe Amadi-Echendu, South Australia
Joseph Mathew, Australia
Jun Ni, USA,
Kerry Brown Australia
Kondo H. Adjallah, France
Lin Ma, Australia
Marco Garetti, Italy
Margot Weijnen, Netherlands
Markus Stumptner, Australia
Ming Liang, Canada
Ming Zuo, Canada
Paulien Herder, Netherlands

Peter Tse, Hong Kong
Roger Willet, Australia
Seppo Virtanen, Finland
Terry Lin, Australia
Tony Hope, United Kingdom
Veeraragavan Amirthalingam, India

Congress Chair

Dr. Woo Bang Lee, Korea Engineering Asset Management Association, Korea

Co-Chair

Prof. Joseph Mathew, CRC for Infrastructure and Engineering Asset Management, Australia

Organising Committee

Chair: Associate Professor Byeongkuen Choi, Gyeongsang National University, Gyeongnam, Korea
Dr. Andy Chae, Korea Institute of Construction Technology (KICT), Korea
Prof. Lin Ma, Queensland University of Technology, Australia
Youngshin Hong, Korean Standards Association, Korea
Dr. Terry Lin, Queensland University of Technology, Australia
Betty Goh, CRC for Infrastructure and Engineering Asset Management, Australia

Contents

Contributors

Haftay Hailay Abraha Faculty of Science and Technology, Centre for Industrial Asset Management (CIAM), University of Stavanger, Stavanger, Norway

Basim Al-Najjar School of Engineering, Linnaeus University, Växjö, Sweden

Joe Amadi-Echendu Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Pretoria, South Africa

Paul Barnes Queensland University of Technology, Brisbane, QLD, Australia

Sandra Beach Brisbane, Australia

Jakob E. Beer Universitetet i Stavanger, Stavanger, Norway

Mehdi Behzad Department of Mechanical Engineering, Sharif University of Technology, Tehran, Iran

Srirama Bhamidipati TPM, Delft University of Technology, Delft, The Netherlands

S. Biswas School of Engineering and Built Environment, University of Newcastle, Callaghan, NSW, Australia

Jukka Borgman Department of Computer Science and Engineering, School of Science, Aalto University, Aalto, Finland

Kerry Brown Southern Cross University, Bilinga, QLD, Australia; Curtin Business School, Bentley, WA, Australia; School of Tourism and Hospitality Management, Southern Cross University, Lismore, Australia

Yvonne Brunetto Cooperative Research Centre for Infrastructure and Engineering Asset Management (CIEAM), Queensland University of Technology, Brisbane, QLD, Australia; Southern Cross University and CIEAM, Coolangatta, QLD, Australia

Wahyu Caesarendra Department of Energy and Mechanical Engineering, Institute of Marine Industry, Gyeongsang National University, Tongyeong, Gyeongnam-do, South Korea

Y. Cao Cranfield University, Cranfield, Bedfordshire, UK

Dan L. Carnahan Rockwell Automation, Mayfield Heights, OH, USA

A. A. Cenna School of Engineering and Built Environment, University of Newcastle, Callaghan, NSW, Australia

Myung Jin Chae Construction Engineering and Management Research Division, Korea Institute of Construction Technology, Goyang-si, South Korea

Tommy H. T. Chan Queensland University of Technology, Brisbane, QLD, Australia

P. Chandra Cranfield University, Cranfield, Bedfordshire, UK

Byeong-Keun Choi Department of Energy and Mechanical Engineering, Institute of Marine Industry, Gyeongsang National University, Tongyeong, Gyeongnam-do, South Korea

Dukki Chung Rockwell Automation, Mayfield Heights, OH, USA

François Clemens Section Sanitary Engineering, Delft University of Technology, Delft, The Netherlands

Craig J. L. Cowled Queensland University of Technology, Brisbane, QLD, Australia

John de Croon AssetResolutions B.V., Zwolle, The Netherlands

Peter Crosby ASC Pty Ltd, Osborne, SA, Australia

Em G. delaHostria Rockwell Automation, Mayfield Heights, OH, USA

Subas Prasad Dhakal Curtin University, Bentley, WA, Australia

Subas P. Dhakal Southern Cross University, Bilinga, QLD, Australia

Riad El-Wardani Centre for Industrial Asset Management (CIAM), Stavanger, Norway

Daniel Emerson Queensland University of Technology, Brisbane, QLD, Australia

Ole-Erik V. Endrerud University of Stavanger, Stavanger, Norway

Rodney Entwistle Department of Mechanical Engineering, Perth, Australia

E. Fabres UFRGS—Control Automation and Robotics Group, Porto Alegre-RS, Brazil

Paul Fishchbeck Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA

Gareth Forbes Department of Mechanical Engineering, Perth, Australia

Jing Gao University of South Australia, Adelaide, Australia

Yangming Guo School of Computer Science and Technology, Northwestern Polytechnical University, Xi'an, China

Per-Erik Hagmark Tampere University of Technology, Tampere, Finland

Abrar Haider Mawson Lakes Campus, University of South Australia, Adelaide, Australia

R. V. B. Henriques UFRGS—Control Automation and Robotics Group, Porto Alegre-RS, Brazil

Paulien Herder Section Energy and Industry, Delft University of Technology, Delft, The Netherlands

E. Hernani-Ros Cranfield University, Cranfield, Bedfordshire, UK

Melinda Hodkiewicz School of Mechanical and Chemical Engineering, The University of Western Australia, Perth, Australia

David Hood Science and Engineering Faculty, Cooperative Research Centre for Infrastructure and Engineering Asset Management, Queensland University of Technology, Brisbane, QLD, Australia

T. E. van der Lei Delft University of Technology, Delft, The Netherlands

Ian Howard Department of Mechanical Engineering, Perth, Australia

Jinfei Hu SEAM Lab, and CNDT Lab, Department of SEEM, City University of Hong Kong, Kowloon, Hong Kong

Seyed Mohammad Jafari Sharif University of Technology, Tehran, Iran

M. G. Jones School of Engineering and Built Environment, University of Newcastle, Callaghan, NSW, Australia

Majid Karimi Department of Mechanical Engineering, Sharif University of Technology, Tehran, Iran

Robyn Keast Southern Cross Drive, Bilinga, QLD, Australia

Hack-Eun Kim Korea Gas Technology Corporation, Yuseong, Daejeon, Korea

Tae Silk Kim University of South Australia, Adelaide, Australia

Andy Koronios University of South Australia, Adelaide, Australia

Helena Kortelainen VTT Technical Research Centre of Finland, Tampere, Finland

Buyung Kosasih Department of Energy and Mechanical Engineering, Institute of Marine Industry, Gyeongsang National University, Tongyeong, Gyeongnam-do, South Korea

Arun Kumar Science and Engineering Faculty, Queensland University of Technology, Cooperative Research Centre for Infrastructure and Engineering Asset Management, Brisbane, QLD, Australia

Jouko Laitinen Tampere University of Technology, Tampere, Finland

Jeroen Langeveld Section Sanitary Engineering, Delft University of Technology, Delft, The Netherlands

Gerard Ledwich Queensland University of Technology, Brisbane, QLD, Australia

Sang Hyun Lee University of South Australia, Adelaide, Australia

T. E. van der Lei Delft University of Technology, Delft, The Netherlands

Fengfeng Li School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD, Australia

Xihui Liang Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada

A. Ligtvoet Delft University of Technology, Delft, The Netherlands

Tian Ran Lin Queensland University of Technology, Brisbane, QLD, Australia

J. P. Liyanage Faculty of Science and Technology, Centre for Industrial Asset Management (CIAM), University of Stavanger, Stavanger, Norway

David P. Lowe Queensland University of Technology, Brisbane, QLD, Australia

Nii Nortey B. C. Lokko Centre for Industrial Asset Management (CIAM), University of Stavanger, Stavanger, Norway

Lin Ma School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD, Australia

Muhammad N. Mahmood University of Southern Queensland, Toowoomba, QLD, Australia

Diaswati Mardiasmo Queensland University of Technology, Brisbane, QLD, Australia

Joseph Mathew Korea Gas Technology Corporation, Yuseong, Daejeon, Korea

Ilyas Mazhar Department of Mechanical Engineering, Perth, Australia

D. Mba Cranfield University, Cranfield, Bedfordshire, UK

Mary M. McGeoch Southern Cross Business School, Southern Cross University, Lismore, Australia

Kristoffer K. McKee Department of Mechanical Engineering, Perth, Australia

Hamid Mehdigholi Sharif University of Technology, Tehran, Iran

Nomsesi Martha Mhlanga Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Pretoria, South Africa

Craig A. S. Moodie Department of Energy and Mechanical Engineering, Institute of Marine Industry, Gyeongsang National University, Tongyeong, Gyeongnam-do, South Korea

Richi Nayak Queensland University of Technology, Brisbane, QLD, Australia

Silvia A. Nelson Southern Cross University and CIEAM, Coolangatta, QLD, Australia

Shawn Nielsen Queensland University of Technology, Brisbane, QLD, Australia

Marko Nieminen School of Science, Department of Computer Science and Engineering, Aalto University, Aalto, Finland

Jun-Seok Oh Korea Gas Technology Corporation, Yuseong, Daejeon, Korea

C. E. Pereira UFRGS—Control Automation and Robotics Group, Porto Alegre-RS, Brazil

L. B. Piccoli UFRGS—Control Automation and Robotics Group, Porto Alegre-RS, Brazil

Jawad Raza Operations and Engineering Services, Apply Sørco AS, Stavanger, Norway

Markku Reunanen VTT Technical Research Centre of Finland, Tampere, Finland

Wouter van Riel Section Sanitary Engineering, Delft University of Technology, Delft, The Netherlands

W. Robinson School of Engineering and Built Environment, University of Newcastle, Callaghan, NSW, Australia

C. Ruiz-Cárcel Cranfield University, Cranfield, Bedfordshire, UK

Charles Sampford Queensland University of Technology, Brisbane, QLD, Australia

E. L. Schneider UFRGS—Control Automation and Robotics Group, Porto Alegre-RS, Brazil

Gavin Shaw Science and Engineering Faculty, Cooperative Research Centre for Infrastructure and Engineering Asset Management, Queensland University of Technology, Brisbane, QLD, Australia

A. K. Shrivastava SEAM Lab, and CNDT Lab, Department of SEEM, City University of Hong Kong, Kowloon, Hong Kong

Antony Sprigg Australian Green Infrastructure Council (AGIC), Sydney, NSW, Australia

Tony Stapledon Science and Engineering Faculty, Cooperative Research Centre for Infrastructure and Engineering Asset Management, Queensland University of Technology, Brisbane, QLD, Australia

Yong Sun School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD, Australia

Sola M. Talabi Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA

Andy C. C. Tan Queensland University of Technology, Brisbane, QLD, Australia

David P. Thambiratnam Queensland University of Technology, Brisbane, QLD, Australia

Kiet Tieu Department of Energy and Mechanical Engineering, Institute of Marine Industry, Gyeongsang National University, Tongyeong, Gyeongnam-do, South Korea

Juha Tiihonen School of Science, Department of Computer Science and Engineering, Aalto University, Aalto, Finland

Amy J. C. Trappey Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan

Wan-Ting Tsao Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan

Peter W. Tse Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong

K. L. Tsui SEAM Lab, and CNDT Lab, Department of SEEM, City University of Hong Kong, Kowloon, Hong Kong

Teuvo Uusitalo VTT Technical Research Centre of Finland, Tampere, Finland

Rick Walters Australian Green Infrastructure Council (AGIC), Sydney, NSW, Australia

Xiaojuan Wang Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong

Justin Z. Weligamage Queensland University of Technology, Brisbane, QLD, Australia

Anna Wiewiora Southern Cross University, Bilinga, QLD, Australia

Ype Wijnia AssetResolutions B.V., Zwolle, The Netherlands; Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

K. C. Williams School of Engineering and Built Environment, University of Newcastle, Callaghan, NSW, Australia

Weiliang Wu Queensland University of Technology, Brisbane, QLD, Australia

Matthew J. Xerri Southern Cross University and CIEAM, Coolangatta, QLD, Australia

Gang Xie School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD, Australia

Audu Enoch Yakubu Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Pretoria, South Africa

Xu Yang Queensland University of Technology, Brisbane, QLD, Australia

Ming J. Zuo Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada

About the Editors

Prof. Byeong-Kuen Choi Byeong-Keun is a Professor at the Department of Energy and Mechanical Engineering at the Gyeongsang National University, Korea. He has held academic positions previously at Arizona State University, USA and at the Pukyong National University. His research interests lie in the research and application of signal processing techniques for diagnosing and prognosing machinery faults. He has published widely in the field. Byeongkuen served as Chair of the WCEAM 2012 Organising and Programme Committees and currently serves on the Board of the Korea Engineering Asset Management Association.

Prof. Woo Bang Lee Dr. Lee is Chairman of the Korea Engineering Asset Management Association and is a Visiting Professor at Pukyong National University. He has served in distinguished positions in industry culminating in the position of Executive Vice President, Project Division for Korea Hydro and Nuclear Power (KHNP) Corporation where he spent three years. Prior to that appointment, Dr. Lee was Vice President, Maintenance Planning and Engineering at KHNP. He has been honoured with several special awards which included the Minister's Award, Ministry of Energy and Resources as well as the special Silver Tower Order of Industrial Service Merit in recognition of outstanding service in the resolution of the radioactive waste disposal facility site. He is a lifetime member of the Korean Nuclear Society and the Korean Society of Mechanical Engineering as well as a Foundation Fellow of the International Society of Engineering Asset Management (ISEAM).

Prof. Lin Ma Lin is an accomplished academic and researcher in integrated asset management platforms, optimal decision support and condition-based asset health prediction at Queensland University of Technology (QUT). She led the Capability Optimisation research programme in the CRC for Infrastructure and Engineering Asset Management (CIEAM) and worked closely with industry partners to address their critical needs. Together with her team, a number of advanced asset management systems, modules and industry utilisation cases have been successfully developed and utilised. Her achievements have received commendation. She is a recipient of the Queensland's Smart State—Smart Woman Award as well as the QUT Vice Chancellor's Award for Excellence in Research. She is a Foundation Fellow of ISEAM.

Prof. Joseph Mathew Joe is a Mechanical Engineer with extensive experience in academia, academic administration, research and research management, and consulting. He led the CIEAM since it was founded in 2003 and specialises in condition monitoring, diagnostics and prognostics of machine systems. He is one of the two founders of the International Standards Organisation's Committee on Condition Monitoring and Diagnostics (ISO TC 108/SC5), which he has chaired since 1996. He also initiated the International Society of Engineering Asset Management (ISEAM) and the World Congress on Engineering Asset Management (WCEAM). He remains as Foundation Chair of ISEAM and General Chair of the WCEAM. Joe has recently been appointed as Interim CEO of the newly founded Asset Institute based at QUT and has been the key facilitator behind the Asset Institute®'s transition from a cooperative research centre.

Chapter 1
Review of Theoretical Foundations for Risk Minimal Operations in Complex Sociotechnical Systems: The Role of Human Error

Haftay Hailay Abraha and Jayantha P. Liyanage

Abstract The shift from the traditional to the modern operational risk assessment and management forms in modern complex sociotechnical systems reflects some transitions from a human-focused perspective to a systemic-focused perspective. A key feature, in the later, is that when a human error is found, it is taken as the starting point for wider analysis of a system and not as the conclusion as it so often is in the human-focused perspective. This paper contributes with a comprehensive review of scholarly work aiming at assessing and elaborating on theoretical foundations of the developments that have been taking place within high-risk industries. Some interesting distinctions could be found, mainly, between realist positions and constructivist epistemologies. Both developments show the relevance of taking into account the philosophical antecedents of applied research in explaining the modern trends and how this is valuable toward new paradigms to resolve risk assessment and management issues relating to complex systems, particularly when the dynamics of human, technological, and organizational issues take into perspective.

1.1 Introduction

Complex interaction between technology, human, and organization characterizes modern complex sociotechnical systems. The level of complexity is evident in the compatibility between the traditionally applied reactive and punitive approach to deviations and errors [1] and the need to develop a reporting system and organization culture that supports learning and reflection rather than punishes. In risk/ safety management, it has been common to develop theories/models from

H. H. Abraha (✉) · J. P. Liyanage
University of Stavanger, 4036 Stavanger, Norway
e-mail: haftay.h.abraha@uis.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_1,

empirical cases of either following accidents or from diagnosis of risk/safety. These empirical works and subsequent theoretical developments led to deeper philosophical issues when questioning the use of the theories/models and their relationships with reality. Some of these questions, in relation to assessing the level of safety of organizations, was "how does the past teach us lessons for the future"? Can models constructed from past experience be useful for predicting the future? Can we do it by reducing phenomena to simple model, formula, or principle (issue of reductionism in science)? Ultimately, all these questions led to a core issue regarding the value of the theories/models for risk/safety.

In another way, empirical research on risk/safety has led to several questions regarding the theory of social systems and the relevance of models for managing system performance. Among these questions is how to understand human action as part of sociotechnical system performance ([2–5] etc.). Approaching a possible answer has involved digging into philosophical positions concerning the role of and relationship between human agent, organizational, and social structure. How personnel (of oil and gas industries, aviation, for example) are enabled and constrained by the organizational and social system they are part of and also take account of the role of human intentionality in producing human errors?

Part of the aim of the paper is to elaborate on one of the questions associated with these philosophical positions as it has not yet been treated much in the field of risk/safety science. The aim of an applied field of research in risk/safety science is not, however, to produce philosophical thoughts but to provide useful models and methods so that minimal risk operation is concretely enhanced, particularly in high-risk industries. As mentioned before, both the empirical works and research led us back to philosophical questions: about causality and teleology, about determinism, about induction and deduction, about empiricism and rationalism, about reductionism and positivism, etc. These are today discussed in many scientific areas, from natural to social sciences. The objective of this paper is how these types of questions underpin and are introduced in the risk/safety management approaches. This paper believes that it is important to address the philosophical questions in the context of risk/safety science as any scientific approach is influenced by its philosophical preconceptions. This review of literature also aims to provide a brief account of the developments that have been taking place in risk/safety management in high-risk industries, by taking human, organizational, and technological issues into perspective.

1.2 Theoretical Foundations

Incident investigation approaches have been rooted in empirical sciences and include theoretical assumptions about how to study the natural and social worlds. Much of the literature is concerned with the details of how to conduct an accident investigation and improve safety, whether in a traditional or systems mode. In this practical literature, there is little discussion of the philosophical assumptions implicit in the methodologies advocated. The greatest attention to the theoretical foundations is found in the relatively recent literature where authors are arguing that a systems approach needs to recognize, as the natural sciences have, that the world is not deterministic.

In risk/safety management, philosophical assumptions have practical consequences, i.e., different philosophies lead to differences in whether investigators aspire to being neutral and objective or recognize that their interpretations of evidence are, to some degree, value and theory laden so that another investigator might have noted different factors as significant; whether they look for universal laws that have general application or for context-specific findings; whether they study mental processes or focus on measurable aspects of behavior; whether they take the individual or the system as the key unit of study; and ultimately whether they aspire to create a safe system where errors do not occur or a system that is good at detecting and learning from successes and failures to be able to anticipated future events. Sections 1.2.1 and 1.2.2 discuss the developments from the traditional human-centered approach (i.e., static and deterministic view) approach to the modern systems-centered approach (i.e., dynamic and probabilistic view).

1.2.1 Human-Centered Approach: Static and Deterministic View

The man who is credited with starting the study of safety in science is Heinrich [6]. In his widely read textbook, *Industrial Accident Prevention,* he states his commitment to a scientific approach which he explains as "to find governing laws, rules, theorems … these exist. They are stated, described, illustrated. After understanding comes application."

His view of scientific study takes *Newtonian physics* as the paradigm case. Key assumptions of this form of *inductivism* are a realist view of the world that the truth of observed facts can be objectively determined and that scientists develop theories by observing the world in a neutral manner and generalizing from their observations to develop abstract laws. In the highest form, these are expressed in mathematical terms like Newton's laws of thermodynamics. Another key aspect of the methodology is *reductionism*: Progress is made by reducing complex phenomena to their constituent parts and studying those and the causal relationships between them.

Heinrich's model of accident causation, which is clearly human-centered, is what has been termed "The Domino Theory." Like a row of dominoes falling, there is one "root cause" that triggers the next and so on, until the accident happens. The process is fully determined, and there is an identifiable first step in the chain.

One criticism of this view is the claim to scientific objectivity in identifying some factor as the "root cause." To isolate one factor as the "root cause" as if its identification were a neutral, scientific process is to ignore the importance of the investigator's own priorities, assumptions, or values in deciding on which chain of causation to focus.

Another criticism of the "root cause" concept is that any cause can also be seen as a consequence of prior causes and so there is no obvious point at which to stop investigating the chain of events. The empirical work by Rasmussen [7] on stop rules provides evidence of the ways that investigators reach a decision to stop exploring further but while these are all reasonable, they cannot be considered neutral or objective.

Within this human-centered approach, the dominant psychology was initially the orthodoxy of the time: *behaviorism*. In their efforts to model themselves on the natural sciences, behaviorists avoided study of mental processes as unobservable and hence unscientific. They sought explanations by reference to observable behavior alone, excluding reference to the beliefs, intentions, or motives of actors or the meaning of their actions. This led to theories explaining actions as shaped by features in the world, with particular stress on the importance of positive and negative re-enforcers.

The objective, neutral view of science came under severe attack in the middle of the twentieth century when studies conducted by philosophers such as Feyerabend [8] and Khun [9] showed how scientists did not begin their study of phenomena with an open mind, making observations that were objectively true, but brought to that study a range of assumptions, beliefs, and values. The assumption that there could be theory-neutral observations of the world as a solid foundation to knowledge was undermined by Popper's arguments [10]. Rather, it was argued that people to some degree *construct* the world they experience as a social, interactive process. Some argue that our perceptions of the world result from the combination of the concepts and beliefs that people bring to it, arising from their social environment, and the experiences the world triggers in our sensory skills so that empirical experience places some limit on what we can construct.

When an incident occurs, it has been a standard and understandable response for people to ask why and how it happened. This leads to an investigation into its causation. In analyzing why something happened, we follow a chain of events back into the past. The behavior of the complex, real world is a continuous dynamic flow which we divide into discrete events in order to identify causes and create explanations. In developing an explanation of an unexpected outcome, we look at previous events and seek to pick out ones which were unusual/abnormal as the focus for our analysis of why the process did not proceed as expected. However, "unusual/abnormal" can only be defined against background knowledge of what is normal in those circumstances. The search back through the history of the adverse outcome has no logical end point. Whatever events are identified, it is often possible to ask further questions about why or how they occurred. To summarize, identification of accident causes is controlled by pragmatic, subjective stop rules. These rules depend on the aim of the analysis, that is, whether the aim is to explain the course of events, to allocate responsibility and blame, or to identify possible system improvements to avoid future hazards [7].

The backtracking continues until the investigator reaches a causal factor that is familiar to him. Technical malfunction is one such familiar cause and leads to solutions that improve the accuracy or reliability of technology. The second, even more common, is human error, usually one whose action was close in time and space to the accident made an error of omission/commission that played a significant causal part in the subsequent accident: "if only this operator had taken the correct action then the accident would not have occurred." Human error has readily been accepted as an adequate explanation. The causal explanation that centers on the human being who took the significant action proximate to the accident offers a familiar and plausible story: The engineer failed to notice the temperature was rising, the operator misunderstood the alarms, etc. However, attributing the adverse outcome to the human operators who were nearest to it in time and space ultimately depends on the judgment by someone that the processes in which the operator engaged were faulty and that these faulty processes led to the bad outcome.

When the approach is satisfied with human error as the explanation, then it produces solutions based on that conclusion. If adverse outcomes are seen as due to human error, it follows that solutions are sought that target human performance in order to improve it and reduce the risk of error. It looks as if a basically safe system has been corrupted by poor human performance (see Fig. 1.1), so that improving safety requires reducing or controlling that human element. As a result, according to Woods [11], three broad approaches to improving performance were identified:

1. *Psychological strategies* using punishments or rewards to shape performance and encourage people to operate at a higher level. Naming, shaming, and blaming those deemed responsible give a powerful message to others about the need to improve the standard of work. Management, too, can introduce strategies that monitor and reward a greater attention to complying with accepted standards of good practice.
2. *Reducing the role of humans as much as possible* in engineering, as automation increases, replacing human operators with machines has been a major solution. Even where individuals cannot be removed from the process, there are ways to reduce their scope for independent decision making by introducing increasingly detailed protocols to provide a step-by-step account of how the operation should be carried out.
3. *Increasing monitoring of the workforce* to ensure they are complying with rules and guidance. As protocols take a more significant part in practice, there is a corresponding need to check that they are being followed and to intervene and punish if deviations are observed.

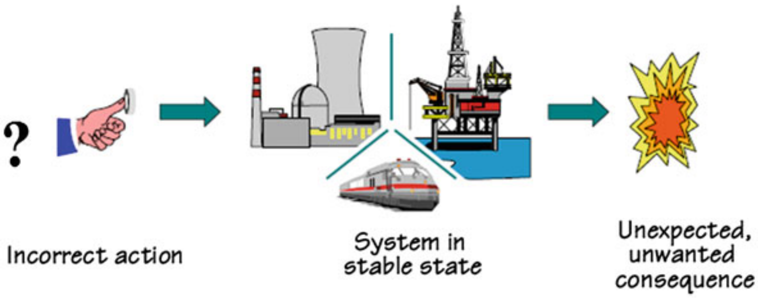


Fig. 1.1 Safe system corrupted by poor human performance

1.2.1.1 Why the Human-Centered Approach Seems So Plausible

The Bias Caused by Hindsight

A major reason for considering that the traditional investigation produces a satisfactory explanation in blaming human error is the bias inherent in examining past events with the benefit of hindsight. Woods lists four reasons why people are ascribed blame for bad outcomes [11]:

- Operators are available to blame. Large and intrinsically dangerous systems have a few, well-identified humans at the sharp end. Those humans are closely identified with the system function so that it is unlikely that a bad outcome will occur without having them present.
- It is difficult to construct a sequence that goes past the people working at the sharp end of the system.
- Human error is so often the verdict after accidents. Studies have consistently shown that people have a tendency to judge the quality of a process by its outcome.
- The evidence about bias in our reasoning gives ground for caution about our reasoning as we investigate back through time to identify the cause of an outcome.

The Fundamental Attribution Error

Another bias in our reasoning that is said to make the human-centered explanation so plausible is the fundamental attribution error [12, 13]. We tend to explain other people's behavior differently from our own. When looking at our own actions, we are very aware of the context in which they were performed and so we tend to offer explanations in terms of that context. When explaining other people, we are most aware of their behavior itself and so focus on our explanations of that rather than the context. Therefore, we attribute a greater causal role to the individual's actions than to the local context in which it took place.

Responsibility and Blame

A final factor that makes the human-centered explanation appealing is that it fits with widespread cultural assumptions about human psychology and morality that focus on individual responsibility for one's actions. The car driver who went through a red light is seen as responsible for taking the wrong action and, as such, is culpable. Identifying the person to blame is psychologically satisfying. It is also reassuring because it suggests that one bad apple has caused the error and the system is basically sound. Although investigators aspire to a no-blame approach and aim to investigate why the person took the action that is now seen as wrong, there are significant legal and moral obstacles to a full blown dismissal of the concept of personal responsibility.

1.2.1.2 The Limitations of the Human-Centered Approach

Besides a concern that human-centered approach was not producing good enough solutions, there was increasing concern that the solutions it produced were themselves inducing new forms of error. The pressure to find a more effective approach was experienced most strongly in high-risk industries, where errors caused the loss of life not just the loss of industrial output. Seeing human error per se as a satisfactory explanation is misleading. It has to be seen relative to the context, not as an absolute, and this is a key feature of the systems approach.

The arguments against the traditional, human-centered approach combine two strands: highlighting the defects in the solutions it has produced and emphasizing how a systems approach offers a more detailed, accurate, and constructive theoretical framework within which to investigate and explain error and safety. The more the risk/safety researchers have looked at the "sharp end," the more they have realized that the real story behind accidents depends on the way that resources, constraints, incentives, and demands produced by the "blunt end" shape the environment and influence the behavior of the people at the sharp end [5].

1.2.2 Systems-Centered Approach

While even quite recent publications refer to the systems-centered approach as "new" [2], the first steps in its development can be seen in the 1940s (e.g., [14]). Since then, it has been further developed, with some significant changes and disputes about theoretical assumptions, and has been adopted in more and more areas of risk and safety management. A more detailed account of the theoretical foundations of the systems approach is provided in Sects. 1.2.2.1 and 1.2.2.2. Before that, brief explanations on some key concepts related to sociotechnical systems are provided.

1.2.2.1 A Sociotechnical System Framework

In contrast to human-centered approach, the systems-centered approach places individuals within the wider sociotechnical system(s) of which they are part. A sociotechnical system is a set of interrelated elements organized for a particular purpose or purposes. It has boundaries, but the boundaries are often permeable—there is movement across them, so the system is exposed to outside forces. For example, each person within a system may be a part of other systems as well. Some of the attributes that warrant a systemic perspective are discussed here.

1.2.2.2 Complexity

The word "complexity" is ambiguous. In this paper, it is linked to complexity theory and nonlinear dynamics. Human-centered approach is deemed to offer too simplistic a model of the work context. This defect has been magnified in recent decades by the growing complexity of current sociotechnical systems in which major accidents occur. As the sociotechnical systems have become more interconnected and complex, simple division is no longer tenable. The functioning of one element has a causal impact on the functioning of others. The role of the human operator has also changed. Many of the tasks previously performed by people are now done by machines, and the human role is to supervise their functioning and intervene if something goes wrong.

The pace of change is another factor to consider: In a traditional work setting, the slow pace of change led to the evolution of fairly stable work procedures and it was easy to define human error with reference to normal practice. In today's work setting, fast pace of changes, fast pace of technological change, new types of hazards, increasing complexity and coupling, more complex relationships between humans and automation, changing regulatory and public views of safety, etc., are making fundamental changes in the etiology of hazards and are creating a need for changes in the explanatory mechanisms used.

This growing complexity leads to the need for an explanation of accidents in terms of structural properties of integrated large-scale systems rather than as isolated links or conditions in a linear causal chain of events [7]. Recognition of the multifactorial nature of causation has highlighted the importance of identifying where in the system the causal factor lies.

1.2.2.3 Seeing People as Understandable and Logical

A key difference between the human-centered approach and the systems-centered approach is that the former is satisfied with the explanation that the accident was caused by human error, while the latter sees human error as the starting point for investigation. Woods points out that "no practitioner intends to make mistakes." The "critical" points and decisions are clear with hindsight but looked different with foresight, given the demands and resources the operator had [15]. When considering how people acted, it is also important to have a realistic idea of human capabilities. The work on human cognitive factors aims to inform our understanding of what standards are likely to be achieved. Simon's concept of "bounded rationality" provides a framework for studying the limitations on our cognitive skills [16]. People use their knowledge to pursue their goals but have finite capabilities and so simplify the tasks they face in order to make them achievable. So they are rational when viewed from the locality of their knowledge, their mind-set, and the multiple goals they are trying to balance.

The aim in a systems approach is to reconstruct the practitioner's point of view within a chaotic, complex, and dynamic system. Analysis is made of the factors that influence human performance, exploring how limited knowledge, a limited and changing mind-set, and multiple interacting goals shape the behavior of people in evolving situations. The task of understanding local rationality is easily compromised by overconfidence. Woods and Cook warn of the need to avoid "the psychologist's fallacy." The phrase was originally coined by the nineteenth-century psychologist William James and refers to the fallacy that occurs when well-intentioned observers think that their distant view of the workplace captures the actual experience of those who perform technical work in context: *distant views can miss important aspects of the actual work situation and thus can miss critical factors that determine human performance in the field of practice.… Understanding technical work in context requires* (1) *in-depth appreciation of the pressures and dilemmas practitioners face and the resources and adaptations practitioners bring to bear to accomplish their goals, and also* (2) *the ability to step back and reflect on the deep structure of factors that influence human performance in that setting* [17].

1.2.2.4 People as the Source of Safety and Manage Risk Dynamically

In systems approach, people are seen not just as a source of error (albeit an understandable or predictable error) but as a source of safety; it is their intelligence and adaptability that is able to identify and intervene when processes are going wrong. Eventually, risk is expected to be a dynamic attribute in a continuously changing context.

Rasmussen made an early case for rethinking the role of practitioners within systems. He pointed out that they do not just follow rules because rules alone are [3] not enough in complex situations to fully determine what should be done. In reality, practitioners make the system work successfully as they pursue goals and match procedures to situations under resource and performance pressure; human work at the sharp end is to "make up for gaps in designers' work." He identified the following factors that are likely to contribute to accidents by reducing the frontline operators' ability to be adaptive and create safety: *undermining the practitioner's ability to be resilient in the face of unanticipated variability;*

producing change, especially change on top of change already underway; providing them with multiple goals that interact and conflict.

In general, the fundamental difference in the two approaches is sketching two mutually exclusive world views: (i) *erratic people degrade a safe system so that work on safety is protecting the system from unreliable people and* (ii) *people create safety at all levels of the sociotechnical system by learning and adapting to information about how we can all contribute to failure* [11].

1.2.2.5 Deterministic View

As discussed in the previous section, the drive to change risk/safety management seems to have come primarily from the failure of the human-centered approach to provide adequate results rather than from any philosophical perspective with its theoretical foundations. There does not seem to be simply one coherent methodology within the broad school of a systems approach. Many of the methodological assumptions of the former approach were carried over. The key difference is in the range of factors analyzed in theorizing the causal pathway that led to the outcome.

Reductionism is still the key strategy of study: reducing complex phenomena to their constituent parts and explaining them in terms of the interactions of those parts: *in human factors and system safety, mind is understood as a box-like construction with a mechanistic trade in internal representations; work is broken into procedural steps through hierarchical task analyses; organizations are not organic or dynamic but consist of static layers and compartments and linkages; and safety is a structural property that can be understood in terms of its lower order mechanisms….this view are with us today. They dominate thinking in human factors and systems thinking, the once pragmatic ideas of human factors and system safety, are falling behind the practical problems that have started to emerge from today's world* [2].

Dekker goes on to summarize three major assumptions: *deconstruction, dualism,* and *structuralism,* which he defines in the following way [2]: (i) *Deconstruction* is the assumption that a system's functioning can be understood exhaustively by studying the arrangement and interaction of its constituent parts. In order to rule out mechanical failure, or to locate the offending parts, accident investigators speak of "reverse engineering." (ii) *Dualism* means that there is a distinct separation between material and human cause—between human error and mechanical failure. Investigations look for potentially broken human components: Was the operator tired, drunk, inexperienced, or "fit for duty"? Human error is then explored in terms of whether the operator was fit for duty. The dualist division of mechanical and human error has become so commonplace we hardly recognize the assumptions implicit in it. Yet, many accidents arise from a combination; changing the design of the mechanical part can reduce the rate of human error: "the error is neither fully human, nor fully engineered" [2]. The dualist assumption contrasts with the alternative approach of locating the causes of accidents in "normal" people doing "normal" work, epitomized by Perrow's concept of normal accidents (see [18]). (iii) *Structuralism*: "the language we use to describe the inner workings of successful and failed systems is a language of structures" [2]. Organizations, for example, are described in terms of a "blunt" end and a "sharp" end [5]. In place of Heinrich's domino model of accident causation, the dominant image here in the deterministic systems-centered investigations has been Reason's "Swiss Cheese Model." He conceptualizes a system as having a series of defense layers to detect and prevent error, but each of those layers is imperfect, i.e., they have holes in them (termed as active and latent failures). The necessary condition for an organizational accident is the rare conjunction of a set of these holes in successive defences, allowing hazards to come into damaging contact with people and assets [5]. This model has flourished for decades and is responsible for many of the lessons now generally adopted. Numerous organizational factors were identified as the causal factors that contributed to the probability of the accident. However, this model does not give a clear explanation how these causal factors combined to provide the circumstances for an accident to take place.

1.2.2.6 Probabilistic View

By sharing the aspiration of using the methods of the natural sciences, brief discussion in the developments of those sciences that have fundamentally altered the theoretical foundations is provided below.

The Nature of Causality

The most significant change has been our understanding of the nature of causality. In the *positivist* and, to some degree, in the *empiricist* approaches, causality is seen as *deterministic*. The causal chain is linear in nature—if A occurs, then B occurs—allowing prediction and control, at least in principle. This deterministic view of the world was dominant in physics until the early twentieth century when developments in quantum mechanics proved that the world was *probabilistic* in nature. This change in the concept of causation has gained widespread acceptance in the rest of the natural sciences.

The deterministic systems approach assumes that a system can be made safe: With sufficient care and attention, the various latent errors can be rectified so that, in theory at least, it becomes possible to design a safe system. The ideal model of a safe organization requires top-down control, specifying how all the constituent elements should interact. It is assumed that if enough data are known, it should be possible to predict what interactions will be caused. Again, in principle at least, it should be possible to develop a set of procedures that covered all aspects of the task.

However, according to the *probabilistic* view of the world, complex systems have *emergent properties* (for example see [4, 19, 20]). An emergent property is one which arises from the interaction between lower-level entities, none of which show it. If you mix blue and yellow paints, the color green "emerges." There is no green in the original blue and yellow paints, and we cannot reverse the process and divide the green back into the two other colors. This raises questions about how far it is possible to predict what will happen as elements in a system come together. Factors that, on their own, are safe may become unsafe as they form a system with others.

This emergent model of accident causation conflicts with Reason's Swiss Cheese Model because for Reason, a latent failure (or hole in a defensive layer) is a constant failure, but in the emergent view, an error may "emerge" from the combination of other factors, none of which is necessarily a latent failure in Reason's meaning of the term but only in combination with other factors does the potential for failure "emerge."

Moving from a deterministic view of the universe to a probabilistic one has radical implications. Wallace and Ross [21] sum up the difference: *instead of a deterministic view in which certain causes inevitably lead to certain effects; in the probabilistic view, certain causes may or may not lead to effects with differing degrees of probability.*

The Importance of the Social

A consequence of viewing complex systems as, to some degree, unpredictable is that it strengthens the need to find out what went on in the run-up to an accident (or what goes on in risk analysis and assessment) in a more detailed manner. Specifically, more attention is paid to the social dimension of sociotechnical systems. The study of human factors in incident investigations had predominantly used cognitive psychology which paid little attention to complex social interactions.

In safety management, this policy was followed through for the most part, but this focus on what goes on individuals' heads has been challenged on the ground that it is the actual performance that needs to be studied, that is, the actions of the person in a context [22]. Studying people in their context highlights the importance of the influence of culture on meaning. This is apparent in making sense of "local rationality"—how people are making sense of their working environment and making decisions. If meaning is seen as, at least in part, the product of social negotiations within a group, then it is important to know how those involved in the error were constructing the work environment in which they were making decisions and acting. It is argued by some that to fully understand how the world looks to practitioners in the field, we have to recognize how people use talk and action to construct perceptual and social order; how, through discourse and action, people create the environments that in turn determine further action and possible assessments, and that constrain what will subsequently be seen as acceptable discourse or rational decisions. We cannot begin to understand drift into failure without understanding how groups of people, through assessment and action, assemble versions of the world in which they assess and act [2].

The importance ascribed to understanding the "local rationality" of the operators, how they made sense of the situation at the time and why they acted as they did, has also led to an interest in the social construction of what counts as an error and what is normal practice. Earlier approaches had taken it as possible to define “error” and “failure” in objective terms that are agreed by all in the organization. Cognitive psychology contains a mechanistic model of the human mind as an information processing system separate from the phenomena it processes. Such a model severely restricts our ability to understand how people use talk and action to construct perceptual and social order; how, through discourse and action, people create the environments that in turn determine further action and possible assessments, and that constrain what will subsequently be seen as acceptable discourse or rational decisions [2]. Accident investigations often focus on whether or not operators complied with procedures as a definition of “safe” practice, but in reality, local cultures often develop in which normal practice deviates from the official manual. This is usually due to difficulties arising in the specific context in trying to comply with the procedures. They are conflicting with other aspects of the work. Local, deviant cultures may, in practice, be good methods of improving the safety of the system and minimizing the clumsiness or inaccuracy of the official procedures. It cannot be decided in advance that all deviations are “errors.”

Mitigating Systems Risk and Improving Safety

Although, in the probabilistic sociotechnical systems approach, systems do not have the property of linear predictability and controllability, they are far from chaotic and have mechanisms that can increase the safety of a system. They are self-organizing, i.e., they have feedback loops, so they are able to learn and adapt. Systems for reporting near-miss incidents can be seen as an example of such a feedback loop that permits learning and modifications to take place before a major accident occurs. The importance of such feedback loops is heightened in this approach because it is assumed that interactions lower in the system will be unexpected and senior management or designers cannot predict all that may occur.

1.3 Discussion and Conclusion

The study of risk and safety management has developed over the decades, with considerable success in identifying and reducing errors and accidents. This paper has reviewed two philosophical perspectives, based on experiences for research in risk/safety sciences. On the one hand, realist positions provide good framework to some researchers on risk/safety, and on the other hand, other researchers on safety use constructivist theory of knowledge as useful approach in their projects. The two perspectives are, however, quite different and imply different positions with regard to ontology and epistemology. Choosing one over the other can be linked to personal conviction and beliefs [23].

The dominant view has been that studies of risk/safety should aspire to the methods of the natural sciences. Since the focus of study has been both on technical and on human contributions to error, the social as well as the natural sciences have been drawn upon, but the tendency has been to prefer those social science schools than model themselves on the natural sciences. Hence, in psychology, there has been a preference initially for behaviorism and, later, for cognitive psychology with less attention paid, until recently, to humanistic or interpretivist schools of thought. In relation to social factors, perhaps the key difference has been whether attention is given to the social dimension or whether the focus is primarily on individuals.

At a philosophical level, there is a strong continuity between the traditional human-centered approach and the deterministic systems-centered approach, i.e., similar methods of study are used (i.e., static and deterministic). However, the focus in the later is broadened out from the human to the human-context dyad. The most interesting conflicts appear within the probabilistic systems-centered approach. There is an emerging school of thought arguing that the developments in natural science involving complexity and nonlinear dynamics undermine the common assumptions about linear causality and predictability. This change has major implications for the whole framework in which error and accident investigations are conducted. More attention is paid to the social system and to understanding how operators, partially at least, construct the context within in which they operate. This school also raises questions about the overall goal of risk and safety management: whether it is possible even in theory to create a system where errors and accidents do not occur or whether the aim should be to design systems that can detect and learn from the inevitable errors that will arise and take proactive measures.

A number of methods of accident analysis (risk analysis) have been developed in different industries. All the methods have features in common. All try to answer the three basic questions: What happened? How did it happen? and Why did it happen? In many high-risk industries, there has been a shift from a traditional, human-centered approach to a systems-centered focus. This fundamental change views the complex interaction of all elements in the system as potentially contributing to both safety and error. Human error is no longer seen as an adequate explanation of a mishap but as requiring closer investigation to find what factors in the wider context contributed to that operator's commission or omission. The range of factors that could be considered is immense, including factors relating to human cognition and emotion, factors in the design of technology and how it shapes human performance, and factors in the organizational culture, resources, priorities, and teams.

While all in risk and safety management have aspired to a scientific approach, there have been differences of opinion on what this entails. Early researchers took a relatively narrow, positivist view, while later researchers embrace a wider range of methods within an empirical framework. More recent writers have highlighted the distinctive features of studying systems as opposed to sets of individuals. Complexity in as much as it implies nonlinear dynamics places limits on prediction (i.e., the cause–effect in complex system is difficult to expect: Different causes can lead to the same effect, same cause leads to different effects, small cause can lead to large effect, and large cause can lead to small effects).

Systems are seen as having emergent properties, arising from the interaction of lower-level entities, none of which show them. There is a greater concern to understand the "local rationality"—how the individual workers are making sense of their environment and their tasks. This leads to a greater interest in interpretivist psychologies and the social construction of the workplace unlike their predecessors who limited the study of human actions to observable behavior or, later, the observable elements of cognition. The focus on systemic vulnerabilities, not individual lapses, slips, etc., has led not only to changes in how accidents are investigated but also to the development of strategies for continual learning in organizations so that weak areas of practice can be identified and improved before a major incident. Recognition of the complex interplay of variables undermines any simplistic, top-down model of policy implementation. The project of adopting a systems-centered approach to mitigate potential vulnerabilities and improve safety in any high-risk organizations has the potential for significant learning and progressive adaptation. Learning how to reduce errors, learning how to improve, and adapting a conscious practice in a dynamic setting are ways of looking at the same central issue. A system that is designed to tolerate dynamism and promote continuous adaption is also one that has effective attributes to explain what is normal or abnormal in a specific context.

References

1. Hoyland S, Aase K (2008) An exploratory study on human, technological and organizational interactions with in health care: safety science monitor http://ssmon.chb.kth.se/vol12/1hoyland.pdf. Accessed 20 Jun 2012
2. Dekker S (2006) The field guide to understanding human error. Ashgate, Aldershot
3. Rasmussen J (1986) Information processing and human–machine interaction: an approach to cognitive engineering. North-Holland, New York
4. Rasmussen J (1997) Risk management in a dynamic society: a modelling problem. http://www.sciencedirect.com/science/article/pii/S0925753597000520. Accessed 03 Mar 2012
5. Reason J (1997) Managing the risks of organizational accidents. Ashgate, Rarnham
6. Heirich WH (1941) Industrial accident prevention. Mcgraw- Hill, New York
7. Rasmussen J (1990) Human error and the problem of causality in analysis of accidents. http://rstb.royalsocietypublishing.org/content/327/1241/449.short Accessed 12 Feb 2012
8. Feyerabend P (1975) How to defend society against science. http://anarcosurrealisti.noblogs.org/files/2010/10/Feyerabend-Paul-How-to-defend-society-against-science.pdf Accessed 15 Feb 2012
9. Kuhn T (1970) The structure of scientific revolutions. The University of Chicago Press, Chicago
10. Popper K (1963) Conjectures and refutations. Routledge, New York
11. Woods D, Johannesen L, Cook L, Sarter N (1994) Behind human error: cognitive systems, computers and hindsight. http://www.dtic.mil/cgibin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA492127 Accessed 21 Feb 2012

12. Dekker S (2002) The field guide to human error investigations. Ashgate, Aldershot
13. Reason J, Hobbs A (2003) Managing maintenance error. Ashgate, Aldershot
14. Garland DJ, Wise JA, Hopkin VD (1998) Hand book of aviation human factors. CRC Press, Mahwah
15. Woods D (2003) Discovering how distributed cognition systems work. In: Hollnagel E (ed) Handbook of cognitive task design. Lawrence Erlbaum Publishers, London
16. Simon H (1996) The sciences of the artificial. MIT Press, Cambridge
17. Woods D, Cook R (2002) Nine steps to move forward from error. http://csel.eng.ohio-state.edu/woods/error/9steps_forward.pdf. Accessed 15 Feb. 2012
18. Perrow C (1984) Normal accidents: living with high-risk technologies. Basic Books, New York
19. Hollnagel E, Woods DD, Leveson N (2006) Resilience engineering: concepts and precepts. Ashgate, Aldershot
20. Leveson N (2004) A new accident model for engineering safer systems. http://www.sciencedirect.com/science/article/pii/S092575350300047X. Accessed 13 Feb 2012
21. Wallace B, Ross A (2006) Beyond human error: taxonomies and safety science. CRC Press, New York
22. Gardner H (1985) The mind's new science: a history of the cognitive revolution. Basic Books, New York
23. Le Coze JC, Pettersen K (2008) Is resilience engineering realist or constructivist? http://www.resilience-engineering.org/RE3/papers/Le_Coze_Pettersen_text.pdf. Accessed 14 Mar 2012

Chapter 2
Review of Theories and Accident Causation Models: Understanding of Human-Context Dyad Toward the Use in Modern Complex Systems

Haftay Hailay Abraha and Jayantha P. Liyanage

Abstract Many accident causation models/theories have been dominating the human factors literature from a range of viewpoints and in a variety of different industrial contexts. However, many of the theories/models have limited applications with respect to capturing the underlying accident causations in modern complex systems. As a result, we seem to be moving toward new incident potentials, particularly in high-risk industries. As observed, major accidents keep occurring that have similar systemic causes in different contexts. This creates serious doubts that the existing analysis methods are capable of discovering the underlying causality in new complex settings, and/or how do we transfer what we learn from past to new contexts. Many of the approaches to safety focus on enforcing 'defenses,' i.e., physical, human, and procedural barriers. This perspective has limited view of accident causality, as it ignores today's changing factors in emerging complex work systems and their interactions that generate various potentially unintended situations and shape the work behavior. Understanding and addressing these causal factors is therefore necessary to develop effective accident prevention strategies. Thus, this paper reviews the popular theories/models, by focusing on the developments in modern industrial environments toward more complex systems and interactive or collaborative operations. Accident models founded on basic systems theory concepts, which endeavor to capture the underlying accident causality and complexity of modern socio-technical systems from a broad systemic view, are analyzed and new insights for future accident analysis efforts are identified.

H. H. Abraha (✉) · J. P. Liyanage
Faculty of Science and Technology, Centre for Industrial Asset Management (CIAM), University of Stavanger, 4036 Stavanger, Norway
e-mail: haftay.h.abraha@uis.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_2,

2.1 Introduction

Many accident causation theories/models with their specific approaches to accident analysis have been dominating the human factors literature. However, many of the theories/models have limited applications with respect to capturing the underlying accident causations in modern complex systems.

Since World War II, the types of systems we are building and the context in which they are being built has been dramatically changing. According to [1–9], the main changes and observation can be highlighted as shown in Table 2.1.

Leveson [1] argues these changes are challenging both our accident models and the accident prevention and risk assessment techniques based on them. In the traditional view, accidents are seen as a result of failure or malfunction of components (machines or human). These failures follow a predefined path of cause and effect. Technical and human factors studies have focused on cause-effect relationships *close in time and space* to the accident sequences. Technical risk assessment relies on expected behaviors of installation components, which can fail or not when fulfilling their function. Potential failures can be introduced in event and fault trees (FTs) with the help of various risk/reliability analyses techniques. Following an accident involving a technical failure, causes are often linked to the component characteristics (its design, its expected reliability, its expected resistance, its level of condition, etc.)—'engineering error paradigm.' Whereas, human errors involved in accident sequences are often linked to shaping factors associated with cognitive models like the SRK Rasmussen's model [8],—'cognitive error paradigm' or other versions of this type of model which includes personal characteristics—'individual error paradigm,' (examples see, [10, 11]).

Human actions can also be introduced in event and FT at the risk assessment stage with the help of human reliability analysis (HRA) techniques such as HEART [12], THERP [13], CREAM [14], and SPAR-H [15]. In modern times, seeing organization as the origin of safety and accidents—'organization paradigm'—has become a trend in research and is described as complementing the technical and human factors oriented studies (examples: [10, 11]). The organizational perspective widens the scope of the technical and human factors approaches and opens the boundaries to more variables than the two others do.

Within the petroleum industry, human factor analyses of drilling and wireline operations have shown that human and organizational factors were the major elements that could explain the causes of various major accidents [16]. The main challenge, however, is the focus in the industry is still related to technical solutions, rather than a systemic looking into the complex interactions between human, technical, and organizational (HOT) issues to see the 'big picture.' This paper will look into the features of emerging complex systems and changes in the working context. It reviews the established accident models and their limitation to capturing the underlying accident causations of modern complex systems. It then analyzes accident models founded on basic systems theory concepts, which endeavor to

Table 2.1 Changes and observations toward new complex systems

Fast pace of technological change Lessons learned over centuries about designing to prevent accidents may be lost or become ineffective when older technologies are replaced with new ones. New technology introduces unknowns into our systems
Changing nature of accidents For example, our increasing dependence on information communications technology (ICT) systems (for example, utilization of real-time data, automatic monitoring, etc.) is creating the potential for loss of information that can lead to unacceptable losses
Decreasing tolerance for accidents The losses stemming from accidents are increasing with the cost, HSE, and potential destructiveness of the systems we build
Increasing complexity and coupling Increased interactive complexity and coupling make it difficult for the designers to consider all the potential system states or for operators to handle all normal and abnormal situations and disturbances safely and effectively, as the risk factors are dynamic and their significance varies as context or interaction changes
More complex relationships between humans and automation Inadequacy in communication between humans and machines is becoming an increasingly important factor in accidents
Changing regulatory and public views of risk and safety In our increasingly complex and interrelated societal structure, individuals no longer have the ability to control the risks around them and are demanding that policy makers or regulators assume greater responsibility for controlling behavior through laws and various forms of oversight and regulation

capture the underlying accident causality and complexity of modern socio-technical systems from a broad systemic view, and new insights for future accident analysis efforts are identified.

2.2 Emerging Complex Systems

Accident history in recent decades has clearly shown that human and organizational factors play a significant role in the risk of system failure and accidents [17], as can be seen in Fig. 2.1.

Complex interaction between technology, human, and organization characterize modern complex socio-technical systems. Modern socio-technical systems are complex, dynamic, and non-decomposable. These three qualities appear from a unique characteristic of these systems: They are emergent. The fact of being emergent means that the complex relationships among causes and effects make unexpected and disproportional consequences to emerge; in other words, certain combinations of the variability of system functional actions as a whole achieve the threshold allowed for the variability of a system function [17].

Lewes [18] proposed a distinction between resulting and emergent phenomena in which the resulting phenomena could be predictable starting from its constituent parts and the emergent phenomena could not.

We are now living in an increasingly interconnected world of socio-technical systems. The multi-scale nature and complexity of these networks are crucial features in understanding and managing the socio-technical systems and the dynamic processes occurring within them.

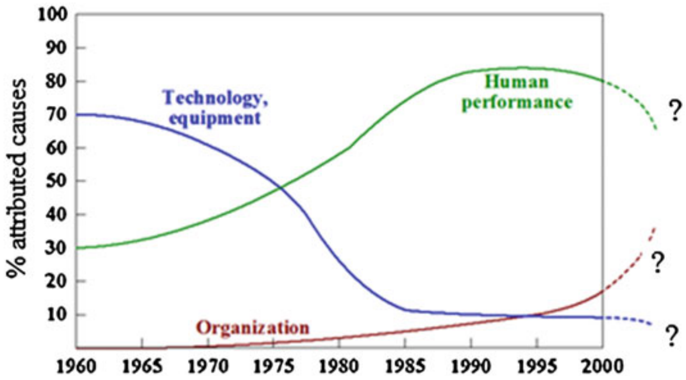


Fig. 2.1 Trends in attributed accident causes (adapted from [17])

In general, complexities of emerging modern socio-technical systems can be summarized as follows:

- Multi-scale complexity: multiple level of social, technical, and organizational scales (micro–macro emergence).
- Multi-agent complexity: distributed roles and responsibilities across different partners.
- Complexities arising from technology divide and nonstandardization: different technological solutions that are not interoperable and nonstandardized applications.
- Complexities arising from changes in culture, work, and power stereotypes.
- Complexities arising due to increasing diversities in work systems.

The difficulty in understanding and addressing human and organizational risk of accidents lies in the fact that they are complex multi-level problems of such socio-technical systems. Due to the complex nature of the socio-technical systems, new techniques for analyzing the problem need to be developed. One should focus on identifying human, technological, and organizational interactions rather than individual influences by combining several theoretical perspectives. A holistic approach, to model the complex and dynamics of human and organizational risk, is required (holism is the idea that the properties of a system cannot be determined or explained by the sum of its components alone). Such approach represents the socio-technical system as networks of dynamic interrelationships among people, knowledge, tasks, resources, technologies, organizations, and the surrounding environment. Example of emerging complexity in aviation industry is shown in Table 2.2.

Table 2.2 Emerging complexity in aviation industry

Example: National Airspace System (NAS)
The National Airspace System (NAS) is an emerging complex socio-technical system. FedEx and UPS have expressed interest in unmanned aircraft (UA)-provided transport of cargo, as they believe there is financial reward in doing so. Military groups have pushed the Federal Aviation Administration (FAA) for more waivers to fly military drones through the NAS on the way to flight testing zones or to patrol the US borders. The addition of unmanned aerial systems (UASs) into the NAS would present a major disruptive technology with emerging new types of risks (i.e., risk of accidents, potential financial and security rewards, etc.). Besides, UASs are not only complex machines with millions of lines of code; they are operated by teams of ground-based pilots. These pilots do everything from manual flight control of the planes (e.g., keeping the wings level) to high-level mission-based planning (selecting waypoints and routing aircraft to interesting spots). The organizational control over UAS and NAS integration is also complex, involving private industry, unions, civilian advocacy groups, all military branches, and civil regulators [32]. At this point, it is unknown what level of control or influence each of these groups should contribute to making decisions about how to integrate UASs, much less than what their role should be in the integrated NAS

2.2.1 Changes in Work Context

The levels and layers within a modern socio-technical system are stressed by a number of external forces and counter forces in their coping with accident risks. The main contextual stressors influencing work performance risks can, for instance, be changing political climate and changing public risk awareness, market conditions and financial pressure, competence and education, and fast pace of technological change [7]. Adaptive responses to these elements have changed the everyday reality of work, the contents of work processes, and the socio-technical systems at traditional workplaces.

The increased use of ICTs as integral parts of manual work and the construction of new distributed industrial organizations change the characteristics of the work: creates new types of communication and improved ability to store and retrieve data and more effective information processing, which all influence modern organization of work [19]. It creates advantages such as automation of work processes; real-time surveillance control; more effective planning and communication; and improved availability of employees. But this 'ICT revolution' also creates pitfalls such as information overload; the increased request of information may disturb sharp-end work, create data security problems, loss of nonverbal communications, etc. Table 2.3 below shows an example of how ICT can influence (positively and negatively) operators at the 'sharp-end' activities of offshore installations.

To sum up, various forms of changes in modern industrial systems; technological, knowledge, organizational, information flow, industrial work life, people, values, etc. have begun to challenge our understanding and interpretation of behaviors and performances. Subsequently, when it comes to accident prevention

Table 2.3 Monitoring work performance at offshore installations

There is an ongoing transition to the concept-integrated operations (IOs) within the oil and gas companies operating in the Norwegian continental shelf (NCS), i.e., use of information technology and real-time data to improve decision-making processes and cooperation across disciplines and organizations. One of the implications of this development is increased use of monitoring of offshore employees. This implies that operators onshore can watch offshore employees' performance by use of camera equipment and monitors. On the one hand, this creates a secure and safe environment as offshore employees have a 'watchful eye' on their performance, making it possible to prevent and stop unwanted actions. It also provides offshore employees with decision support from onshore experts. On the other hand, monitoring might lead to workers feeling uncomfortable of being evaluated all time and even a sense of mistrust

analyses, there is still such a considerable reliance on the domino model and barrier metaphor. The review and discussion in the following sections are based on the following perspectives on organizational accidents and resilient organizations with respect to modern developments and contexts.

Theses perspectives are the following:

- The domino and energy and barrier perspective understands and prevents accidents by focusing on dangerous energies and means by which such energies can reliably be separated from vulnerable targets [19–21].
- The information processing perspective, taking Reason's theory of Swiss cheese model (SCM) as a starting point [10, 11, 22], understands an organizational accident as a breakdown in the organizational flow and interpretation of information that is linked to physical events.
- The decision-making perspective, with a focus on handling of conflicting objectives, e.g., Rasmussen's [9, 23] model of activities migrating toward the boundary of acceptable performance.
- Perrow's [5–7] theory of normal accidents, which explains a situation where accidents are normal outcome due to emerging complex systems that are tightly coupled and interactively complex.
- The theory of high-reliability organizations (HRO) [24] is grounded in intensive studies of organizations that have demonstrated an outstanding capacity to handle fairly complex technologies without generating major accidents. Important concepts from this research tradition are organizational redundancy and a capacity of organizations to reconfigure in adaptation to peak demands and crisis.
- The theory of resilience engineering (RE) is a new concept developed as a response to shortcomings in traditional safety management approaches. It focuses on a socio-technical system's ability to attend to, monitor, and cope with performance variability [2].

2.3 Review of Established Accident Models/Theories with Respect to Modern Systems

Accident causation model, like any model, is a simplified representation of reality. It should, however, highlight the most important features of a phenomenon in order for it to be useful in identifying risks of accidents. Deficiencies of each model to cope with the changing world are the driving forces for the emergence of a multitude of different approaches. The pros and cons of each model should be identified, and therefore, there is a need for several accident models. On the one hand, identification of the causal relationships in accidents has been a research interest for a long period. For example, accident statistics with selected main cause(s) of each accident has been gathered. On the other hand, it has turned out that the multitude of contributing factors and the dynamic and complex nature of the chain of events leading to the accident make it very difficult to find the causal relationships from statistics. Three different types of accident causation models can be distinguished today:

- sequential accident models,
- epidemiological accident models, and
- systemic accident models.

2.3.1 Sequential Accident Models

One of the earliest accident causation models is the *Domino Theory* of accident models proposed by Heinrich in the 1940s [20] and implies the linear one-by-one progression of multiple chain of events (domino effect) leading to the accident and finally its consequence, e.g., an injury.

FT and event tree (ET) models are standard risk analysis methods for component failure analysis. Such approaches, however, have serious limitations in the analysis of complex socio-technical systems (e.g., man–machine interaction). Sequential models assume that the cause-effect relation between consecutive events is linear and deterministic.

2.3.2 Epidemiological Accident Models

The pioneering model with multi-causality was the *Energy model*, by Gibson [19], and it was later developed by Haddon [21]. The core of energy model lies in the fact that the consequences of an accident are always based on the transfer of energy (in one or another form: mechanical, chemical, thermal, electrical, etc.), which is affected by a barrier. The widely used concept of a *barrier*, which protects the asset from hazardous effects of energy, is another key concept in the energy model and has had an important effect on many other accident models.

Bowtie model can also be considered as epidemiological accident model, which is built with the combination of a FT model and an ET model; thus, it integrates the elements and options affecting the probability/frequency of a critical event (an accident) with its consequences. A bowtie model demonstrates how a critical event may have several precursors as well as several consequences, and hence, it accounts for multi-causality.

Reason's [10, 11] *SCM* of defenses is a major contributor to this class of models and has greatly influenced the understanding of accidents by highlighting the relationship between *latent* and *active* causes of accidents. Understanding of the causal factors of an accident with the linkage to human and organizational errors was enhanced with the structural division of the human performance on the SRK level [25].

2.3.3 Systemic Accident Models

These models attempt to describe the characteristic performance on the level of a system as a whole, rather than that of specific cause-effect 'mechanisms' or even epidemiological factors [26]. The goal of systemic accident model is to describe the dynamic and nonlinear behavior of risks within a socio-technical system. This holistic viewpoint of complex systems allows us to understand the underlying mechanisms of evolution of socio-technical systems. Some of the basic models that fall in this group include the following (see Sect. 2.4 for elaboration):

1. Rasmussen's socio-technical framework
2. Normal accident theory (NAT) and HRO
3. RE

2.4 Review of Alternative Approaches to Accident Prevention

Accidents in complex high-risk operations, such as aviation, petroleum, and nuclear, are frequently explained as a result of system failures; however, few methods exist that can adequately be used to investigate how the variability of individual, technical, and organizational performance in combination may lead to an adverse outcome [26]. The science of complexity offers a paradigm that is a viable alternative to positivist, reductionist approaches to accident analysis/risk assessments. Put simply, complex systems are characterized by having multiple interacting parts that exhibit nonlinear behavior leading to unpredictability and being made up of nested systems that are open and mutually affecting with each level exhibiting system properties that emerge out of the interactions of the parts. Risk analysis/accident analyses of such wholes must be systemic and should focus on the interactions between the parts of the system while recognizing patterns produced by self-organization and feedback loops that lead to adaptive behavior.

2.4.1 Rasmussen's Socio-technical Framework

Rasmussen [9] claims that sequential and epidemiological accident models are inadequate for studying accidents in modern socio-technical systems. Rasmussen advocates system-oriented approach based on *control theory* concepts, proposing a framework for modeling the organizational, management, and operational structures that create the preconditions for risk/accidents (see Fig. 2.2a, b)

According to Rasmussen [9], it is not possible to establish procedures for every possible condition in complex and dynamic socio-technical systems. In particular, this concerns emergency, high-risk, and unanticipated situations. Decision-making and human activities are required to remain between bounds of workspace, which are defined by administrative, functional, and safety constraints.

Rasmussen [9] argues that in order to analyze the safety in a work domain, it is important to identify the boundaries of safe operations and dynamic forces that may cause the socio-technical system to migrate toward or across these boundaries. Due to the combined effect of management pressure for increased efficiency and a trend toward least effort, Rasmussen argues that behavior is likely to migrate toward the boundary of unacceptable risk. The exact boundary between unacceptable and acceptable risk is not always obvious to the actors in complex systems where different actors attempt to optimize their own performance without complete knowledge as to how their decisions may interact with decisions made by other actors. At each level in socio-technical hierarchy, people are working hard to respond to pressure of cost-effectiveness, but they do not see how their decisions interact with those made by other actors at different levels in the system. Rasmussen claims that these uncoordinated attempts of adaptation are slowly but surely 'drifting into failure.'

2.4.2 Normal Accident Theory and High-Reliability Organizations Theory

There is some overlap between these theories, and this paper just wants to highlight key elements and possible lessons to be learned and applied in the field of risk/accident analyses. NAT [5–7] introduced the idea that in some systems accidents are inevitable and normal. Such system accidents involve the

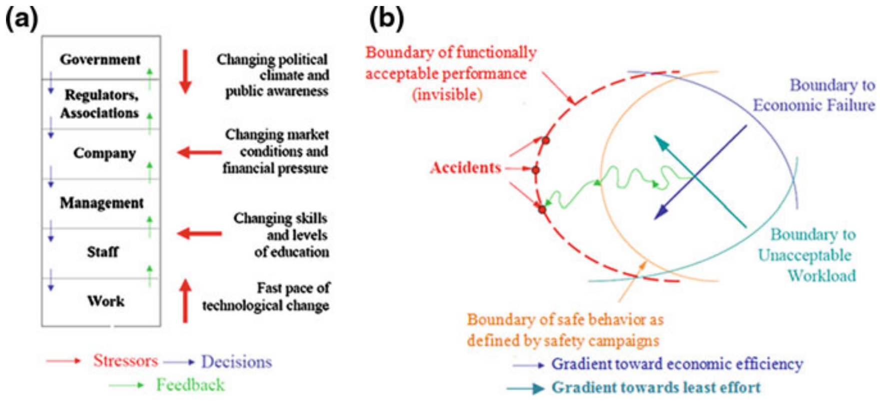


Fig. 2.2 Rasmussen's socio-technical framework: **a** structural hierarchy; **b** systems dynamics and boundaries of safe operation (adapted from [9])

unanticipated interaction of multiple failures. Perrow uses two dimensions in a two-by-two table to indicate that different systems may need different ways of organizing. If the system is both interactively complex and tight coupled, there is no possibility to identify unexpected events, and the system should be abandoned. In such systems, simple, trivial incidents can cascade in unpredictable ways and disastrous consequences. The changes in working context, described above in Sect. 2.2.1, have resulted in increased vulnerabilities at most workplaces over the last twenty years. Therefore, it may be important to reduce interactive complexity and tight coupling in the design of workplaces.

HRO researchers claim to counter Perrow's hypothesis on high-risk, complex systems [24]. HRO is based on studies on organizations that successfully handle complex technologies. The cost of failures in these organizations is unacceptable for the society. Main characteristics for HROs include managing complexity through (1) mindfulness, (2) the use of organizational redundancy, and (3) the ability to reconfigure spontaneously. Some theorists suggest that both the NAT and HRO theories oversimplify the cause of risk (e.g., [27, 28]). HRO theory underestimates the problems of uncertainty; conversely, NAT recognizes the challenges associated with uncertainty but underestimates the potential for strategic intervention.

2.4.3 Resilience Engineering

From the classical definition of safety as 'freedom of unacceptable risk,' through safety seen as a dynamic nonevent (HRO) to the ability to predict, plan, and act to sustain continuous safe operation, the RE School presents an alternative or supplementing perspective. Instead of focusing on failures, error counting, and decomposition, we should address the *capabilities to cope with the unforeseen*. The ambition is to 'engineer' tools or processes that help organizations to increase their ability to operate in a robust and flexible way [26].

A systemic view is encouraged to understand how the system as a whole dynamically adjusts and varies to continue safe operations. The focus is on the proactive side of safety management and the need to make proper adjustment for anticipation, updating of risk models, and effective use of resources.

While the majority of definitions focused on the capability to cope with failures providing a reactive approach, RE focused on the ability to adjust prior to or following a failure. RE explores ways that enhance the ability of the organizations to be robust and flexible and make the organizations prepared to cope with the unexpected.

The premises for RE are the following: (1) *performance variability is normal*, in the sense that performance is never stable in complex socio-technical systems and (2) many adverse events could contribute to a success or to a failure. These adverse events are the result of *adaptations* to cope with complexity; (3) safety management shall take into account both hindsight (reactive) and the ability of the system to make proper adjustments to anticipate potential threats, monitor risk, revise risk models, and to use resources proactively.

In this context, the qualities (i.e., anticipation, attention, and response, A-A-R) required for a system to be resilient are illustrated in Fig. 2.3. Adaptation is comprised of knowing what to expect (anticipation), knowing what to look for (attention), and knowing what to do (response). All the three factors (A-A-R) are continuously active. This state of alertness is made possible by constant updating of knowledge, competence, and resources [26].

2.4.4 Discussion and Future Implications

The sequential and epidemiological models have contributed to the understanding of accidents; however, they are not suitable to capture the complexities and dynamics of modern socio-technical systems (pros and cons are highlighted in Table 2.4).

In contrast to these approaches, systemic models view accidents as emergent phenomena, which arise due to the complex and nonlinear interactions among system components. These interactions and events are hard to understand, and it is not sufficient to comprehend accident causation by employing the standard techniques in risk analyses, i.e., by analyzing the failure modes of individual components using techniques such as FTA and FMECA, or relating the accident to a single causal factor. Since the standard techniques concentrate on component failures, they cannot adequately capture the dysfunctional interactions [1] between individual components operating without failure.

Rasmussen's framework has been comprehensively and independently tested on the analysis of two Canadian public health disasters [29] and on the Esso gas

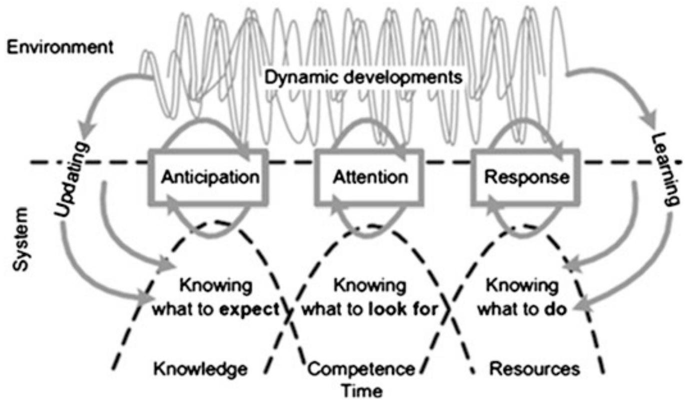


Fig. 2.3 Required qualities of a resilient system (adapted from [26])

plant explosion accident in Australia [30]. These case studies demonstrate the validity of Rasmussen's framework to explain the accident causation a posteriori. Further research is needed to extend this framework to predict accidents and to explore the applicability to risk and safety analysis of critical modern socio-technical systems.

NAT and HRO theories provide different explanations of safety and accident causation in complex organizations and offer alternative strategies for safety and risk management.

RE is emerging as a new paradigm in safety management, where 'success' is based on the ability of organizations, groups, and individuals to anticipate the changing shape of risk before failures and harm occur [26]. Complex systems exhibit dynamic behavior and continuously adapt their behavior to account for the environmental disturbances. Such system adaptations cannot be preprogrammed during system design [31].

In general, system-based theoretical approaches provide frameworks for modeling the technical, human, organizational, and environmental/contextual factors in modern socio-technical systems, including complex interactions among the system components. System-based accident models/theories will further contribute to capturing of future value potentials represented by the dynamic developments of ICT's now reshaping modern socio-technical systems. The combinations of Rasmussen's, NAT, HRO, and RE perspectives could strengthen the ability to discover and be prepared for unexpected situations in modern socio-technical systems such as the ICT-enabled integrated operations (IO) emerging in the petroleum activities in the North Sea.

Table 2.4 Summary comparison of the different approaches to accident prevention strategies

Theories/models	Scope	Benefits	Limitations
Domino theory[a]	Event-based accident model Static, deterministic	Works well for losses caused by failures of physical components or human errors in relatively simple systems	Cannot comprehensively explain accident causation in modern socio-technical systems
Formal methods (mathematically based techniques)[a]	Considers the human failure as an integral part of probabilistic safety/risk assessment (PSA/PRA)	PRA has become starting point for addressing human factors and led to the development of HRA Easy and convenient to use Identify key cause-consequence relationship	Human and organizational factors (HOF) cannot be adequately addressed
Swiss cheese model (SCM)[b]	Linear combination of active failures and latent conditions (to several event chains) Accidents were seen as occurring due to several events that coincide Static, deterministic	Look for the accident beyond the proximate causes, which is advantageous in the analysis of complex systems that may present multiple-causality situations Numerous organizational factors were identified as the casual factors that contributed to the probability of the accident	Difficult to describe and understand how multiple factors can come inline simultaneously to produce something as disastrous as a blowout Inadequate to capture the dynamics and nonlinear interactions between system components in complex socio-technical systems
Normal accidents theory (NAT)[c]	Defines two related dimensions of a system-interactive complexity and coupling	Made an important contribution in identifying these two risk-increasing systems characteristics Emphasizes the opportunity for interaction between (normal) faults	Unnecessarily pessimistic to possibility of effectively dealing with interactively complex and tightly coupled systems Hard to define links and interactions

(continued)

Table 2.4 (continued)

Theories/models	Scope	Benefits	Limitations
High-reliability organization (HRO)[c]	Prioritize safety Promote culture of reliability Use of organization learning Extensive use of redundancy	Continuous focus on finding small characters the 'unexpected,' the entire staff time to be vigilant Avoid simplifications	HRO oversimplifies the problems The measure of 'mindful operation' seems to require high-discipline organization
Rasmussen's socio-technical framework[c]	Safety as an emergent property impacted by decisions of all actors, at all levels, not just frontline workers alone Dynamic work practices	Attempt to model the dynamics of complex socio-technical systems	Raises the need for the identifications of the boundaries of safe operation, making these boundaries visible to the actors and giving opportunities to control behavior at the boundaries
Resilience engineering[c]	'Tractable' and 'intractable' Variability (both negative and positive)	RE points out that new knowledge is needed to further improve safety Continuous attention Focus on being proactive	It can be questioned whether RE is new paradigm within research

[a] sequential; [b] epidemiological; [c] systemic

2.5 Conclusion

Organizations today are stressed by a number of complex and dynamic factors in their environment. A key feature of complex systems is their 'emergent' behavior (emergent phenomena have no fixed position in time and space, and one cannot trace their causes directly). Rasmussen's, NAT, HRO, and RE have all been developed within and used in a context of complex high-risk socio-technical systems. Theories from such risk research domains are therefore important contributors to discourses on risk management approaches, as they invite to consider whether new models and approaches can supplement and improve current accident analysis approaches.

The emerging complexity of modern socio-technical systems (e.g., smart assets) poses a challenging area of multi-disciplinary research in the development of new safety/risk analysis and accident models involving researchers from various disciplines (i.e., founded on a concept of extensive cross-disciplinary knowledge sharing). Thus, there is a convincing need for researchers to step outside their traditional boundaries (i.e., researchers must work in an integrated and efficient manner across their geographical, organizational, and professional boundaries) in order to capture the complexity of modern socio-technical systems from a broad systemic view for understanding the multi-dimensional aspects of risk/safety and modeling socio-technical system accidents.

References

1. Leveson N (2004) A new accident model for engineering safer systems. Saf Sci 42(2004):237–270
2. Hollnagel E, Woods DD, Leveson N (2006) Resilience engineering: concepts and precepts. Ashgate, Aldershot
3. Hopkins A (2005) Safety, culture and risk: the organizational causes of disasters. CCH, Sydney
4. Kirwan B (2001) Coping with accelerating socio-technical systems. Saf Sci 37(2001):77–107
5. Perrow C (1984) Normal accidents: living with high-risk technologies. Basic Books, New York
6. Perrow C (1994) The limits of safety: the enhancements of a theory of accidents. J Contingencies Crisis Manag 2(4):212–220
7. Perrow C (1999) Y2K as a normal accident. Paper presented at international conference on disaster management and medical relief, Amsterdam, June 14–16
8. Rasmussen J (1986) Information processing and human–machine interaction: an approach to cognitive engineering. North-Holland, New York
9. Rasmussen J (1997) Risk management in a dynamic society: a modeling problem. Saf Sci 27(2):183–213
10. Reason J (1990) Human error. Cambridge University Press, Cambridge
11. Reason J (1997) Managing the risks of organizational accidents. Ashgate, England
12. Kirwan B (1994) A guide to practical human reliability assessment. Taylor & Francis, London
13. Bell BJ, Swain AD (1983) A procedure for conducting a human reliability analysis for nuclear power plants, NUREG/CR-2254, US Nuclear Regulatory Commission, Washington, DC
14. Hollnagel E (1998) Cognitive reliability and error analysis method (CREAM). Elsevier, Oxford
15. NASA Revisited (2006) Theory, analogy, and public sociology. Am J Sociol 112(2):353–393. The University of Chicago Press, Chicago
16. Vignes B, Liyanage LP (2011) Integrity in drilling wireline operations in offshore assets: a case on the potential error causation paradigms. Paper SPE138462 published at the SPE European health, safety and environment conference in oil and gas exploration and production, Vienna
17. Hollnagel E (2004) Barriers and accident prevention. Ashgate, Hampshire
18. Lewes GH (2005) Problems of life and mind. First series: the foundations of a creed, vol 2. University of Michigan Library Reprinting Series, Ann Arbor
19. Gibson JJ (1961) Contribution of experimental psychology to the formulation of the problem of safety: a brief for basic research. In: Behavior approaches to accident research. Association for the Crippled Children, New York, pp 77–89
20. Heinrich HW, Petersen D, Roos N (1980) Industrial accident prevention. McGraw-Hill, New York
21. Haddon W Jr (1980) The basic strategies for reducing damage from hazards of all kinds. Hazard Prevent 16:8–12
22. Turner BA (1978) Man-made disasters. Wykeham, London

23. Rasmussen J, Svedung X (2000) Proactive risk management in a dynamic society. Swedish Rescue Services Agency, Karl
24. LaPorte TR, Consolini PM (1991) Working in practice but not in theory: theoretical challenges of “high-reliability organizations”. J Public Adm Res Theory (J-PART) 1(1):19–48
25. Rasmussen J, Jensen A (1974) Mental procedures in real-life tasks: a case study of electronic troubleshooting. Ergonomics 17(1974):293–307
26. Hollnagel E, Pariès J, Woods DD, Wreathall J (2011) Resilience engineering in practice: a guidebook. Ashgate, Aldershot
27. Sagan S (1993) Limits of safety: organizations, accidents, and nuclear weapons. Princeton University Press, Princeton
28. Marais K, Dulac N, Leveson N (2004) Beyond normal accidents and high reliability organizations: The need for an alternative approach to safety in complex systems. http://esd.mit.edu/symposium/pdfs/papers/marais-b.pdf. Accessed 20 Jun 2012
29. Woo DM, Vicente KJ (2003) Socio-technical systems, risk management, and public health: comparing the North Battleford and Walkerton outbreaks. Reliab Eng Syst Safety 80:253–269
30. Hopkins A (2000) Lessons from Longford: the Esso gas plant explosion. CCH, Sydney
31. Hollnagel E (2006) Resilience—the challenge of the unstable. In: Hollnagel E, Woods DD, Leveson N (eds) Resilience engineering: concepts and precepts. Ashgate, Aldershot, pp 629–634
32. Stringfellow MV (2010) Accident analysis and hazard analysis for human and organizational factors. Massachusetts Institute of Technology, Massachusetts

Chapter 3
On Engineering Assets and Technology Supply Capacity

Joe Amadi-Echendu and Nomsesi Martha Mhlanga

Abstract Capital development of an oil and gas refinery involves the establishment of a large base of engineering assets. In developing countries, the question arises as to the capacity of local firms to supply the required technological capabilities while operating under regimes that vibrantly promote indigenization and empowerment legislations. By analyzing the capability of local vendors, this paper provides some insight into the perplexing challenges presented by the need to create and build local science, technology, and innovation (STI) capacity. Despite a vigorous empowerment regime, the local market more or less features agents, distributors, and vendors of sources of technological capabilities that are located well outside the sociopolitical environment of the refinery project.

3.1 Introduction

The oil and gas exploration, production, refining, and distribution value chain involves a wide range of tangible and intangible technological capabilities that may be defined in terms of intellectual capital, knowledge, skills, and sophisticated engineering assets. For brevity, a high-level classification of the engineering assets includes drilling rigs, production platforms, refineries, and distribution depots, and these in turn, comprise several tiered levels of classification in terms of infrastructure, specialized processing units, static and rotating equipment, control, instrumentation and telecommunication, and piping systems. The engineering assets may be further resolved into various levels of equipment and components plus their associated technologies. In this regard, the development of a major engineering asset such as an oil refinery involves a wide range of sources

J. Amadi-Echendu (✉) · N. M. Mhlanga
Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Pretoria, South Africa
e-mail: joe.amadi-echendu@up.ac.za

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_3,

of knowledge and skills, designers, constructors, manufacturers and suppliers of hardware and software components, equipment, infrastructure, process technologies, services, and systems. The globalization of supply chains means that the sources of the technological capabilities required may be domiciled outside the economic, social, and political jurisdiction of the geographical location of the refinery. This is particularly important in a developing country context, where empowerment issues tend to take on more relevance within the mix of economic, environmental, political, social, and technology considerations for the development of a major capital project such as an oil refinery.

Although empowerment, that is, the process of enabling maximum individual and/or collective access to authority, influence, and power (see [1]), is an issue in many jurisdictions, however, in some developing African countries (e.g., Nigeria, South Africa, and Zimbabwe), the significance of empowerment is often stipulated in terms of legislative provisions that are generally geared toward indigenization and local participation in economic activities. For these countries, it has become an imperative to promote knowledge and skills acquisition and capacity building so as to enhance economic development and social emancipation. Whereas it is intuitively obvious that global supply chains have wide ramifications on indigenization and empowerment legislations, however, the local challenge is how to harness the extensive sources of the technological capabilities required for the capital development of an oil refinery.

This paper briefly examines the local capacity for the supply of equipment and components necessary for establishing a technology-intensive engineering asset such as an oil refinery, in a manner that is consistent with indigenization sentiments and compliance with empowerment legislation in a developing country.

3.2 Capacity Challenges

The proceedings of the 2008 Global Forum on STI Capacity Building for Sustainable Growth and Poverty Reduction convened by The International Bank for Reconstruction and Development/The World Bank (see [2]) indicate a general consensus that the globalization impetus means that "…science, technology and innovation (STI) capacity building…is an absolute necessity" for developing countries. This view is summarized in Fig. 3.1, depicting five dimensions for STI capacity building.

A cursory review (see [3]) suggests that legislation that promote indigenization and empowerment in developing African countries (e.g., Nigeria, South Africa, and Zimbabwe) are intended to address the five dimensions for STI capacity building. At least, three of the STI capacity dimensions are pertinent to and thus may be regarded as vital requirements for the development of an oil refinery located in a developing country viz:

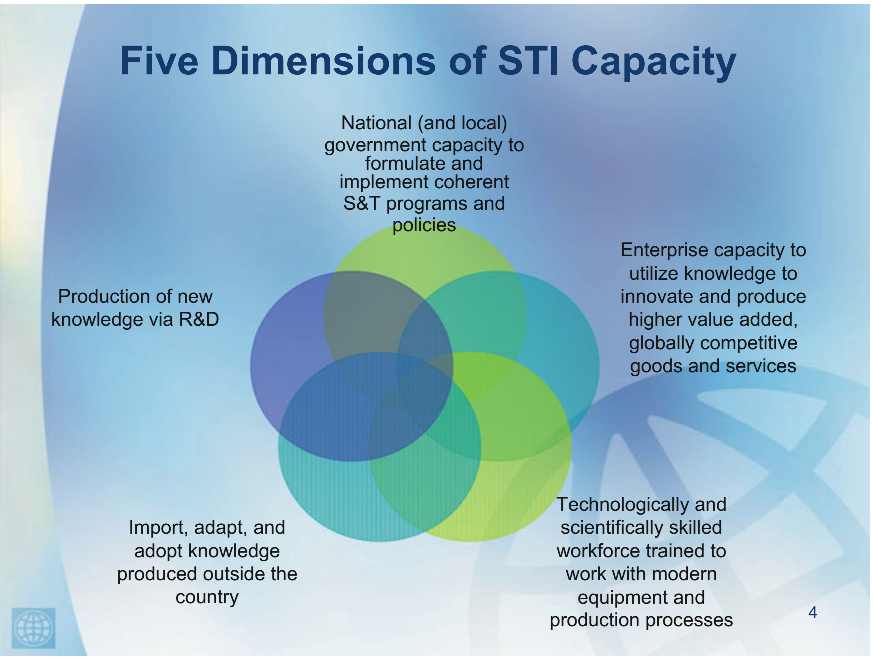


Fig. 3.1 Five dimensions for STI capacity building

1. The importation, adaptation, and adoption of applicable *oil refining* knowledge produced outside the country;
2. The availability of "technologically and scientifically skilled workforce trained to work with modern equipment, processes, and *systems for oil refining*";
3. The existence of "enterprise capacity to utilize knowledge to innovate and *develop* higher value added, globally competitive" *oil refining capability*.

With respect to the supply chain, the first STI capacity challenge is the availability of local sources for the technological capabilities. That is, how many manufacturers and vendors are locally available to supply the knowledge, skills and engineering asset components, equipment, infrastructure, and systems required for developing a new refinery in our case study country?

The study published by Ren [4] does relate to dimensions (1) and (2) of STI capacity above, even though the discussion focuses on technological capacity with regard to the perplexing imperatives of innovation and ecological footprint. As illustrated in Fig. 3.2, the paper highlights from a sustainability viewpoint that there are interrelationships between "innovators and networks for innovation in petrochemical processes." This can be regarded as a depiction of the scope of technological capacity in terms of knowledge acquisition, absorption, transfer, and application. With regard to the development of a major engineering asset such as an oil refinery, the innovators can be interpreted as sources while the network is the recipient of the knowledge. Thus, the second STI capacity challenge is whether or not there is a local innovation network to accentuate the know-how required to develop a globally competitive oil refinery.

In developing countries, the existence of an enterprise capacity may be constrained by *macro*-, *meso*-, and *micro*-level factors as classified in Forss et al. [5] and summarized in Fig. 3.3. In the context of this paper, the *micro-level actors, competence, and capability* more or less reiterate the intent in indigenization and empowerment legislation in developing countries. This is complementary to the *physical infrastructure condition* at the *meso*-level and further constrained by *institutional market regulation* at the *macro*-level. Hence, the third STI capacity challenge at the *micro*-level is market access, i.e., whether or not a locally competitive market exists for the supply of knowledge, skills and engineering asset components, equipment, infrastructure, and systems required to build the refinery.

The engineering asset schematic layout for a typical oil refinery (see [6]) illustrated in Fig. 3.4 shows a number of processing units, each comprising its own set of technological capabilities. For the purposes of our study, we identify two categories of technological capabilities as follows:

1. Tangible technological capabilities—that is, engineering asset components such as equipment, infrastructure, and firmware systems; hence, from a supplier market stance, we view equipment, infrastructure components, and firmware systems as *products*.
2. Intangible technological capabilities—that is, know-how that combines knowledge, skills, and organizational flair. Again, from a supplier market stance, we view the intangible technological capabilities as *services*.

Based on our categorization of technological capabilities, our study then examines the first and third STI challenges within a developing country regime that vigorously promotes empowerment legislation. Instead of evaluating its impact, we have rather presumed that the empowerment legislation influences the supplier market, thus the rest of this paper describes an analysis of the availability of a locally competitive market for suppliers of products and services required for the development of a major oil refinery in our developing country location.

3.3 Research

The conceptual model illustrated in Fig. 3.5 is a hierarchical depiction of the oil and gas business from a supply market view of the associated technological capabilities.

For the particular refinery development latently referred to in our study, the applicable empowerment legislation stipulated that 30 % of the services and products procured for the project must be from local suppliers. In order to assess the availability of a competitive local market for the required technological capabilities, we carried out a search for manufacturers, suppliers, and vendors

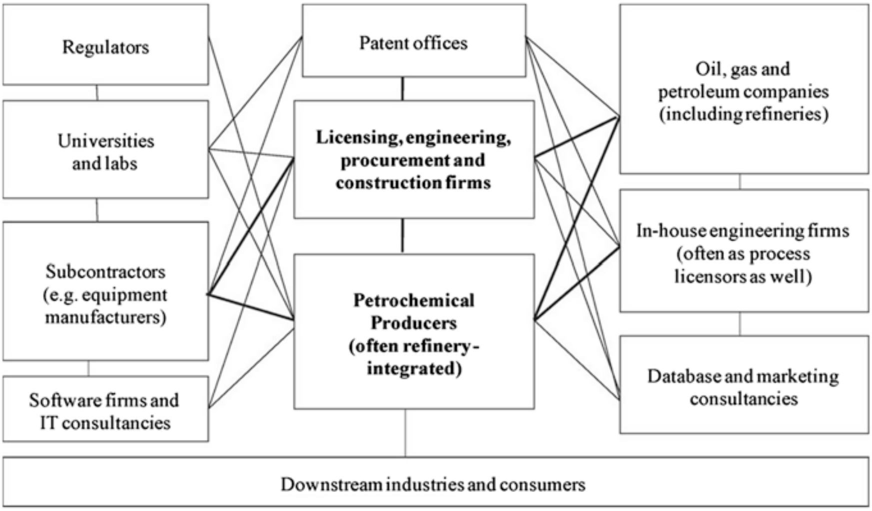


Fig. 3.2 Innovation networks in petrochemical processes (*Source* [4])

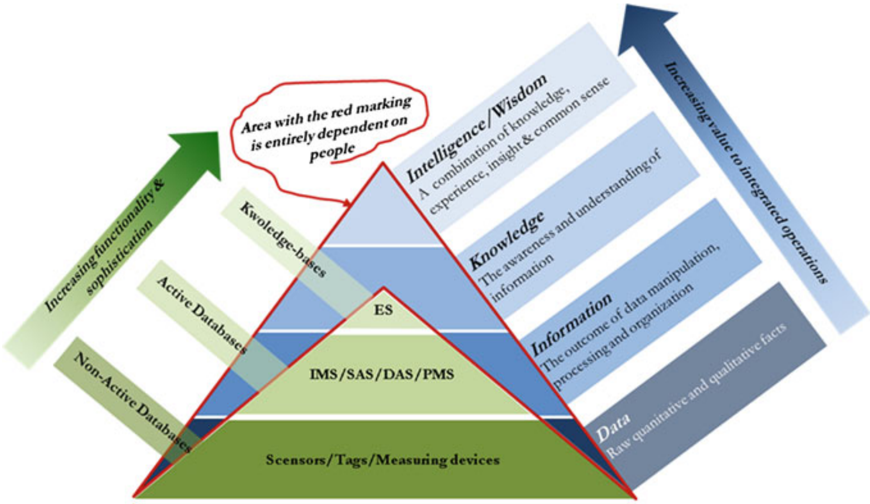


Fig. 3.3 Factors influencing the development of an oil refining enterprise (*Source* [5])

of infrastructure components, equipment, and systems specified in the engineering design of the refinery. Our research data about suppliers of equipment and refinery process technologies were obtained from publicly available databases of companies, plus vendor lists of two local petrochemical companies, and a transnational engineering design firm. Based on our understanding of the technological

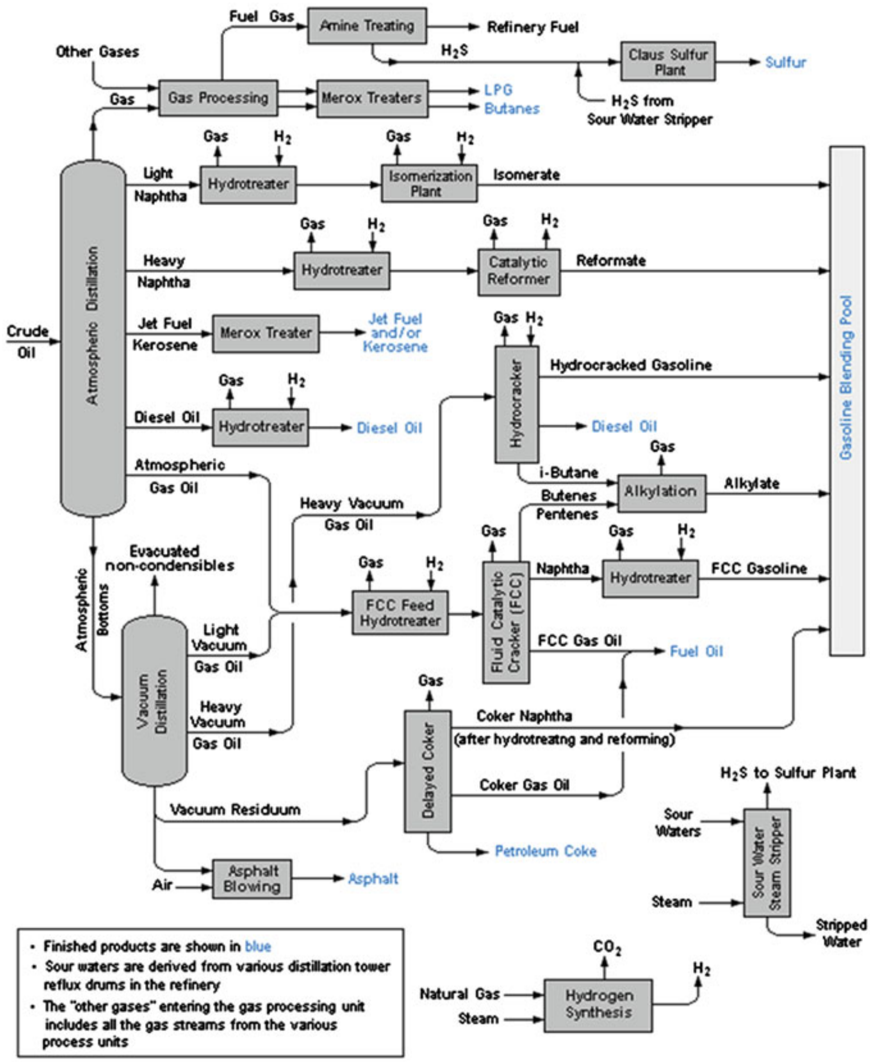


Fig. 3.4 Schematic flow diagram of an oil refinery (*Source* [7])

capabilities specified in engineering design documents of the new refinery, the search for suppliers resulted in the data arranged as shown in Fig. 3.6.

It is remarkable that the data shown in Fig. 3.6 may be biased in the sense that a count was recorded for each instance that a supplier was cited against the G-coded grouping of the technological capabilities. Thus, the count of the number of suppliers may be bloated considering that a supplier may supply items in a number of the groupings. In this regard, each G-code group was resolved into subgroups of engineering asset classes, equipment types, firmware systems, and services. This

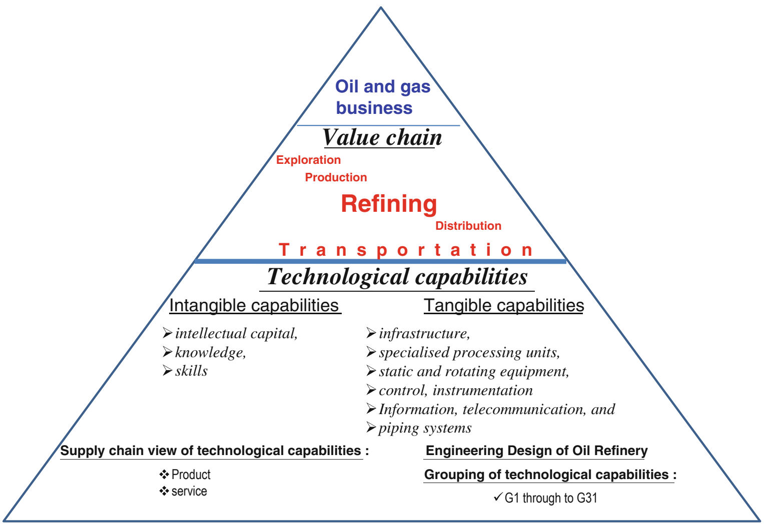


Fig. 3.5 A conceptual hierarchical model of oil and gas technological capabilities

Classification of technological capabilities	Number of Suppliers identified
G.1. FURNACES	19
G.2. EXCHANGERS	154
G.3.PRESSURE VESSELS	138
G.4. ROTATING EQUIPMENT - PUMPS	426
G.5. SPECIAL EQUIPMENT	285
G.6. SOLID SPROCESSING EQUIPMENT	40
G.7. UTILITY EQUIPMENT	40
G.8. PIPING SYSTEMS	968
G.9. STRUCTURAL	86
G.10. ELECTRICAL	603
G.11. INSTRUMENTATION AND CONTROLS	839
G.12. FIRE AND SAFETY	221
G.13. CHEMICAL AND CATALYST	28
G.14. COATINGS	64
G.15. CIVIL SUBCONTRACTS	245
G.16. FIELD SUBCONTRACTS	230
G.17. FOOD AND ACCOMODATION	64
G.18. EQUIPMENT RENTAL	306
G.19. FUEL/LUBRICATION	45
G.20. SITE RECLAIMATION	196
G.21. ADMINISTRATION/HR/SECURITY	54
G.22. COMMODITIES/CONSUMABLES	151
G.23. BUSINESS SERVICES	48
G.24. CONSTRUCTION SERVICES & TOOLS	39
G.25. OFFICE SERVICES	65
G.26. MEDICAL SERVICES	13
G.27.TECHNICAL SERVICES	93
G.28. THIRD PARTY INSPECTION	13
G.29. EPC&M	34

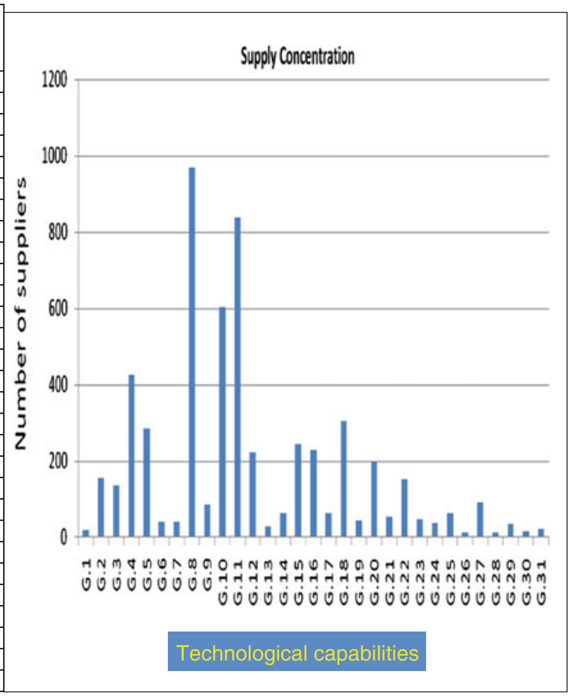


Fig. 3.6 Number of suppliers identified for technological capabilities required for our case study refinery

resulted in 355 subgroups of technological capabilities. We also augmented our data collection scope through a survey that was sent to suppliers that had valid local contact details. Regrettably, only 22 out of the 198 suppliers contacted provided feedback. In addition to basic information about location and nature of business, suppliers were asked to indicate their primary operational status, i.e., agency, consulting, contracting, distribution, or manufacturing. Furthermore, suppliers were asked to indicate the range of items in the identified subgroups that they can supply locally, irrespective of the source of the technological capabilities.

The feedback obtained revealed that 15 out of the 22 responding suppliers are located in one geopolitical region of the developing country, even though the refinery is to be cited about 900 km away. The feedback shows that 14 suppliers indicated that they were original equipment manufacturers and six of whom also operated as agents, consultants, or distributors. Three of the 22 responding suppliers operated purely as agents/distributors, only one supplier operated as a contractor, while the remaining four operated as distributors/consultants.

Figure 3.7 illustrates the range and scope of technological capabilities that can be supplied by the suppliers that responded to our survey. In both graphs, the vertical axis is a nominal scale that represents range and scope, respectively.

In comparison to Fig. 3.6, it is evident from Fig. 3.7 that many of the technological capabilities required for the refinery project could not be provided by local suppliers. In fact, from the upper graph in Fig. 3.7, the respondents indicated that they could supply items (within some of the 355 subgroups) in 16 of the 31 major groupings. Also, the lower graph shows that only five suppliers could supply items in four or more subgroups of the technological capabilities, about seven suppliers could supply items in two or three subgroups while the remaining ten respondents could supply items in one or two subgroups. Thus, the combination of range and scope is well below the supply capacity requirements for the refinery project.

3.4 Discussion

Data obtained from our research suggests that there is low availability of local suppliers of the requisite technological capabilities for the development of the oil refinery. From our database searches, we identified more than 200 suppliers each for at least ten of the major (G-code) groupings and more than 20 suppliers each for 29 of the 31 G-code capabilities. Even if the 198 suppliers that had valid local contact details had responded to our survey, the reality is that this number still represents low availability of the supply market for the 31 major groups of technological capabilities required. In terms of the first STI challenge, the implication is that this represents a major issue for the case study country, particularly with regard to indigenization sentiments, compliance with, and the enforcement of the provisions in the empowerment legislation, albeit the stipulation is that 30 % of the services and products procured for the specific refinery project must be from local suppliers.

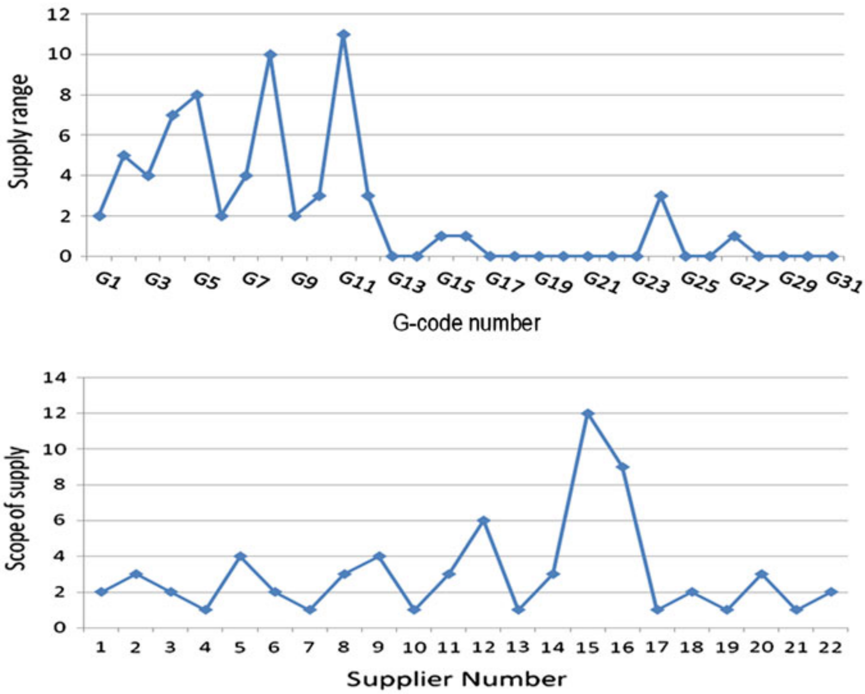


Fig. 3.7 Local supply range and scope of technological capabilities

A similar inference holds for the third STI challenge with regard to the existence of a competitive supply market for the technological capabilities that are locally available. First, it calls to reason that many of the 198 locally listed suppliers may only be operating as agents or distributors of technological capabilities which are sourced outside our case study country. This view is supported by the fact that only five of the 22 respondents indicated that they operated solely as manufacturers, albeit of a very narrow range of G-code groups of technological capabilities. What this also suggests is that the remaining 17 respondents source most of their technological capabilities outside of the country. It further gives some credibility to the view that those suppliers that claim that they manufacture locally may actually be sourcing the intangible aspects of the technological capabilities outside of the case study country. With primary data provided by only 22 respondents, it is remarkable that coupling the narrow supply range with the very limited scope of supply portends a weak local market that may not be competitive. Again, this begs to question what type of impact that indigenization sentiments and empowerment legislation may have on the supply of technological capabilities for the development of the engineering assets that constitute the oil refinery.

The findings of low availability and uncompetitive local supply capacity implicitly suggest that most of the technological capabilities are sourced externally beyond the confines of indigenization campaigns and empowerment legislation. Thus, another aspect for further investigation to complement our continued research is the issue of technology transfer as embodied in the second STI challenge—availability of "technologically and scientifically skilled workforce trained to work with modern equipment, processes and…" systems applicable for oil refining.

References

1. Page N, Czuba CE (1999) Empowerment: what is it? J Extension 37:5
2. Watkins AJ, Ehst M (2008) Science, technology, and innovation: capacity building for sustainable growth and poverty reduction. A World Bank publication, Washington, DC
3. Dunmade I (2002) Indicators of sustainability: assessing the suitability of a foreign technology for a developing economy. Technol Soc 24:461–471
4. Ren T (2009) Barriers and drivers for process innovation in the petrochemical industry: A case study. J Eng Technol Manage 26:285–304
5. Forss K, Lundström M, Saasa O, Temu F (2003) Enterprise development programmes in Tanzania and Zambia. Department for Finance and Corporate Development. Sida Evaluation 03/36. Swedish International Development Cooperation Agency, Stockholm
6. Salary R, Jafari Nasr MR, Amidpour M, Kamalinejad M (2008) Design of oil refineries hydrogen network using process integration principles. Iran J Chem Chem Eng 27:4
7. http://en.wikipedia.org/wiki/Oil_refinery

Chapter 4
Asset Operations: Non-productive Times During Oil Well Drilling

Joe Amadi-Echendu and Audu Enoch Yakubu

Abstract Weather, well bore characteristics, tools and equipment failures, and other sources of risk aggregate into non-productive times (NPT) which, in turn, adversely impact on the efficiency and costs of drilling crude oil wells. The imperative is to increase efficiency and reduce the cost of drilling operations, especially for offshore locations where the high temperature and pressure conditions mean that the window between reservoir pore pressures and fracture gradients can be quite narrow. By analyzing data on drilling rigs from three operators, the paper provides insight into some of the prevalent reasons why NPT account for more than 30 % of the budget for oil well construction in the geographical area under study.

4.1 Introduction

Non-productive times (NPT) that arise while drilling for crude oil can result in the escalation of well construction costs, especially in offshore locations. NPT may be defined as an unplanned event that prolongs the drilling operations schedule [1]. Delays encountered during oil well construction can increase the risk of overexposure of drilling fluids such as water-based mud (WBM) or synthetic oil-based mud (SOBM). Overexposure of drilling fluids can cause well bore damage in the rock formations, and this, in turn, can reduce reservoir throughput [2]. The interaction between the drilling mud and the rock formations can result in the onset of irrecoverable instability, and in the worst case, may ultimately result in the loss of the well.

J. Amadi-Echendu (✉) · A. E. Yakubu
Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Pretoria, South Africa
e-mail: joe.amadi-echendu@up.ac.za

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_4,

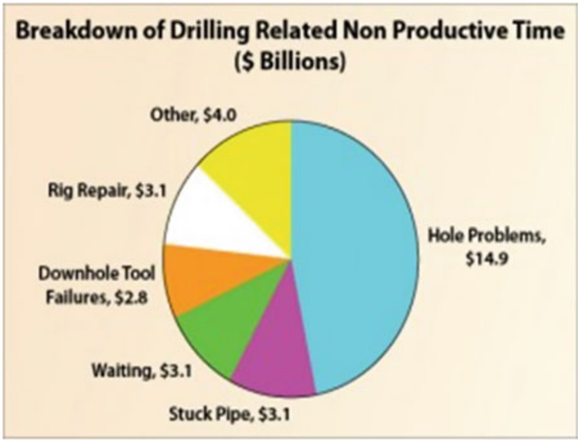


Fig. 4.1 Estimated NPT cost elements for 2010 (*Source* [4])

Although delays are expected in drilling operations, however, the cumulative effect of NPT over a program of drilling projects may adversely influence the number of wells that can be drilled in a given time frame. For a marginal oil field, NPT may reduce the economic viability as well as result in a huge cost to the environment with devastating consequences [3]. Thus, eliminating NPT provides opportunity for considerable cost savings. Sources of NPT include, for example, downhole tool failures, rig repairs, waiting on weather, pulling of dulled bits, running and cementing casings, and wireline logging. Figure 4.1 shows a breakdown of estimated costs from sources of NPT [4].

4.2 Sources of NPT in Drilling Operations

Figure 4.2 shows equipment and infrastructure typically deployed for drilling operations. NPT accumulates from delays arising from equipment and/or accessory malfunction due to hazards encountered when the rig machinery is commissioned for the construction of an oil well.

An outline workflow for drilling operations is depicted in Fig. 4.3. This includes:

1. modeling the formation lithology and prediction of well trajectory;
2. acquisition, installation, and commissioning of rig equipment;
3. drilling, data collection, and progress monitoring.

Although NPT may be experienced during each stage of drilling operations, however, a major source of risk and cause of NPT are vibrations that occur when the drillstring (pipe, collar, and bit) interacts with the rock formations as drilling progresses. Real-time monitoring and evaluation of drillstring vibrations are being applied to mitigate such risks [5].

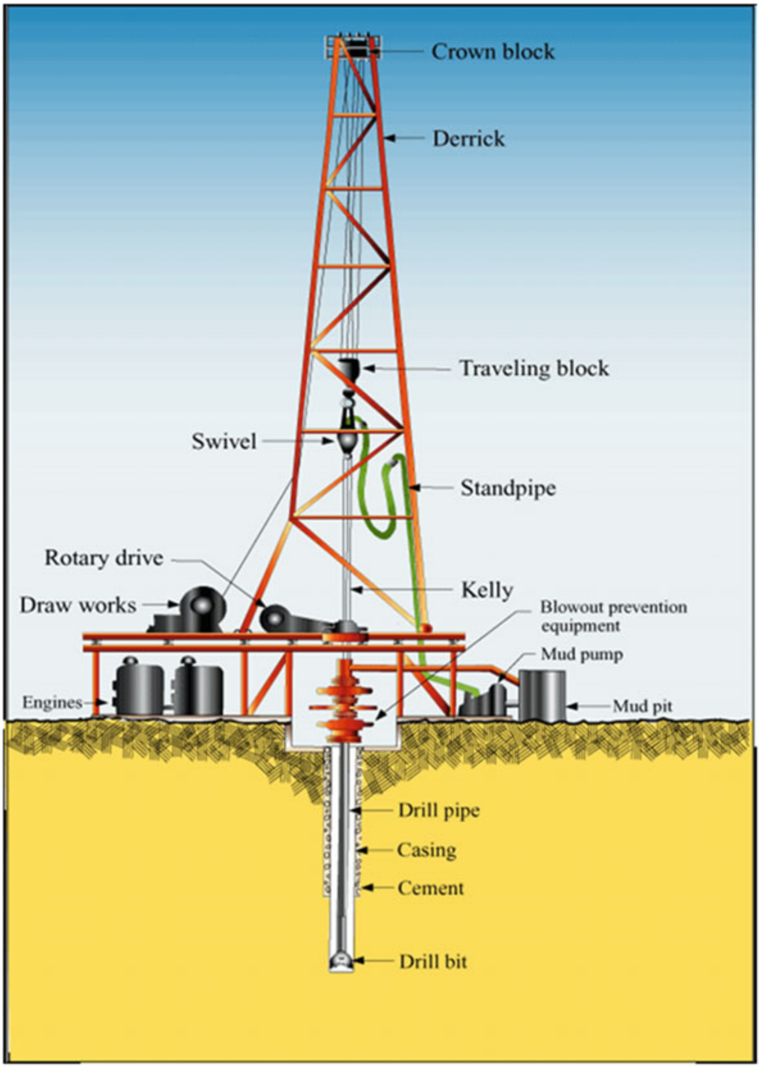


Fig. 4.2 Drilling rig schematics (*Sources* www.conservation.ca.gov; http://visual.miriam-webster.com)

Other typical sources of risk and causes of NPT are summarized in a set of benchmark metrics [6] listed in Table 4.1. It is remarkable that particular sources of risk may prevail over others in different geographical regions of oil and gas resource locations. Therefore, approaches for mitigating NPT risks vary depending on the prevalent sources. In the Norwegian continental shelf, Andersen et al. [7] examined 13,700 incidents of equipment failures between 1985 and 1991 to establish the relationship between failure modes and NPT. Top drive, drillstring,

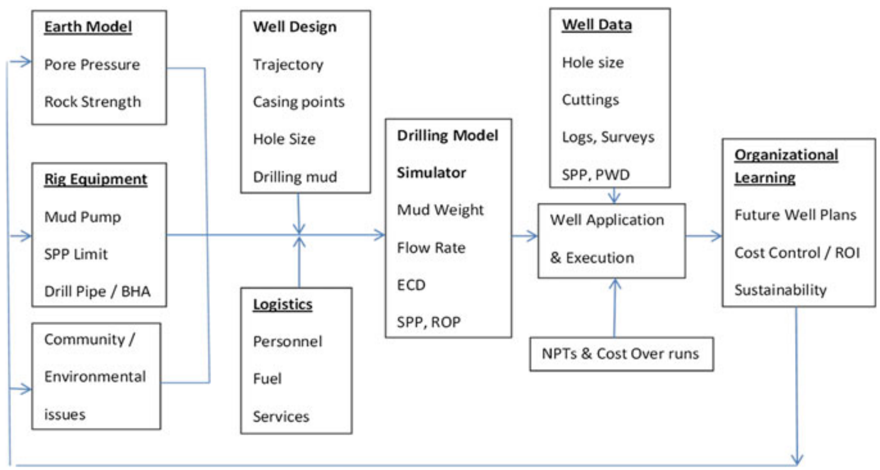


Fig. 4.3 Drilling operations summary

Table 4.1 Sources of NPT (*Source* [6])

Stuck pipe	Twist off	Kick	Directional correction
Lost circulation	Sloughing shale	Weather delay	Cement squeeze
Well bore instability	Equipment failure	Rig failure	Mud/chemical
Shallow gas flow	Shallow water flow	vibration	Casing or wellhead failure

and blowout preventer failures were prominent contributors to NPT and focal to continuous improvement initiatives.

4.3 Research

This paper examines sources of NPT for three drilling companies operating in the Niger Delta. We obtained data regarding NPT by selecting oil well construction projects from each of the three drilling companies that participated in our study. The total number of oil wells available from the three companies is 962, but we obtained information about 135 wells evenly located on land, swamp, and offshore environments. This includes data from both drilling and financial departments of the companies. For brevity, we focus here on the dominant causes and effects of NPT as identified from the data we collected.

Figure 4.4 is a summary of the dominant sources of NPT and the relationship between NPT and “red money”, that is, additional costs committed to the drilling operation as a result of NPT. For the particular region under study, data arising from the 135 wells indicate that the dominant causes of NPT may be grouped under

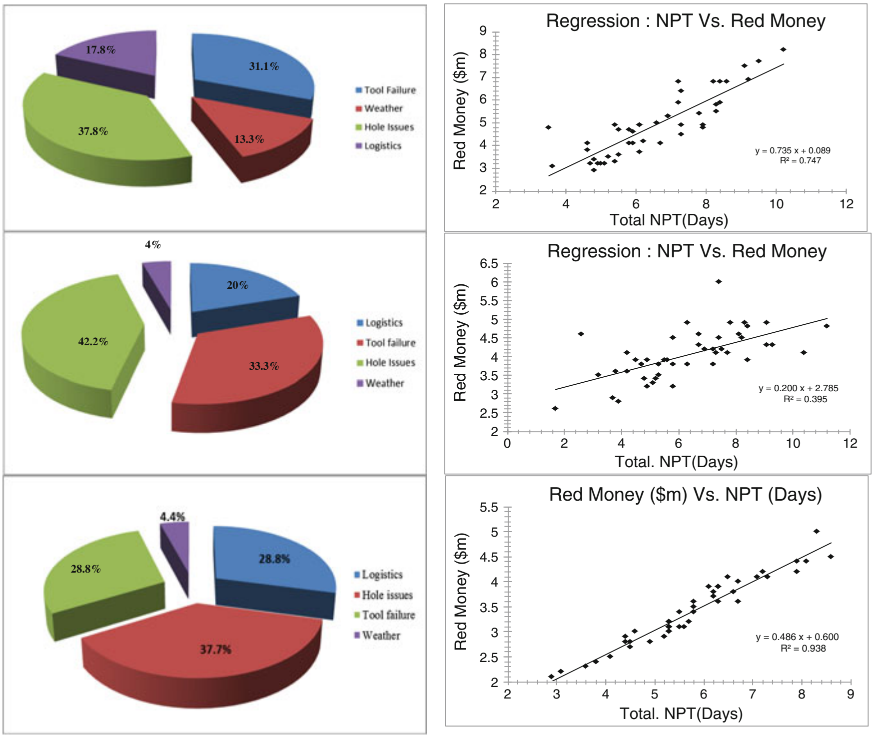


Fig. 4.4 Sources of NPT and effects of NPT on drilling costs

1. logistics issues,
2. hole issues,
3. tool failure, and
4. weather effects.

Our data show that between these causes, weather appears to have the least impact on NPT. It is perplexing that the effect of weather on NPT for company X is more than double that for the other two companies, albeit that all the wells are located in the same geographical region and within a radius not exceeding 300 miles. It is not surprising that logistics issues have a noticeable impact on NPT since associated infrastructure is very weak in the geographical region. It is interesting that tool failures appear to be very significant contributors while hole issues are dominant causes of NPT for the 135 oil wells studied.

The results we obtained concur with the view that NPT increases costs of drilling operations beyond the initial approved expenditure. Subject to the accuracy of the original data recorded, information obtained from each of the three companies supports the view that there is a considerable degree of correlation between NPT and red money.

4.4 Discussion

Identification of the dominant causes of NPT and its correlation with red money present opportunities for continuous improvement and possible reductions in oil well construction costs for the three operators in our study. The dominance of hole issues and tool failures means that greater attention needs to be placed on obtaining relevant, valid, accurate, and sufficient geomechanical data. With more wells being drilled in the geographical region, analysis of a large database of geomechanical data should provide the means to describe an optimum range of drilling parameters applicable for the region such as mud weight, rheology, and flow rate weight on bit, drill speed, and circulating pressures.

A further analysis of the types of tool failures provides opportunities for improving the relationship between drilling operators and their associated service providers. Advances in high resolution 3-D seismic and 4-D time-lapse imaging is making it possible to drill directional wells with tighter well path control to reduce environmental impact and ecological footprint. Mitigating the operational and financial risk associated with tool failures requires concerted efforts and synergy between the operators and service companies involved in drilling operations for the construction of oil wells.

References

1. Reid D, Rosenberg S, Montgomery M, Sutherland M, York P (2006) Drilling hazard mitigation and high value well construction techniques. In: Offshore technology conference, OTC 18084. Houston, Texas, USA, 1–6 May 2006, pp 1–4
2. Hong W, Soliman MY, Towler FB (2008) Investigation of factors for strengthening a wellbore by propping fractures. In: Drilling conference, SPE 112629. Orlando Florida USA, 4–6 Mar 2008, pp 1–3
3. McAndrews KL (2011) Consequences of Macondo: A summary of recently proposed and enacted changes to US offshore drilling safety and environmental regulation. In: SPE Americas E & P health, safety, security and environmental conference. Houston Texas USA, SPE143718-PP, 21–23 Mar 2011, pp 1–2
4. Hamayun R, Rich D (2011) Enhanced use of real-time centers can optimize well planning, drilling and geosteering. The drilling contractor magazine, March/April 2011 (issue)
5. Craig DA, Hanley MC, MacFarland B, Shearer D, King P (2009) A proven approach to mitigating drilling vibration problems in offshore Western Australia. In: International petroleum technology conference, IPTC 13399. Doha, Qatar, 7–9 Dec 2009
6. Dodson JK (2010) Drilling performance benchmark metrics. Dodson Data Systems.com
7. Andersen LB, Lembourn PR (1994) Analysis of interruption and lost rig time due to equipment failure in offshore drilling operations. In: Drilling conference, SPE/IADC 27511. Dallas, Texas USA, 15–18 Feb 1994, pp 1–4

Chapter 5
More Reliable Profitable and Competitive Production via Effective e-Maintenance: A Case Study

Basim Al-Najjar

Abstract Profitable production demands continuously improved maintenance decision accuracy for reducing/avoiding failures and unplanned stoppages due to their consequences. More accurate decisions prolong life length of components, and consequently machines, and maintain the production running longer. When a condition-monitoring (CM) value exceeds a significant (warning) level, it demands a clear understanding of what happened and how it will develop in the next future to avoid failures. In addition to CM it demands reliable information concerning the probability of failure, residual lifetime, and when is the most profitable time of conducting maintenance. Companies strive to reduce production cost in order to increase the possibility of offering customers lower prices and generating additional competitive advantages. But applying new technologies for enhancing maintenance, production performances, and company competitiveness counters many problems in industry. In this paper, a new innovative e-maintenance decision support system (eMDSS) is introduced; the problems facing successful implementation of eMDSS based on a case study are introduced and discussed. Solutions to avoid the problems facing successful implementation of eMDSS are suggested and discussed. eMDSS offers a unique opportunity to achieve just in time dynamic and cost-effective maintenance by selecting the most profitable time for maintenance.

It offers innovative solutions to increase maintenance profitability by enhancing maintenance decision accuracy via:

1. Identifying and classifying problems based on their shares in production losses.
2. Predicting the value of CM parameter, e.g., vibration, level in the near future.
3. Estimating the probability of failure and component remaining life in the near future.
4. Estimating the most profitable time for maintenance action.

B. Al-Najjar (✉)
School of Engineering, Linnaeus University, 351 95 Växjö, Sweden
e-mail: basim.al-najjar@lnu.se

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_5,

5. Selecting the most profitable maintenance action.
6. Mapping and following up maintenance and production performances.
7. Identifying and following up the most profitable investment in maintenance.

eMDSS combines technical and economic data describing future situation in addition to the current and past data, which is necessary for production and maintenance successful planning. The major conclusion is that applying eMDSS is possible to reduce failures appreciably, prolong life length of components/equipment, and perform profitable maintenance.

5.1 Introduction

It is necessary for production and maintenance managers that maintenance should be dynamic to be suited for the changes in the production process cost-effectively and to enable with high decision accuracy achieving just in time maintenance [1]. One way to do that is to enhance the accuracy of maintenance decisions so that we can select the most profitable time for conducting maintenance actions. Maintenance is, in many cases, considered as a necessary evil, while it instead can be considered as a profit-generating center since it is closely related to company internal efficiency [2]. According to [3], significant savings (up to 20 %) can be achieved in a company by reducing losses in the production process. Also, according to the [4], the maintenance costs represent the next largest after energy costs.

Many publications (see e.g., [5–7]) emphasis that many software modules and tools are available and utilized for developing knowledge bases which can be used for fault diagnosis. In [8], the authors claim that combined production and maintenance models lead to significantly greater reward (about 25 %) compared to a traditional method. Using an efficient maintenance policy, such as vibration-based maintenance (VBM) policy for rotating machines, leads to fewer planned stoppages and failures, lower level of spare part inventory, and a smoother production process. It will also lead to a higher-quality and more profitable production process especially in process, chemical, energy, and recently in manufacturing engineering industry [2]. To be able to monitor, map, analyze, assess, predict, and improve the outcome of different maintenance actions properly, it is necessary to gather and use the data covering relevant disciplines, i.e., technical, economic, and managerial data. Data gathering and analysis processes can even be more trouble-free and cost-effective if it is computerized in a form such as decision support system (DSS). This allows following up production and maintenance performance more frequently.

Therefore, planning and executing maintenance actions cost-effectively, and answering where, how much and why an investment in maintenance may have the best economic payoff can easily be achieved [9]. In general, Maintenance actions can be planned and performed with respect to three concepts and strategies:

- only at failure to restore a machine to as good as before,
- at regular time does not matter whether they are needed or not, or
- based on the condition of the machine which means do nothing until it is required, i.e., do maintenance when there are symptoms showing damage initiation and development to avoid failures. Performing maintenance on demand reduces unnecessary stoppages, prolongs production time, and reduces economic losses.

5.2 Problems Facing Implementation of eMDSS

A case study was conducted in Sweden, and it consisted of four companies covering different branches, e.g., engineering manufacturing, production machine manufacturer, and paper and pulp manufacturer. The major focus of this case study was to identify the potential problems that may counter a successful implementation of DSS. We consider one e-maintenance decision support system (eMDSS). It is standalone system which consists of four modules that can operate individually and also can be integrated for more synergies. The final results of this case study can be summarized by the following:

1. Interface problems concerning eMDSS-communications with the current databases, software programs, and systems. This problem is general and not specific to a particular DSS.
2. Data gathered and processed by eMDSS consist of technical and economic data which make it difficult for a company to accept it as a Web-based service due to data security concern. This is why it is more preferable having eMDSS as a standalone system separated from Internet to increase data security especially economic information.
3. In some cases, we found that the company does not apply a full scale of condition-based maintenance (CBM) strategy and this generates difficulties when applying eMDSS.
4. One of these serious shortages in companies' applications is they did not always acquire reasonable CM levels describing normal operation, potential failure (when damage is initiated and under development), and imminent failure level which is required to start planning maintenance actions.
5. The data in the company are distributed in several databases, and in order to access these data, the company demands to switch between software programs, computer systems, and maybe buildings. It makes data accessibility and integration not easy especially when more than one department in the company is involved.
6. In some cases, we found that the CM software program is not used properly due to the lack of knowledge and experience. For example, they cannot reset new vibration alerts, or conduct reliable diagnosis or interpretation of the vibration signals/spectrum.

7. Outsourcing of maintenance and CM measurement analysis increases the difficulties of applying eMDSS because data, experience, and databases are distributed between two or more companies and make it very difficult to access.
8. Maintenance in many companies is considered as cost-center which makes new investments for enhancing maintenance decision accuracy something unnecessary especially when the company tries to reduce maintenance cost.

5.3 Solution to Ease Implementation of eMDSS

During last decades, companies have had a problem of losing knowledge and experience due to many reasons, such those reasons related to personnel:

1. Personnel leave the company to a competitor, or
2. Retired or die.

To make the investment in eMDSS acceptable and its installation problem-free, we considered the following solutions:

1. eMDSS can work standalone manually with its database which eliminates the need to be integrated with the current software systems and databases if it is not necessary.
2. But, for simplicity in operation and for saving time for the maintenance engineers, it is also possible to be automated. The automations concerns gathering all measurements of, e.g., vibration, automatically that required for analysis and diagnosis. (see Fig. 5.1):

   - Visualizing trend and the behavior of the deterioration process (vibration in the past and current situation),
   - Predicting the vibration level in the near future, such as next planned stoppage or measuring opportunity,
   - Presenting a graph showing real measurements and predicted values for following up and understanding changes in the condition of components/equipment,
   - Assessing the probability of failure of a component/equipment and its residual life, and
   - Reporting.

3. When the company is not really using CM/CBM in full scale, eMDSS is prepared to cover the shortages in their applications through providing effective alerts, diagnosis, and prognosis.
4. The major problem to be solved is how to integrate eMDSS with the available and relevant software system without generating unavailability of the CM software program. It is solved by using an intermediate database for gathering the measurements. Preparation of eMDSS can then be done at the office of system supplier and not in the end user offices to avoid disturbances in

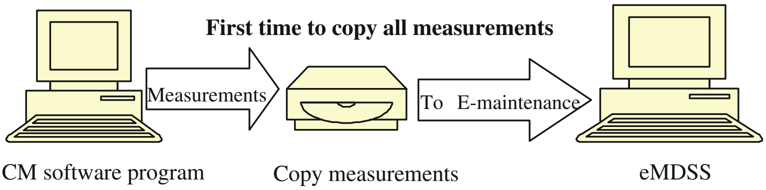


Fig. 5.1 Copy and accommodation of all CM measurements to be processed for eMDSS at the system supplier site

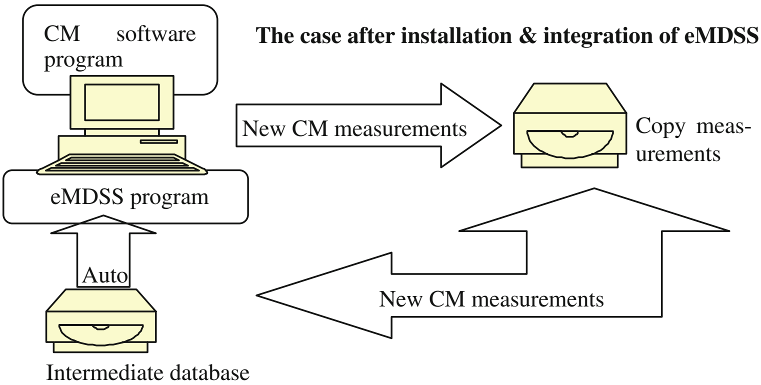


Fig. 5.2 Installation and integration of eMDSS and CM software program for analysis of new CM measurements

production, systems, and performance efficiency. When eMDSS is properly prepared for integration, it is then integrated with the relevant software programs at the company (see Figs. 5.1, 5.2).

5. eMDSS ability to highlight maintenance contribution in company short-term/long-term profit eases company investments.
6. Acquiring a knowledgebase in addition to the database encouraging a company to invest in maintenance because the knowledge and experience can now kept in-house does not matter whether the personnel still working there or not.

5.4 eMDSS Characteristics, Features, and Purposes

The major characteristic of eMDSS is reliable in enhancing maintenance decision accuracy and cost-effectiveness that are needed for reducing economic losses, enhancing company profitability and competitiveness. eMDSS is characterized as follows:

1. It provides relevant and reliable CBM supported by special data- and knowledgebase.
2. It provides relevant and high-quality data concerning past, current, and future condition of components/equipment and criteria for more accurate, dynamic, and cost-effective maintenance decisions.
3. It uses dynamic decision rules, i.e., rules that can be accommodated per case.
4. It combines technical and economic data for achieving a wider and reliable holistic view.
5. It provides a reliable opportunity to easily follow up damage development and changes in operating conditions.
6. It is a flexible DSS that can be accommodated per company and add and remove modules so that each company can build its own eMDSS.
7. It can be used for different objectives, such as
   - more accurate maintenance decisions,
   - mapping, analyzing, and improving maintenance and production processes by identifying and prioritizing problem areas,
   - identification and assessment of losses in production time,
   - simulation of different technically suitable alternatives solutions in order to select the most cost-effective, and
   - following up, controlling, and assessing maintenance contribution in company profit.
8. It supports continuous and cost-effective improvements of maintenance and production performances.

eMDSS features provide reliable answers to the following questions:

1. How long is the residual life? (The time left until failure that is necessary for production planning).
2. What is the probability of failure in the next future?
3. What is the value of the CM parameter in the next future? It is necessary for successful planning of production and maintenance actions.
4. What is the most suitable solution? It is usual that when we have a problem, we may find several alternative solutions. But, the major problem becomes then how to answer; which of these alternatives should be chosen?
5. Is the selected solution really profitable? A theoretical selection of the most profitable solution may not be confirmed in reality. This is why follow up such decisions is crucial from economic and experience point of views.
6. What are the most profitable investments in maintenance? It is to do it right from the first time from both technical and economic perspectives.
7. What is maintenance contribution in company profit? It is important to highlight maintenance role in the company business and follow up the investments done in maintenance for improving its performance whether it was profitable.

Table 5.1 Tools and toolsets included by eMDSS (see also Fig. 5.3)

Toolsets	Tools	Features and function
Toolset 1 to enhance accuracy of maintenance decisions	PreVib (see Fig. 5.4)	To predict the level of a CM parameter at the close future, e.g., at the next planned stoppage or measuring moment. It is to avoid sudden and dramatic changes in the vibration level which may lead to a catastrophic failure
	ProLife (see Fig. 5.5)	To assess the probability of failure and residual lifetime of a component/ equipment at need or when its CM level (vibration level) is significantly high
Toolset 2 to Select the most cost-effective maintenance action	AltSim (alternative simulation) (see Fig. 5.6)	To select the most cost-effective alternative maintenance solution when several applicable solutions/actions are considered. It uses intelligent motor to examine the cost-effectiveness of maintenance solutions. Well-defined economic and technical criteria are used
Toolset 3 to map, follow up, and assess maintenance cost-effectiveness	MainSave (maintenance savings) (see Fig. 5.7)	To monitor, map, analyze, follow up, assess, and control maintenance cost-effectiveness, i.e., maintenance contribution in company profit. It is a reliable tool for securing cost-effective maintenance actions

eMDSS is a software system, and its major purpose is to achieve dynamic and cost-effective maintenance decisions, which can be approached through a combination of the sub-purposes:

- identifying and prioritizing of problem areas in production process,
- achieving more accurate maintenance decisions,
- localizing and judging new investments in maintenance, and
- mapping, following up, analyzing, assessing, and controlling of maintenance economic impact on company business.

eMDSS is built based on many research results that have been achieved at Terotechnology, Linnaeus University, and E-maintenance Sweden AB, Sweden, during the last years, which are published in well-known international journals and conferences' proceedings (see e.g., [10]). It consists of three toolsets where every toolset consists of one to three tools with different functions (see Table 5.1 and Figs. 5.3, 5.4, 5.5, 5.6, 5.7).

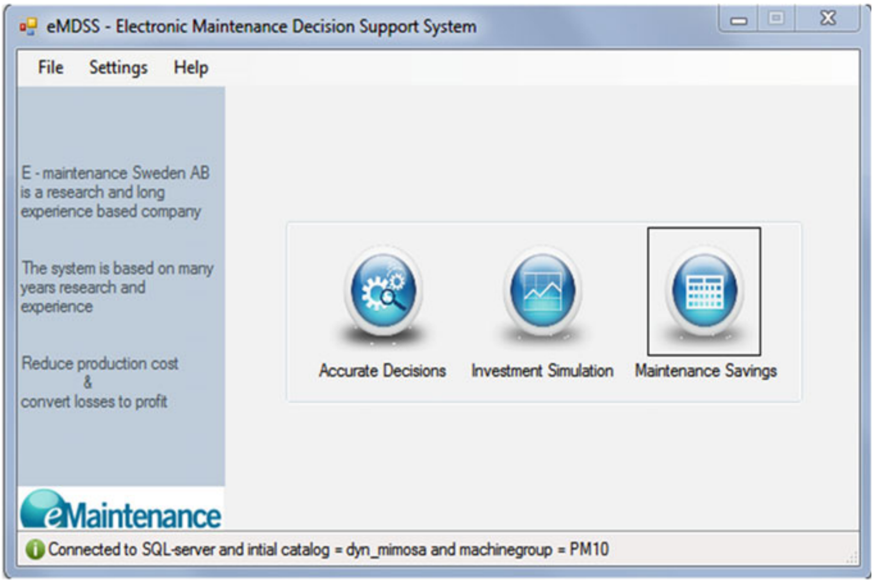


Fig. 5.3 Main menu of eMDSS software program

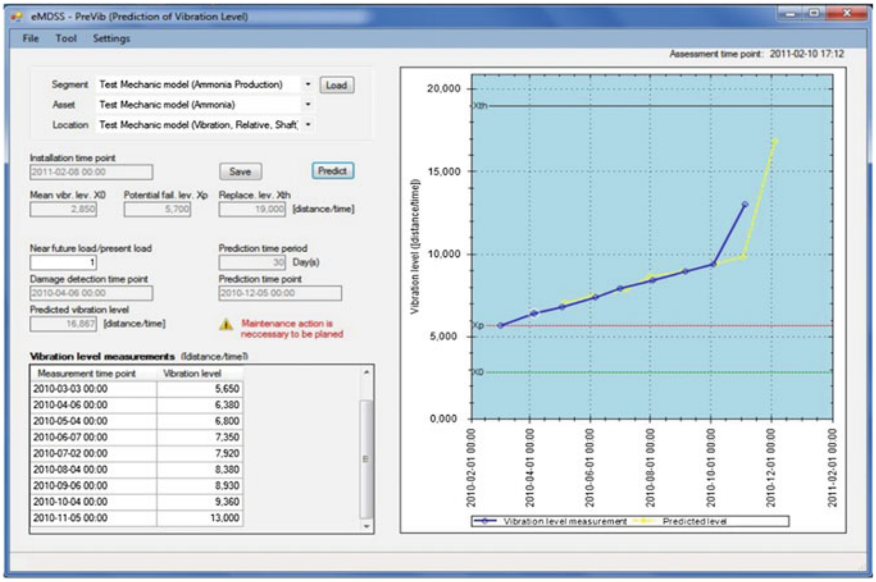


Fig. 5.4 PreVib software module

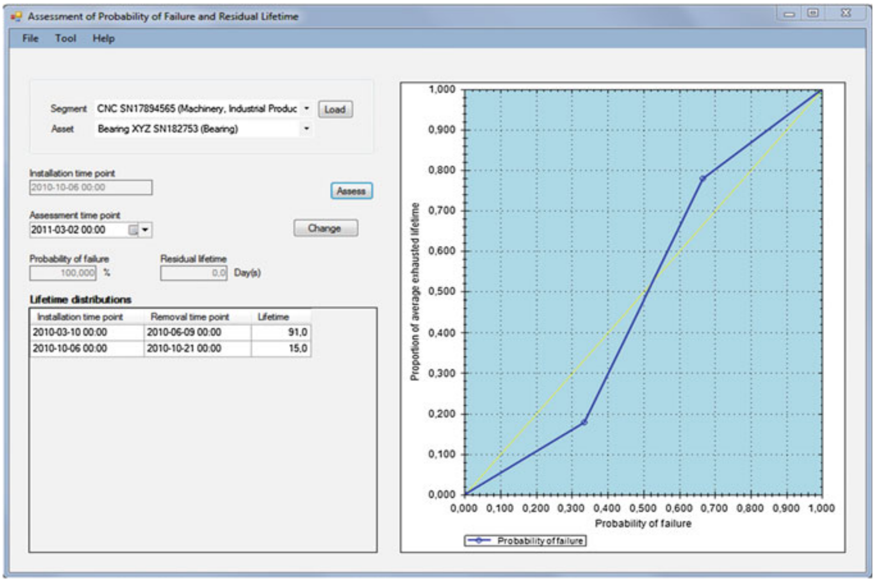


Fig. 5.5 ProLife software module

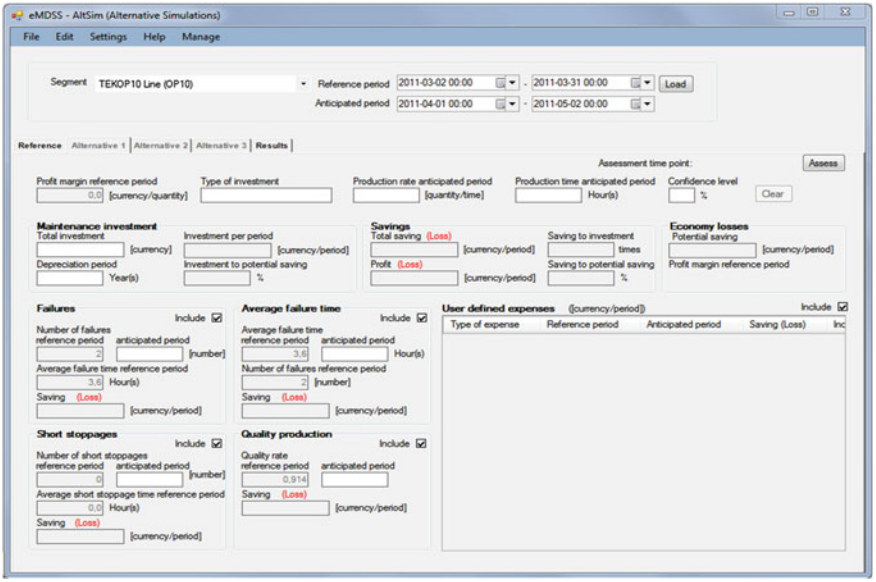


Fig. 5.6 AltSim software module

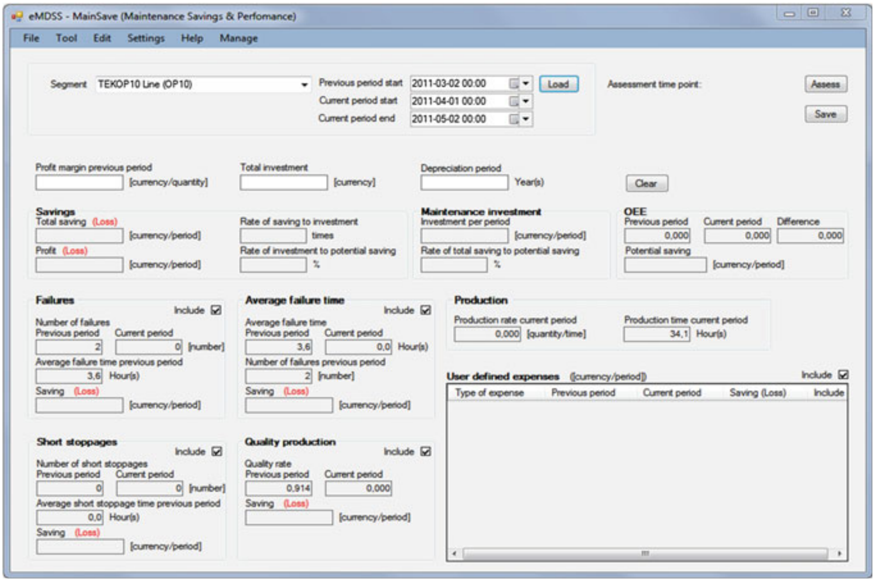


Fig. 5.7 MainSave software module

5.5 Summary and Conclusions

Using eMDSS, it is possible to act at an early stage in both tactical and strategical levels for fulfilling company strategical goals for continuous improvement of its profitability and competitiveness. Applying eMDSS makes it possible to handle real-time data, analysis, and decisions and give necessary information about maintenance and other working areas to the decision makers. It provides better data coverage and quality which are essential for improving knowledge and experience in maintenance and thereby aid in increasing the competitiveness and profitability of a company. Predicting the CM level at the next planned measuring time or planned stoppage reduces the risk of an unexpected deviation in the condition of significant components. This would reduce the number and duration of planned and unplanned stoppage. Also, the probability of failure of a component, such as a rolling element bearing, can be kept very low (close to zero) which has a direct impact on company productivity and, hence, its competitiveness. Applying eMDSS, several traditional problems facing implementation of MDSS can be solved reliably. It is also possible to show how maintenance affects the profit of a company in a way that is hard to do with the available tools and techniques even if the data required are available. Development of relevant and traceable KPIs for technical and economic maintenance control has made this possible.

Acknowledgments This study was partly funded by EU-IP DYNAMITE (Dynamic Decisions in Maintenance) 2005–2009 and partly by E-maintenance Sweden AB.

References

1. Al-Najjar B, Alsyouf I (2003) Just in time maintenance: a way to improve companies' competitiveness. Conference business excellence, University of Minho, Portugal, pp 596–603, 11–13 June 2003
2. Al-Najjar B (2007) The lack of maintenance and not maintenance which costs: a model to describe and quantify the impact of vibration-based maintenance on company's business. IJPPM 55(8):260–273
3. Djurdjanovic D, Lee J, Ni J (2004) Watchdog agent—an infotronics-based prognostic approach for performance degradation assessment and prediction. Adv Eng Inform 17:109–125
4. Garg A, Deshmukh SG (2006) Maintenance management: literature review and directions. JQIME 12(3):205–238
5. Fu C, Ye LQ, Liu YQ, Iung B (2002) MAS-based model of intelligent control-maintenance management system (ICMMS) and its applications. Inter Conf Mach Learn Cybern 1:376–380
6. Cerrada M, Cardillo J, Aguilar J, Faneite R (2007) Agent based design for fault management systems in industrial processes. Comput Ind 58:313–328
7. Nadakatti M, Ramachandra A, Kumar ANS (2008) Artificial intelligence-based condition monitoring for plant maintenance. Assembly Autom 28(2):143–150
8. Sloan TW, Shanthikumar G (2000) Combined production and maintenance scheduling for a multiple-product, single-machine production system. Prod Oper Manage 9(4):379–399
9. Al-Najjar B (2009) A computerised model for assessing the return on investment in maintenance; following up maintenance contribution in company profit. In: Proceedings of WCEAM Sep 2009. Greece, Athens, Springer-Verlag London Limited, London, pp 137–145
10. Holmberg K, Jantunen E, Adgar A, Mascolo J, Arnaiz A, Mekid S (2010) Strategies for cost-effectiveness. E-maintenance. Springer-Verlag London Limited, London

Chapter 6
Infrastructure Delivered Through Networks: Engagement of Stakeholders

Sandra Beach

Abstract Public awareness of infrastructure projects, particularly roads, many of which are delivered through networked arrangements, is high for several reasons. Such projects often involve significant public investment; they may involve multiple and conflicting stakeholders and can potentially have major environmental impacts. To produce positive outcomes from infrastructure delivery, it is imperative that stakeholder "buy in" be obtained particularly about decisions relating to the scale and location of infrastructure. Given the likelihood that stakeholders will have different levels of interest and investment in project outcomes, failure to manage this dynamic could potentially jeopardize project delivery by delaying or halting the construction of essential infrastructure. Consequently, stakeholder engagement has come to constitute a critical activity in infrastructure development delivered through networks. To guide infrastructure governance networks in the undertaking stakeholder engagement, this paper proposes and tests a domain-based model of stakeholder engagement.

6.1 Introduction

Throughout Australia, the construction and redevelopment of public infrastructure continues to be a key factor in economic development. Since the late twentieth century, there has been a shift away from delivery of engineering infrastructures, including road networks, solely by government. Consequently, a range of alternative delivery models have emerged in engineering infrastructure planning, construction, and management. Among these new models, governance networks, in which governments and the industry and community sectors jointly cooperate to

S. Beach (✉)
2 George St., Brisbane, QLD 4001, Australia
e-mail: sandra.beach@qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_6,

provide infrastructure services, have become a widespread delivery mechanism. However, little is known about how these networks engage with stakeholders inside the network and those outside of the network who have the potential to affect or be affected by the infrastructure being delivered.

This paper examines how stakeholders were engaged by three road infrastructure delivery networks in Queensland, Australia. In particular, the paper examines how the relative priorities assigned to stakeholders by these networks relate to stakeholder engagement activities actually undertaken. Furthermore, the paper identifies the major factors that influence both prioritization and engagement of stakeholders.

6.2 Stakeholders

The nature of road construction raises a number of challenges for stakeholder engagement. Road projects may cover considerable distances, traverse different local government boundaries, and have different types of impact on property owners. As a result, stakeholder concerns may therefore vary significantly, requiring specialized and highly time-consuming responses. Acknowledging the need to engage with directly affected stakeholders such as those facing loss of amenity due to factors such as noise, pollution, and reduced property values, there are also likely to be "hard to reach" [1, 6] or reluctant stakeholders who are difficult to identify and connect with.

Yang and Lim [2] have identified that there is a need to establish new approaches that integrate and synthesize the different perspectives of multiple stakeholders within new frameworks of local governance not previously addressed [3]. This is not a simple activity given that road infrastructure may be planned, constructed, and managed through arrangements that involve the public and private sectors working jointly within networks [4]. Furthermore, the scale of infrastructure projects may also impact on stakeholder engagement. While large infrastructure projects may stretch across many local government areas and communities, smaller scale annual works programs may be more localized.

The differing levels of interest may also lead to shifting levels of commitment and contribution to projects and their outcomes [5]. Dealing with these dynamics through stakeholder engagement, however, offers the possibility that potential risks associated with stakeholders can be more effectively managed [6]. As El-Gohary et al. [7] note: "involvement with stakeholders can be a decisive factor that can 'make or break' a project." However, it remains a continuing challenge for governance networks to identify appropriate stakeholders, to determine when and how to engage with them, and to effectively manage these relationships to achieve results and derive benefits.

6.2.1 Stakeholder Management

The process of dealing with a diverse set of stakeholders, possessing different attributes and interests in multiple issues in the delivery of infrastructure, is a complex undertaking. A number of steps which are important in effectively managing stakeholders [8] are presented as an integrated framework in Fig. 6.1.

Focusing on step 1, stakeholder definition and identification have been approached from numerous perspectives [9], including a managerial perspective (why stakeholders and organizations interact) [8] and normative perspective (with whom should organizations interact) [10]. In this study, stakeholders are defined as groups or individuals who can have either an actual or potential effect on governance network outcomes [11, p. 869]. The second step, stakeholder classification, has also been approached in a variety of ways in an effort to prioritize stakeholders [12].

However, one classification system, the stakeholder identification and salience model [11], has particular resonance in the governance network context because of its focus on power dependency and legitimacy in relationships which are also important factors for governance networks [13]. The model theorizes that there are seven stakeholder types comprised of different combinations of power, legitimacy, and urgency. Depending on the combination, stakeholders are accorded a different priority and treated accordingly. The third step, stakeholder engagement, refers to the practices undertaken by governance networks to communicate with stakeholders, involve them in network activities and develop relationships with them [14, 15].

6.2.2 Proposed Stakeholder Engagement Domains

While there have been a number of studies examining the role of stakeholder salience [16, 17], the issue of how stakeholder salience relates to engagement is under researched. Therefore, to answer the question of "How governance networks engage with stakeholders," this study proposes a domain-based model of stakeholder engagement as depicted in Fig. 6.2.

Within the proposed model, it is theorized that there are two Domains of stakeholder engagement. Domain 1 comprises the core network participants who are attributed high priority within the network. Furthermore, there is ongoing engagement within this Domain such that network participants take part in relationship-based processes that focus on collaboration. However, such processes occur at low-frequency intervals.

In Domain 2, it is proposed that stakeholders who are peripheral to the core network are engaged sporadically by the network as issues arise or resources are required. In this Domain, stakeholders would be of moderate-to-low priority and be engaged through a range and frequency of processes. The major difference

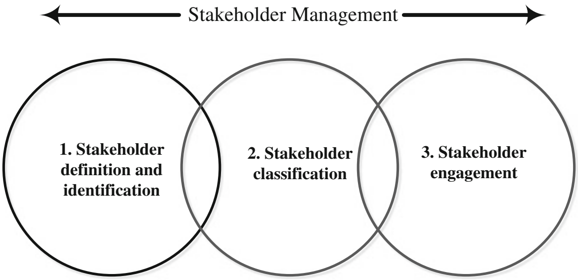


Fig. 6.1 Stakeholder management framework

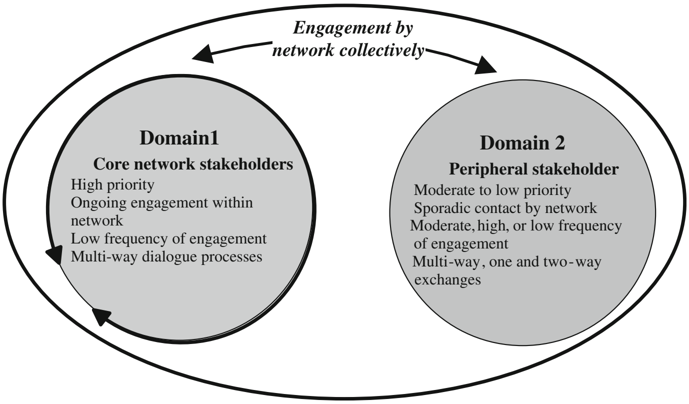


Fig. 6.2 Proposed stakeholder engagement domains

between the two Domains is that in Domain 2 stakeholders are outside of the network boundary, unlikely to be involved the network's governance arrangements and are engaged on a sporadic rather than continuing basis.

6.3 Methodology

To obtain a better understanding of how stakeholder engagement was undertaken by governance networks, three regional road group (RRG) delivery networks at different geographic locations in Queensland, Australia, were selected as case studies:

- Far North Queensland Regional Road Group
- Wide Bay Burnett Regional Road Group
- Northern South–East Queensland Regional Road Group.

The RRGs function as regional decision-making bodies which make use of both bureaucratic and networked structural arrangements to deliver small, but politically significant, regional works programs. RRGs comprise two interlinked groups, one of which focuses on engineering issues and the other operates at the political level. This approach distinguishes between the technical and political aspects of regional road delivery. This approach incorporated a number of key features: It brought together political and technical actors in collaborative decision-making processes and provided a mechanism for long-term road planning that transcended local and state election cycles.

A purposeful sample [18] of key informants from each network was selected for this study. The sampling strategy was designed to obtain a mix of organizations represented in the network, differing occupational categories and long-standing and "new" key informants. Three methods of data collection were employed: stakeholder analysis, structured interviews, and semi-structured interviews. Documentary evidence was used as additional source of information to confirm interviewee accounts and to identify inconsistencies. By undertaking this type of triangulation [19] across data sources, the validity of the results was enhanced.

6.4 Emergent Themes

In testing the proposed model of stakeholder engagement domains across the three cases, a number of issues emerged. Two major issues are discussed in this paper: stakeholders identified and stakeholder engagement combinations. The final issue, Domains of stakeholder engagement, is discussed in the conclusion section.

6.4.1 Stakeholders Identified

The networks identified a number (29) of stakeholders with whom they had varying levels of contact. However, as can be seen from Fig. 6.3, the majority of these stakeholders (69 %) were focused on transport management, indicating a strong focus on transportation issues. The remaining stakeholders had a diversity of roles including regional and economic development and environmental management.

Surprisingly, each of the networks identified a number (27) of stakeholders with whom they had no contact. Figure 6.4 shows the roles and frequency with which these stakeholders occurred. The networks were aware of a broad range of stakeholders despite having no current contact with them; however, road users and community groups are an omission from this list.

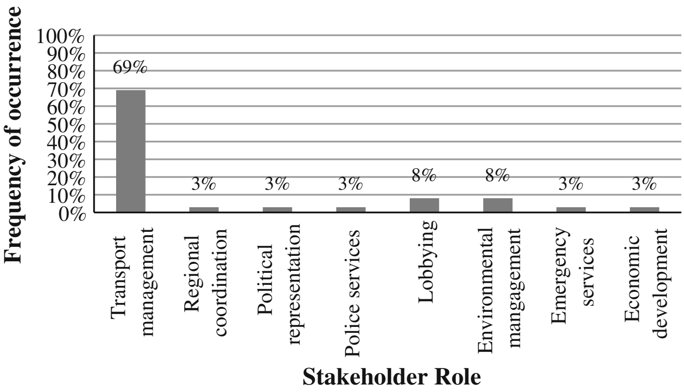


Fig. 6.3 Roles of stakeholders with whom networks had contact

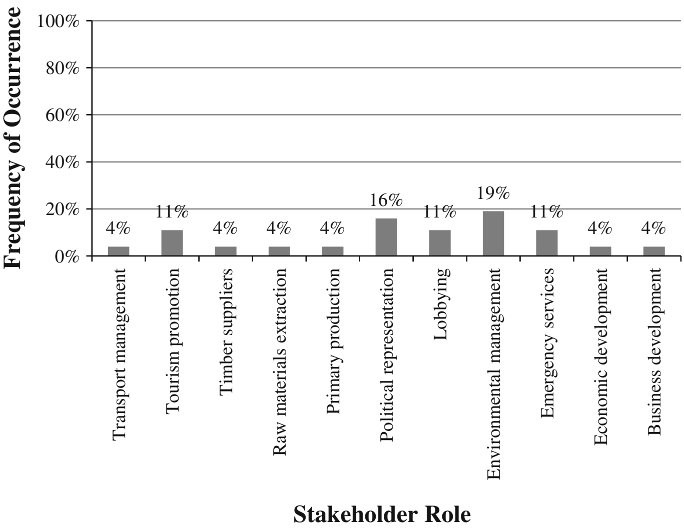


Fig. 6.4 Roles of stakeholders with whom networks had NO contact

In summary, it can be seen from Figs. 6.3 and 6.4 that the networks identified a broad range of stakeholders as having relevance to their activities; however, they had no contact with 49 % of these stakeholders. Of those stakeholders with whom the networks had contact, the majority (69 %) of these were transport management related. Thus, transport management stakeholders had a predominant role in the networks and among others; road users and communities were not identified as having a role at all.

Table 6.1 Combinations of stakeholder engagement

Engagement processes	Engagement frequency		
	Low	Moderate	High
Multi-way dialogue	Core network stakeholders (69 %)	Peripheral stakeholders (3.5 %)	
Two-way exchanges	Peripheral stakeholders (10 %)	Peripheral stakeholders (3.5 %)	Peripheral stakeholders (3.5 %)
One-way communication	Peripheral stakeholders (7 %)	Peripheral stakeholders (3.5 %)	

6.4.2 Stakeholder Engagement Combinations

Across the networks, seven combinations of stakeholder frequency and type of engagement processes were with stakeholders. As can be seen from Table 6.1, almost 70 % of stakeholders, i.e., core network stakeholders, were engaged through multi-way dialogue processes that occurred at low frequency.

The remainders, i.e., the peripheral stakeholders, were engaged through six different combinations of engagement. Therefore, the predominant combination of stakeholder engagement undertaken by the networks was multi-way dialogue processes that occurred at low frequency.

6.5 Conclusions

This paper concludes that there are three possible Domains of stakeholder engagement by infrastructure delivery networks as identified in Fig. 6.5. However, only Domain 1 is in broad use suggesting that infrastructure delivery networks engage intensively with internal stakeholders and exclude external stakeholder from participating in important activities that may have negatively affected them.

Accordingly, stakeholder engagement by infrastructure governance networks follows an "inner" and "outer" group pattern differentiated by network membership. This extent of solidification of network boundaries represents a potential problem for road delivery governance networks because it buffers the network from external stakeholders [12]. As such, these networks have no mechanism for detecting changes in the stakeholder environment that may be important in road project prioritization.

This paper also evidences that externally imposed "rules" embedded in the networks' institutional arrangements effectively prescribed the focus, frequency, and processes of stakeholder engagement undertaken by the networks. Although preexisting relationships within the networks facilitated stakeholder engagement in

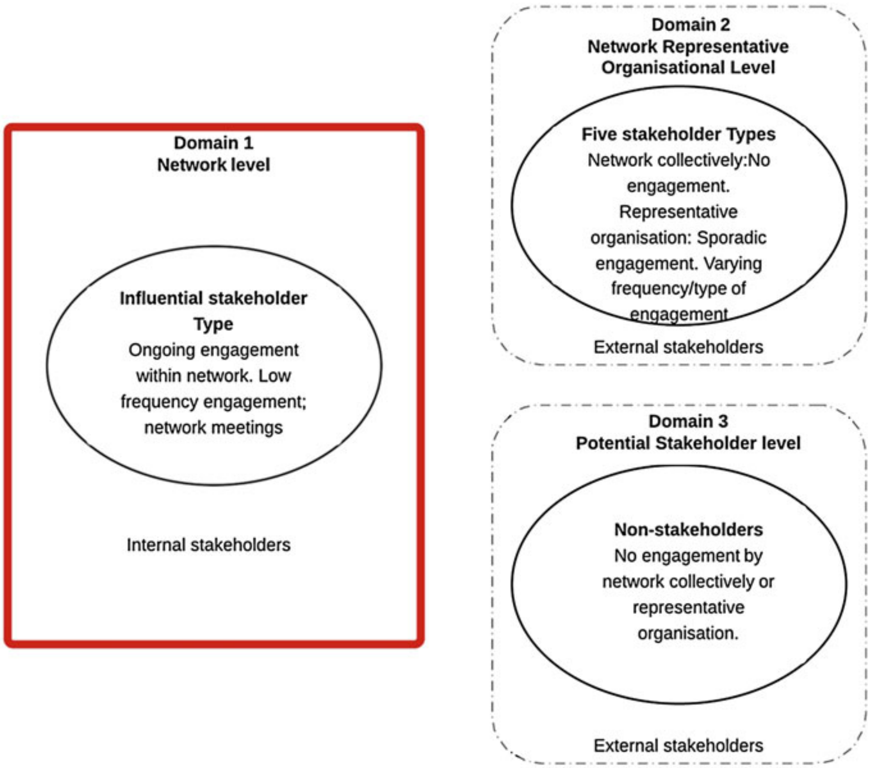


Fig. 6.5 Actual stakeholder engagement domains

some instances, institutional "rules" were found to be more influential. These results suggest that stakeholder engagement within infrastructure delivery networks is likely to be stable and routinized as a result of following practices embedded in the institutional framework. However, this stability provides little incentive for the networks to extend engagement beyond network boundaries and build mutually beneficial relationships with external stakeholders.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established, and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Brackertz N, Zwart I, Meredyth D, Ralston L (2005) Community consultation and the 'hard to reach' concepts and practice in victorian local government. Institute for Social Research, Swinburne University of Technology, Melbourne

2. Yang J, Lim S (2008) Reality check–the identification of sustainability perception and deliverables for australian road infrastructure projects. Paper presented at the New Zealand society for sustainability engineering and science, Blueprints for sustainable infrastructure conference, Auckland, NZ
3. Tavares A, Camões P (2010) New forms of local governance: a theoretical and empirical analysis of the case of Portuguese municipal corporations. Public Manage Rev 12(5):587–608
4. Chinyio E, Olomolaiye P (2010) Introducing stakeholder management. In: Chinyio E, Olomolaiye P (eds) Construction stakeholder management Wiley-Blackwell, Oxford, UK
5. Newcombe R (2003) From client to project stakeholders: a stakeholder mapping approach. Constr Manage Econ 21(8):841–848
6. Bourne L, Walker D (2008) Project relationship management and the stakeholder Circle™. Int J Managing Proj Bus 1(1):125–130
7. El-Gohary N, Osman H, El-Diraby T (2006) Stakeholder management for public private partnerships. Int J Project Manage 24(7):595–604
8. Freeman R (1984) Strategic management: a stakeholder approach. Pitman, Boston
9. Miles S (2012) Stakeholder: essentially contested or just confused? J Bus Ethics 108(3):285–298
10. Donaldson T, Preston L (1995) The stakeholder theory of the modern corporation: concepts, evidence and implications. Acad Manag Rev 20:65–91
11. Mitchell R, Agle B, Wood D (1997) Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts. Acad Manag Rev 22(4):853–886
12. Friedman A, Miles S (2006) Stakeholders: theory and practice. Oxford University Press, New York
13. Klijn E, Koppenjan J, Termeer K (1995) Managing networks in the public sector: a theoretical study of management strategies in policy networks. Public Adm 73(3):437–454
14. Chinyio E, Akintoye A (2008) Practical approaches for engaging stakeholders: findings from the UK. Constr Manage Econ 26(6):591–599
15. Harrison S, Mort M (1998) Which champions, which people? Public and user involvement in health care as a technology of legitimation. Soc Policy Adm 32(1):60–70
16. Agle B, Mitchell R, Sonnenfeld J (1999) Who matters to CEOs? An investigation of stakeholder attributes and salience, corporate performance, and CEO values. Acad Manag J 42(5):507–525
17. Parent M, Deephouse D (2007) A case study of stakeholder identification and prioritization by managers. J Bus Ethics 75(1):1–23
18. Patton M (2002) Qualitative research and evaluation methods, 3rd edn. Sage Publications Inc., Thousand Oaks, California
19. Miles M, Huberman A (1994) Qualitative analysis: an expanded sourcebook. Sage Publications, Thousand Oaks, California

Chapter 7
Sustainability Performance in Modern Industrial Context: From Firm-Centric to Network-Based Assessment for Realistic Impact

Jakob E. Beer and Jayantha P. Liyanage

Abstract The popular discussion of sustainability, its importance, and the need for its systematic assessment opened up many research venues and has led to the development of a variety of methods to measure and improve sustainability performance. While there have been some genuine efforts on developing life cycle-based research regimes, most of these efforts and subsequent solutions have seemingly been limited to the products and processes within the firms'-specific physical or organizational boundaries. In this paper, the authors argue that concentration on firms' boundaries to improve sustainability performance is ineffective. This paper, therefore, brings a network-based approach into perspective for a systemic and dedicated sustainability performance improvement that extends beyond the prevailing firm-based approach. Shortcomings of firm-based approaches are highlighted, and how they can be avoided in a network setting will be outlined.

7.1 Introduction

The unabated importance and popularity of sustainability as a topic in industry and research have led to development and diffusion of a host of frameworks to assess and report sustainability [8]. Apart from life cycle-based approaches such as life cycle assessment (LCA), many of the popular frameworks for sustainability performance assessment are based on a firm-centric, if not production site-centric perspective; Global Reporting Initiative (GRI), EMAS, ISO 14031, and Sustainability Balanced Scorecard [9, 16, 24] are some prominent examples.

J. E. Beer (✉) · J. P. Liyanage
Universitetet i Stavanger, 4036 Stavanger, Norway
e-mail: jakob.e.beer@uis.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_7,

In this paper, we argue that the concentration on local sustainability reporting neglects the fact that much of today's value creation in manufacturing takes place in networks [17, 29, 35, 37]. Concentration on firm-level reporting, let alone production site-level reporting, fails to address the complex interdependencies in modern manufacturing networks [4, 27] and thus leaves great potential untapped. After outlining some key attributes of the transition process, we will discuss some arguments as to why we believe purely firm-centric sustainability performance approaches cannot provide the basis for realistic assessment and significant improvement.

The results have been developed within the EU-funded project Sustain Value. One of the outcomes of the project is a sustainability framework for complex manufacturing networks. The results discussed in this paper lay out the groundwork for the framework (cf. [2] for a more detailed discussion of the framework).

7.2 Development Toward Network-Based Value Creation and Its Performance Impact

Networks have become the dominant form of value creation in some industries such as automotive [17, 37] and consumer electronics [35], and in fact, it often is networks and not firms that compete with each other [14].

> Competencies necessary for survival and growth are now as rooted in network architectures and ecosystems as they are in the fiber of individual companies. Strategies are no longer the limited purview of a company-focused view of value and control over its resources. Rather they stretch as far as the boundaries of the global networks within which companies must find their place [28].

The existence of networks is quite distinct from what is normally termed 'supply chains.' Supply chains do exist, but the chain model normally holds true only when adopting a product-based flow perspective. For management, relevant actions take place in a network of firms consisting of several streams of supply that can have strong interrelations [37]. Performance effects resulting from network membership—though difficult to separate from other performance effects and even more difficult to quantify [15]—can thus be expected to have implications beyond what Forrester [10] proposes for supply chains. The rather simplistic 'chain notion' can thus be misleading as there are effects in networks that do not exist in chains [3].

Prahalad and Krishnan [29] describe the development from the firm-centric 'Ford model' with high degree of vertical integration toward global networks. In several industries, it can be observed that focal companies have outsourced major shares of value-adding activities to suppliers and service providers. Two different approaches can be used to explain the emergence of complex manufacturing networks: exogenous factors (resource distribution etc.) and endogenous factors (firm and network properties) [21]. Given the complexity of many modern products, individual firms are unable to provide all the necessary skills required, hence an increased number of entities are involved in product and service provision. The higher number of entities involved leads to an increased need for coordination [20]. Complexity is increased due to a high rate of technological innovation, the introduction of new technologies incumbents are not familiar with (such as hybrid technology for cars [33]), shorter product life cycles [18] which, in turn, require faster time to market [20, 23, 38], changing customer expectations as well as changing political conditions in both different customer (e.g., China) and supply markets (e.g., the Arab spring). Other reasons for this development can be seen in strategies to cope with risk in volatile markets [20, 36], benefits from lower labor costs at suppliers [18], and scarcity of qualified personnel [22]. Drawing on transaction cost theory and social network theory, Jones et al. [20] identify four conditions for networks to emerge: demand uncertainty with stable supply (as unstable demand requires flexibility which networks can provide through faster and cheaper reallocation of resources), asset-specific exchanges (as they create dependency between parties), complex tasks under time pressure (as it increases the need for coordination), and frequent exchanges among parties (as it allows 'human asset specificity' to develop and increases the level of tacit knowledge among the parties involved) [20]. While Prahalad and Krishnan [29] describe the general development toward ad hoc relationships between those that request and those that provide resources, other authors propose the notion of long-term network relationships that provide advantage over market resource allocation [38]; the widest-spread notion may be that the development outlined above leads to structural embeddedness [12].

Figure 7.1 depicts the development toward complex networks.

To achieve significant improvements in sustainability performance, it is necessary to have the 'right' network partners. Certain improvements cannot be realized with misaligned interests or objectives in the network and network partners certainly have impact on an individual firm's performance [14]. If a manufacturing firm's product development department aims at using more durable components to extent product lifetime, for instance, they depend on their suppliers' capabilities to produce parts that meet the new specs.

The arguments presented above suggest that value creation has changed and in increasingly many industries happens within networks. It seems reasonable to ask what implications this has for individual firms.

There has been some research on how to assess performance impact from networks on individual firms (cf. [15, 30]). Assessment of performance impact on firms, let alone quantitative assessment, turns out to be difficult for the separation from other performance effects is hardly possible. Put in different terms, the dependent variable —the performance outcome—is not *only* a function of the network but of other variables, too, and the firm may change the network itself through its membership [31].

Considering the problem of separation of performance effects, it is even hard to prove whether there is any effect from the network at all, although common sense suggests just this. Within the domain of social network research, some concepts have been developed that prove helpful for understanding the interaction of the

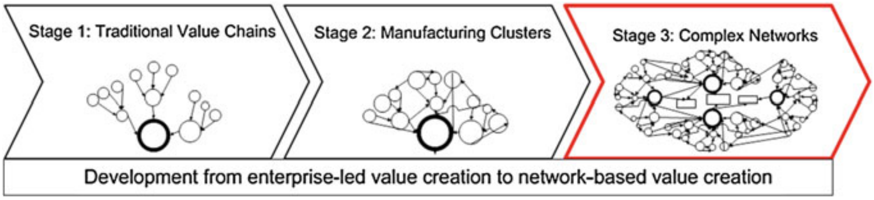


Fig. 7.1 Development toward complex networks

firm with its network and the performance effects that may result from the interaction. Freeman [11] has discussed the concept of centrality of firms in networks as a proxy of power. Combining the concept of centrality with the concept of density—a measure for the relative number of ties in a network—, Rowley [32] characterizes a focal firm's ability to manage stakeholder pressure. He proposes, for instance, that the combination of high density of the network and low centrality will result in a 'subordinate' role of the firm whereas high centrality of the firm combined with low density of the network would provide the firm with a 'commander' position. Although somewhat simplistic, the model proposed by Rowley [32] is able to demonstrate an important aspect of firm interaction with networks: The distribution of power and how may enable or restrict a firm in its options. Thereby, Rowley [32] extends the work done earlier by Oliver [26] and puts it into a more general context (Oliver's focus is on firm responses to institutional pressure). Besides 'structural configuration of ties' (e.g., strong vs. weak ties as indicator of embeddedness, and cohesive vs. bridging ties) and 'partner profiles' (partners' status etc.), centrality is also one of Gulati et al.'s [15] dimensions suggesting a certain level of performance impact from the network on the individual firm.

7.3 The Inadequacy of the Firm-Centric Perspective

7.3.1 Weaknesses of Firm-Centric Approaches

To allow a focused discussion about performance values, it is necessary to have a clear definition of system boundaries based on which the performance is to be measured. A system can be defined as

> (…) a set of bounded interrelated elements with emergent properties [which it] represents within the context of a paradigm [34].

The definition of system boundaries has various implications for sustainability performance. It determines, for instance, which stakeholders are considered in the performance assessment, thus laying out the fundamental design of the assessment process. With respect to the development of manufacturing from firm-centric value creation to network-based value creation, a broader set of stakeholders will need to be included in the system boundaries for sustainability performance management. A vague or 'wrong' definition of system boundaries may result in the creation of incentives counter to sustainability performance objectives. Two extreme theoretical positions for system boundaries can be identified in the present context of manufacturing. On the one hand, the legal and/or physical boundaries of a manufacturing site can represent system boundaries for performance management. This is a common setup, and several frameworks are based on the idea of measuring sustainability performance on the firm level. On the other hand, system boundaries can be drawn along the entire life cycle of a product across several firm boundaries, consumers, and other stakeholders. To evaluate sustainability performance along this path is the idea life cycle assessment is based upon. While the former concept can be described as a static perspective, the latter can be described as a 'flow' perspective. During its life cycle, the product flows through several 'open systems' [39], each of which could constitute a separate unit of analysis for sustainability performance from a local static perspective.

There are several reasons as why we believe the firm-centric or production site-level approach, as pursued by several prominent sustainability frameworks, to be insufficient. The weaknesses concern two main activities of performance management: (1) the assessment and (2) the improvement actions. The objective of performance assessment is to draw a meaningful and representative picture of the firm's performance. Information about performance is essentially worthless, though, without subsequent improvement actions. When firms confine themselves to local assessment and improvement, they are unlikely to achieve significant improvements as those may stretch beyond firm boundaries. Below, we summarize arguments against a purely firm-centric or production site-centric approach:

- **Limited Share of Production Site on Environmental and Societal Impact**: As outlined above, manufacturing firms normally are part of a network consisting of formal contractual relationships with suppliers, service providers, and customers, all directly involved in value creation. An assessment of local sustainability performance does not include those value-adding activities performed by suppliers and service providers outside the boundaries of the manufacturing firm. On the one hand, focusing on one particular manufacturing site as relevant measurement system allows relatively easy identification of responsibilities, stakeholders, and of system outputs, such as toxic waste, CO_2 emission, or waste water generation. Questions can be raised with regard to the relevance of these results, however. The sustainability impact of an individual production site may even be insignificant. Even in energy-intensive industries such as automotive, focusing on the sustainability performance of one particular production site (e.g., car assembly site on OEM stage) does not tell much about the true impact of the production process and the product on environment and society. It ignores energy-intensive logistics [1], manufacturing activities on different value-adding stages (e.g., tier-1 or raw material extraction), customer use, and end-of-life cycle treatment. If major value-adding activities remain outside performance management scope, the assessment loses significance and hardly creates a meaningful impression of the true sustainability impact. Sturgeon [35] discusses the case of the US electronics industry. He points out how major manufacturing activities have moved from OEMs to global suppliers. Assessing sustainability performance of OEM sites which mainly maintained product development and marketing capabilities would fail to create meaningful information about sustainability performance. The same applies to much of the automotive industries, where most OEMs have kept only a relatively small share manufacturing (often only about 30 %) and in their production sites essentially assemble the parts and modules supplied by manufacturing partners [19]. Not only may the significance of the measurement results be limited but also the opportunity for improvement. The latter depends on the particular industry and competitive position of the firm, however.

- **Risk in Pursuing Improvement Measures**: Firms-pursuing measures to improve sustainability performance without involving network partners have to carry the risk of failure alone. The introduction of a new business model, invest in new facilities and machines, or changed processes bear the danger of ineffectiveness which may result in economic losses. Own objectives and objectives among partner firms may be disparate, rendering the improvement measures ineffective. Narayanan and Raman [25] have described how disparate incentives in supply networks can lead to inferior performance. Objectives and incentives of partner firms may be misaligned once processes of the focal firm change if they are not involved in or informed about the change procedure.
- **Limited Resources of the Firm**: While resources and motivation of individual companies may be sufficient to *assess* sustainability performance, they may not in every case be sufficient to achieve significant *improvements*. SMEs in particular may be over-challenged without guidance from more experienced partners. It is certainly a fair statement to say that most likely the majority of SMEs does not have personnel with educational background or practical experience in sustainability.
- **Limited Capabilities of the Firm**: Traditionally, education both in engineering and in management has hardly (if at all) included sustainability or ethics as topics. Only recently have more universities started to include such fields in their curricula. It will take some time until graduates will have dispersed so that SMEs will benefit from their knowledge. With multiple firms of the manufacturing network cooperating for sustainability improvements, knowledge gaps could be overcome and individual firms could be driven to performance they would otherwise be incapable to reach.
- **Misleading Incentives**: Confining sustainability performance assessment on a firm or production site perspective creates the incentive to outsource those activities that score particularly bad for the performance assessment. Moving those activities outside system boundaries will let the firm (or its production site, respectively) appear in a more sustainable way. In some industries, there has already been a significant shift of value-added upstream supply networks with OEMs trying to get rid of certain risks inherent to their industry such as

highly fluctuating demand [35]. The result is that important drivers of bad sustainability performance will lie outside responsibility and control. It does not solve a single problem from a global perspective; it is just the local performance that looks better as negative performance drivers lie outside system boundaries.

- **Limited Foresight of the Firm**: Individual firms—and SMEs in particular—may lack the foresight necessary to understand the implications of their actions in terms of sustainability. Being in frequent touch with other firms of the value network can help create the awareness for sustainability; having a partner firm in the network with experience in improving sustainability performance can be inspiring. Until that is going to happen, there is reason to assume that most SMEs lack awareness and foresight.

While all of the arguments mentioned above are important to consider, the first argument—limited impact of the production site on environment and society—appears to be the most relevant one. Its relevancy rests on the shift from firm-centric value creation to network-based value creation [28, 29], which has altered the distribution of risks and benefits along manufacturing companies, their suppliers, and customers [4].

On the other extreme, however, life cycle-based approaches try to integrate the entire life cycle of a product and its impact on environment and society by integrating several open systems in the assessment. While this may work well in theory, this approach will be cumbersome or will turn out not feasible in practice. Assessing sustainability performance along the entire life cycle of a product is only possible with very rough simplifications and under certain assumptions that render the results somewhat arbitrary. For SMEs which represent the vast majority of all businesses, conducting comprehensive life cycle assessment does not seem feasible. So essentially, between the two extreme positions—a firm-centric or production site-centric approach on the one side and a life cycle approach on the other—there is a trade-off between relevance and feasibility. The relevance of 'local' sustainability performance depends on the industry and the specific market position of the firm (in terms of both its market power and its vertical position in the value-network) but is generally lower than of an assessment including several stages of the product life cycle.

7.4 Taking into Account Network Complexities in Performance Context: A Road map

The question remains how the network perspective—which would need to include several value-adding entities without creating the need for a full-blown life cycle assessment—can be taken into account for sustainability performance assessment. If there is no external entity that could impose a network-wide performance assessment on all actors, firms in the network have to tackle the issue on their own.

Since focal firms—and SMEs in particular—will normally not be able to dictate the conditions of performance assessment to all network entities, dyadic relationships will need to be in the focus of attention.

Pointing out network complexities and proposing an assessment of dyads may sound like a contradiction; indeed, such an assessment would fall short of expectations with respect to emergent effects resulting from complex systems behavior. For a practical approach, however, it seems adequate to focus on the most important ties of the firm. Also, there is no need for a scientific analysis as to which ties are the most important—each firm will recognize the right ties intuitively when provided with a guiding framework. As pointed out earlier, individual firms are somewhat limited in their options to improve sustainability performance. Significant improvements—as they could be achieved, for instance, through radically different business models, use of environmentally friendly transportation modes, use of environmentally friendly extraction methods for raw material, use of non-toxic and recyclable materials and components, etc.—involve cooperation of suppliers and service companies. In particular, it requires aligned objectives, a match of relevant capabilities and healthy, trustful partnership (sometimes referred to as relational embeddedness; cf. [13]). Therefore, internal levers to improve sustainability performance cannot be assessed separately from certain network conditions without losing important information about causes and effects. Donabedian [5–7] therefore proposes a three-part approach (to quality management in networks) which has been adopted by Provan and Sydow [30] for performance assessment in networks. It requires that causal links are established between network conditions, internal performance levers, and common performance outcome. We follow the general idea and propose it for assessment of sustainability performance in manufacturing networks (for a detailed discussion, cf. [2]).

7.5 Conclusion

In this paper, we have argued that focusing on an individual firm or production site to assess and manage sustainability performance is conceptually wrong and does not reflect the realities of many modern manufacturing environments. A great proportion of value is created in networks of firms; accordingly, sustainability performance is influenced by a variety of network entities. We presented arguments as to why focusing on a single firm as unit of analysis may not provide meaningful results in the assessment, and we have explained why significant improvements in sustainability performance may not be achieved by single firms on their own. Literature both on performance as well as on sustainability is still focused on individual organizations; the boundaries of the organization, however, do not reflect the distribution of risks and opportunities inherent to complex networks.

Therefore, we propose to include an assessment of the most important dyadic relationships with respect to such characteristics as objective alignment, partnership health, and capability matching. These characteristics provide the basis for cooperative work on improving sustainability performance and are causally linked to internal performance levers, i.e., opportunities to improve sustainability performance that have their origin within the firm's organizational boundaries.

Acknowledgments The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°262931.

References

1. Beer JE, Liyanage JP (2011) Sustainability risks for global production clusters in the automobile industry: a case of supplier networks. In: Proceeding of the APMS 2011 conference. value networks: innovation, technologies and management, Stavanger, Norway
2. Beer JE, Liyanage JP (2012) Principal components of a governing framework for sustainability in manufacturing networks. In: Towards integration and interoperability in engineering asset management. Proceedings of the 7th world congress on engineering asset management. Springer
3. Beer JE, Schumacher T, Liyanage JP (2012) Network performance impact on supply reliability in the automobile industry. In: Proceedings of the 12th PICMET conference. Vancouver, Canada
4. Crook C (2009) Complexity theory: making sense of network effects. In: Kleindorfer PR, Wind Y (eds.) the network challenge: strategy, profit, and risk in an interlinked world. Wharton School Publishing, Philadelphia, pp 207–223
5. Donabedian A (1988) The quality of care: how can it be assessed? J Am Med Assoc (JAMA) 260(12):1743–1748
6. Donabedian A (2005) Evaluating the quality of medical care. Milbank Q 83(4):691–729 (reprinted from Milbank Memorial Fund Q 44(3, Pt. 2):166–203 (1966))
7. Donabedian A, Bashshur R (2002) An introduction to quality assurance in health care. Oxford University Press, Oxford
8. Feng SC, Joung C (2010) An overview of a proposed measurement infrastructure for sustainable manufacturing. In: Proceedings of the 7th global conference on sustainable manufacturing
9. Figge F, Hahn T, Schaltegger S, Wagner M (2002) The sustainability balanced scorecard—linking sustainability management to business strategy. Bus Strategy Environ 11(5):269–284
10. Forrester JW (1958) Industrial dynamics: a major breakthrough for decision makers. Harvard Business Rev 36(4):37–66
11. Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–239
12. Granovetter MS (1985) Economic action and social structure: the problem of embeddedness. Am J Sociol 91(3):481–510
13. Granovetter MS (1992) Problems of explanation in economic sociology. In: Eccles RG, Nohria N (eds) Networks and organizations structure, form and action. Harvard Business School Press, Boston, pp 25–56
14. Gulati R (1998) Alliances and networks. Strateg Manag J 19:293–317
15. Gulati R, Dialdin DA, Wang L (2002) Organizational networks. In: Baum JAC (ed) Blackwell companion to organizations. Blackwell Publishers, Oxford, pp 281–303
16. Hahn T, Wagner M (2001) Sustainability balanced scorecard. Von der Theorie zur Umsetzung. Centre for Sustainability Management (CSM), Lüneburg
17. Hensel J (2007) Netzwerkmanagement in der Automobilindustrie. Erfolgsfaktoren und Gestaltungsfelder. Deutscher Universitäts-Verlag|GWV Fachverlage GmbH, Wiesbaden

18. Holweg M (2008) The evolution of competition in the automobile industry. In: Parry G, Graves A (eds) Build to order: the road to the 5-day car. Springer, London, pp 13–34
19. Ihme J (2006) Logistik im Automobilbau. Logistikkomponenten und Logistiksysteme im Fahrzeugbau; mit zahlreichen Tabellen. Hanser, München
20. Jones C, Hesterly WS, Borgatti Stephen P (1997) A general theory of network governance: exchange conditions and social mechanisms. Acad Manag Rev 22(4):911–945
21. Kenis P, Oerlemans L (2008) The social network perspective: understanding the structure of cooperation. In: Cropper S, Ebers M, Huxham C, Smith Ring P (eds) The Oxford handbook of inter-organizational relations. Oxford handbooks. Oxford University Press, Oxford, pp 289–312
22. Kloepfer I, Sonnet C (2010) Wo sind all die Ingenieure hin? Frankfurter Allgemeine Zeitung (Aug 2010)
23. Kuhn A, Hellingrath B (2002) Supply Chain Management. Optimierte Zusammenarbeit in der Wertschöpfungskette. Engineering. Springer, Berlin
24. Möller A, Schaltegger S (2005) The sustainability balanced scorecard as a framework for eco-efficiency analysis. J Ind Ecol 9(4):73–83
25. Narayanan VG, Raman A (2004) Aligning incentives in supply chains. Harvard Bus Rev 82(11):94–102
26. Oliver C (1991) Strategic responses to institutional processes. Acad Manag Rev 16(1):145–179
27. Parunak HVD, Savit R, Riolo RL (1998) Agent-based modeling vs. equation-based modeling: a case study and users' guide. In: Proceedings of the First international workshop on multi-agent systems and agent-based simulation. Springer, London, UK, pp 10–25
28. Prahalad CK (2009) Creating experience: competitive advantage in the age of networks. In: Kleindorfer PR, Wind Y (eds) The network challenge: strategy, profit, and risk in an interlinked world. Wharton School Publishing, Philadelphia, pp 25–36
29. Prahalad CK, Krishnan MS (2008) The new age of innovation: driving cocreated value through global networks. McGraw-Hill, New York
30. Provan KG, Sydow J (2008) Evaluating inter-organizational relationships. In: Cropper S, Ebers M, Huxham C, Smith Ring P (eds) The Oxford handbook of inter-organizational relations. Oxford handbooks. Oxford University Press, Oxford, pp 691–716
31. Provan KG, Fish A, Sydow J (2007) Interorganizational networks at the network level: a review of the empirical literature on whole networks. J Manag 33(3):479–516
32. Rowley TJ (1997) Moving beyond dyadic ties: a network theory of stakeholder influences. Acad Manag Rev 22(4):887–910
33. Schmidt-Carré A (2010) Allianzen mit Akkubauern. Financial Times Deutschland (Dec 2010)
34. Shehabuddeen N, Probert D, Phaal R, Platts K (1999) Representing and approaching complex management issues. Part 1: role and definition. Centre for Technology Management Working Paper Series. Institute for Manufacturing at the University of Cambridge Engineering Department, Cambridge
35. Sturgeon TJ (2002) Modular production networks: a new american model of industrial organization. Ind Corp Change 11(3):451–496
36. Tsay AA (1999) The quantity flexibility contract and supplier–customer incentives. Manage Sci 45(10):1339–1358
37. Urban G (2007) Das ganze Zuliefernetzwerk im Griff—Innovations- und Effizienzpotenziale nutzen. Von der Zulieferkette zum Zuliefernetzwerk. In: Gehr F, Hellingrath B (eds) Logistik in der Automobilindustrie. Innovatives Supply Chain Management für wettbewerbsfähige Zulieferstrukturen. Springer, Berlin, Heidelberg, pp 1–4
38. Uzzi B (1997) Social structure and competition in interfirm networks: the paradox of embeddedness. Adm Sci Q 42(1):35–67
39. von Bertalanffy L (1950) The theory of open systems in physics and biology. Science 111(2872):23–29

Chapter 8
Principal Components of a Governing Framework for Sustainability Performance in Manufacturing Networks

Jakob E. Beer and Jayantha P. Liyanage

Abstract There are a variety of existing frameworks that specify what to measure and how to improve on sustainability performance and there is constant discussion on indicators for performance assessment. Current practice mostly relates to firm-centric approaches, however and thereby neglects the realities of a networked manufacturing environment. In this paper, we argue for sustainability performance assessment that extends beyond firm boundaries and incorporates dyadic characteristics that affect internal performance levers. It thereby links performance outcome, internal performance levers, and network conditions through causal relationships. The framework takes into account existing interdependencies between different firms within a manufacturing network and thus provides the means for more realistic sustainability assessment.

8.1 Introduction

In the modern industrial context, there is a growing tendency to capitalize on business networks as a means to ensure commercial gains and to realize competitive advantage [3]. In this paper, we introduce the principal components of a governing framework for sustainability performance assessment that can address network interdependencies. The structure of the framework is derived from the necessity to capture not only internal performance levers—the EFQM Excellence Model and DIN 14031 (standard no. 14031 by the German Institute for Standardization) call these "enablers" or "management performance indicators," respectively—and performance outcomes—"results," as they are called in EFQM, or "operational performance indicators," as referred to in DIN 14031—but also to

J. E. Beer (✉) · J. P. Liyanage
Universitetet i Stavanger, 4036 Stavanger, Norway
e-mail: jakob.e.beer@uis.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_8,

include an assessment of the most important dyadic characteristics with respect to a firm's ability to trigger changes beyond the narrow boundaries of its own organization, thereby involving other relevant parties in the network. Including an assessment of network relationships has gained importance due to the development toward network-based value creation in several manufacturing sectors [3]. Nevertheless, we have not come across any framework for sustainability performance that grants network relationships the importance it deserves. The literature both on performance management [30] and on sustainability (e.g., [9, 17, 31]) is predominantly focused on internal actions of the firm disregarding manufacturing partners, service companies, and most other stakeholders who may well have decisive impact on the firm.

Acknowledging that there is a lack of frameworks that take into account performance dynamics between network actors, we have sought to identify the relevant items that can be used to derive meaningful conclusions from an assessment of dyadic network relationships. Literature on social networks and organizational theory (such as transaction cost theory) prove to be helpful in gaining a better understanding of the impact of network membership on individual firm performance. In addition to the network view, we identified internal performance levers that would influence a firm's sustainability performance.

8.2 Framework Design

8.2.1 Network Impact on Firm Performance

The framework consists of three main performance categories reflecting (1) *the network conditions*, (2) *internal performance levers*, and (3) *performance outcome*. The second and the third categories can also be found in other frameworks such as EFQM ("enablers" and "results") and ISO 14031 ("management performance indicators"/MPI and "operational performance indicators"/OPI). The underlying concept—the distinction between internal performance levers and performance outcome—is similar as these frameworks try to separate the actions that are actually performed in the organization from the results of these actions that can be measured in common terms such as financial returns, water usage, CO_2 emission, or employee health.

The question remains how the network the firm is related to can be taken into account. There are two challenging questions arising when network relationships are to be included in performance assessment: First, based on what criteria can we decide whether or not an organization belongs to the network? Second, how can we separate performance effects on the firm resulting from its network participation from effects resulting from other influencing factors? The answer to the first question depends on how we define "network." The case is rather clear if only those entities are considered part of the network that have contractual relations with other firms. Where this is not the case, network membership is difficult to define [34]. Some authors propose a broader definition and emphasize informal social properties as defining feature. Network membership would then be about binding social contracts between the different actors; although legal contracts may exist, their importance is considered as secondary [20]. Uzzi [45] has observed the two types of network relationships existing simultaneously. On the one hand, firms maintain arm-length relationships with suppliers and service companies characterized by ad hoc transactions and individual effort that aims at the fulfillment of contractually defined requirements. On the other hand, firms in a network may exchange more fine-grained information on a more frequent basis beyond what is legally required with trust being one of the most important properties of the (long-term) relationship. The latter refers to what Granovetter [12] refers to as embeddedness. Johnsen et al. [18] point out that reasoning behind actions in network relationships can easily be confused and what seems like altruistic "partnership" behavior beyond the narrow boundaries of the legal contract might rather be explained by other variables. In the automobile industry, for instance, it is not unusual that OEMs support suppliers that are facing financial problems and agree to temporary price hikes [44]. This behavior could mistakenly be interpreted as partnership loyalty, whereas it is a mere necessity for OEMs to avoid supply shortages. Some authors consider a common purpose of their actions the defining property of firm networks [34], and coordination of their actions also provides argumentative ground for firms to represent networks [1]. After all, there is no definitive answer to the question of boundary specification and it depends on the specific purpose of the research or the undertaking how network boundaries are to be defined. In terms of performance management, the decision rests on the judgment of the responsible manager who needs to have a good understanding of the interrelations and interdependencies in "his" network. The specification of boundaries has considerable influence on the structural properties of the network [23] and may thus alter results of performance assessment significantly.

The second question addresses the performance impact on the firm resulting from its network membership as opposed to other influencing factors. Here again, there is no precise answer or general wisdom that could guide the process of performance assessment. The problem of identifying the performance outcome of network membership is further complicated by the question of network boundary specification as discussed in the previous paragraph. Firms can benefit from their network membership in several ways, such as better quality work and higher reliability of network partners as compared to other transaction-based ad hoc business partners [45], better access to information and other resources, and status benefits from highly reputed network partners [15]. It remains unclear how these benefits can be measured in terms of improved outcome. The same problem applies to risk resulting from network membership. Also, firms may not just belong to one but to several networks at the same time, all of which can affect individual firm performance simultaneously. Firm survival rates and stock market effects have been used as proxy to measure effects of network membership [15]; however, these indicators are hardly sophisticated enough and barely reliable.

Donabedian [6–8] has proposed a framework that includes the structure of the organization, the internal processes, and the outcome. His concept rests on the idea that the structure determines the effectiveness of the processes which, in turn, lead to a certain outcome. In order to explain performance results measured in the category "outcome"—which do not tell much by themselves—it is necessary to assess the underlying processes that led to the particular outcome. When the processes are to be judged with respect to their effectiveness, it is necessary to take into account the underlying organizational structure that partly determines how the processes are shaped. With respect to the importance of network partners in many manufacturing environments, an analogous concept can be used to integrate a network perspective with the performance framework. Provan and Sydow [36] have attempted to modify Donabedian's framework to match it with network requirements. When there is a strong correlation between positive attributes in structure and process on the one hand and outcome on the other hand, then this suggests that performance gains can be traced back to network participation. In the opposite case, when performance gains can be reported without particularly strong structure or process indicators, then performance gains are likely to be independent of network participation [36]. Hence, the framework attempts to identify and measure the performance impact of the network, albeit not directly. We have adopted this concept and modified it according to the requirements we identified in our research.

8.2.2 Network Conditions

The first category—network conditions—includes three performance dimensions: *objective alignment*, *capability matching*, and *partnership health*. These dimensions characterize the condition of the dyadic relationship to other network actors and are meant to provide an impression of the network effectiveness. Effective network relationships are considered a requirement for significant changes toward better sustainability performance. The rationale is that manufacturing firms that find themselves embedded in a network [12] are limited in their opportunities to achieve considerably better sustainability performance if these depend on the goodwill, the capabilities, and the objectives of manufacturing partners. Regarding the extent to which manufacturing happens in the network context with multiple different firms involved, it seems crucial to involve (or, at least, consider) network partners in the decision-making process. The category network condition covers the relationships to network partners and thereby indicates the potential for collaborative performance improvements.

Partnering firms are most effective when they have "complementary but distinctive sets of innovative capabilities" [15]. *Capability matching* is a measure of how well firms in the network are able to match each other's requirements with the capabilities they can provide. It also describes their ability to communicate on the same level with the same language and the same understanding. If circumstances change, companies need to adapt. It may be difficult for firms to keep up effective cooperation if they adapt at a different pace [4]. In effect, capability matching describes the partners' competencies to innovate and change, or more general: the basis to achieve their common objectives. If capabilities provided by one firm do not match capabilities required by another, transaction costs will increase if both entities are tied to each other through a binding contract or if switching costs are high. If certain achievements in terms of sustainability performance depend on another firm's contribution—as it may be the case if an OEM decides to change its products' characteristics toward higher durability and use of non-toxic and certified material and thus needs a major supplier to change some of the components it supplies—then in the worst case all internal effort can be rendered useless if the supplier is not able to provide the necessary capabilities.

The competencies to achieve common objectives are worthless; however, if the network partners in fact diverge in their objectives. Hence, the second dimension is *objective alignment*. Objectives among firms in the network can be identical, complementary, independent, or conflicting. Objective alignment describes how well the firms in the network are able to align their internal objectives with their partner firms' objectives. Common objectives may require that incentives and motivation across the network are aligned as well. Since objectives among different partners may be conflicting, alignment with one partner may lead to conflict with another which creates the need for supportive guidelines or tools to trade-off objectives. Aligning incentives in manufacturing networks requires managers to embrace the idea that both different firms and different functions within firms are driven by differing motivations [29]. Though certainly not all objectives can be stated in clear terms as some objectives are of rather vague nature, a clear definition of the main objectives is vital as they represent the foundation of performance management. Performance is always described as the relationship between expectations (as represented by objectives) and deliverables. Referring to Kilduff and Tsai [24], Provan and Kenis [35] call networks that are constituted based upon common goals "goal-directed networks." In these networks, the emerging performance of the network can outdo performance achievements by individual organizations that act independently [35]. Firms can be part of different independent networks, and the firm's objectives may change depending on the requirements each network imposes on the firm. Objectives of one "goal-oriented network" may differ from objectives of another "goal-oriented network." Apparently, such situations can lead to performance trade-offs and tensions but also to synergies [25]. Probably some form of governance is necessary for goal-directed networks to become effective and efficient [35]. Narayanan and Raman [29], for instance, propose monetary incentives to align network objectives.

Some of the stakeholder theory literature is helpful for a better understanding of how objective alignment in network relationships work. Jones et al. [21] have developed the concept of "stakeholder culture" to predict how firms with certain ethical attitudes will respond to stakeholder claims. Network partners are one category of stakeholders, and they confront a focal firm with claims and expectations just like any other stakeholder. Provided with limited resources, the management of the focal company has to prioritize the claims to which it will respond. To prioritize stakeholder claims and make a decision, criteria are necessary. Several authors have proposed different criteria, such as power and interest [10], power, interest and attitude [28], and power, legitimacy, and urgency [27]. Jones et al. [21] extend the decision criteria introduced by Mitchell et al. [27] by providing a framework to predict which of the three criteria for stakeholder prioritization management will value most based upon their stakeholder culture. Their findings may allow to draw an analogy from the more general question of stakeholder salience to the more specific question of network partner objective trade-offs.

The third performance dimension of network effectiveness we propose to characterize the potential for sustainability performance improvements in manufacturing networks is *partnership health*, a measure of quality and health (or the lack thereof) of a dyadic relationship. This dimension addresses the socio-emotional layer of the dyadic network relationship, sometimes referred to as “relational embeddedness” [13, 20, 45]. Trust is one of the key concepts, along with information exchange and common problem-solving motivation [45]. Other authors [5] have tried to identify and describe items that constitute a “healthy” relationship between partnering firms. Because in a “networked” environment many manufacturing firms depend on their partners to a greater extent than they used to in a world of firm-centric value creation [33], they also share risks and gains to a greater extent. A high level of trust will thus be crucial for significant achievements in sustainability performance, especially when risk is involved as in the case of the transition to a new business model (from product sale to service, for instance). Social mechanisms such as care for reputation and collective sanctioning in case of abuse of trust and rules can become major factors for network effectiveness [20]. The level and distribution of risks and gains may be defined if there are formal agreements; however, one of the characteristics of many networks is that there is no such formal bureaucratic structure defining the collaboration and coordination between firms [20]. Gulati [14] refers to this type of relationship as “strategic alliance”; Jones et al. [20] call them “network governance” (other authors use the term “network governance” differently, cf. [11, 35]. In their terminology, network governance describes the mode of coordination between network entities. Atkinson et al. [2] suggest that there are both explicit (as in the case of formal agreements) and implicit contracts that define what firms expect from each other (the authors use the more general term stakeholder to describe the entities of their research; cf. [2, #422:27]. Hence, the distribution of risk and gains may remain rather vague [16]. Also, the visibility of risk in collaborative endeavors may be low and difficult to measure [16]. The importance of shared expectations and beliefs across different network actors is sometimes referred to as “macroculture” [20].

These three dimensions of network effectiveness represent the foundation for any significant achievement in a production setting that involves more than one firm, i.e., in a manufacturing network. Improvements in sustainability performance typically do involve more than one firm if the achievements are meant to be significant. We claim that assessing sustainability performance only locally contributes to sub-optimization. As pointed out earlier, the performance of a firm is influenced by the networks to which the firm belongs [14]; to achieve overall improvement in sustainability performance in manufacturing network, it will be necessary to involve value-adding partners. By making sense of effective network relationships, firms can extend their level of control and thereby reduce uncertainty and vulnerability [32]. It is important to point of, however, that not all internal performance levers may be affected by the network conditions to the same extent (e.g., governance is likely to be less affected than product and service development; see below).

8.2.3 Internal Performance Levers

The second main performance category comprises internal performance levers. Drawing on factors that have been identified as being influential for sustainability performance, five dimensions have been identified: (1) product and service development, (2) strategy and business model, (3) performance management system, (4) organizational culture, and (5) governance.

Product and Service Development: To a great extent, the later sustainability performance of a product or service can be influenced during the development phase [37, 46]. The durability, maintainability, usability, recyclability, and weight of a product are but a few characteristics that are decided on during product development. Cost for changes in the product design and of other main product characteristics may be extraordinarily high once supply contracts and assembly lines have been set up and tools been developed. Therefore, it seems crucial to start considering sustainable solutions early in the process. The more different actors are involved, the more complicated later changes will be. Also, product and service development is an area that demonstrates the importance of well-working network relationships: The expertise and experience of pro-active partner firms in the network can enable the focal firm to steer the development toward much better performance than it would be possible without external support. Vice versa, failure in product development efforts for sustainable product characteristics may be explained by a lack of support or conflicting goals of suppliers and service companies.

Strategy and Business Model: Both the strategy and the business model of a firm can be inherently damaging for the society and the environment. Firms that make their money by producing products that are designed to fail within a short period of time (planned obsolescence) and have to be discarded do neither serve the environment nor the broader society if their production, distribution, and scrapping require large amounts of energy and involve poisonous or rare material. In the course of the economic breakdown following the bankruptcy of several major banks in the USA, several incidents have revealed how part of the (unofficial) strategy and business model of some banks was to sell financial products of no value to unwary investors [43]; throughout most industries, cases can be found of firms that made money based upon unethical or environmentally unfriendly business models and strategies. At the same time, changing a business model may require major efforts not only by the focal firm but also by other firms in the networks being involved in the business. Changing the business from sale of products to sale of services may even lead to severe resistance from suppliers that see their business threatened. This, in turn, demonstrates again how much network relationships influence internal decisions of the focal firm.

Performance Management System: The performance management system creates the incentives scheme for the individuals who are working in the firm. It thereby has strong influence on the individual behavior and the objectives each individual pursuits [41]; "people concentrate on whatever is measured" [26]. The performance management system provides the means for turning strategy into operations [42]. Creating the right incentives to foster the actions intended is not an easy task; literature provides a broad variety of examples for incentive schemes that have led to unintended behavior of individuals. Performance management systems could actively encourage individuals to take stakeholder needs into account and thereby enhance social performance; or they could create incentives for selfish and egoistic behavior and actively limit the opportunities for ethical behavior. Hence, a carefully designed performance management system can be considered an important driver of a firm's sustainability performance.

Organizational Culture: Schein [38, 39] defines

> Culture is a pattern of shared tacit assumptions that was learned by a group as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems.

From this definition it follows that culture has important influence on the way individuals in a firm behave and act. Likewise, organizational culture and attitude influences the extent to which firms are able and willing to partner with other firms in the network [1] (and thus is interrelated with the performance category network conditions). Culture can increase the awareness of individuals in the firm for environmental or social issues; likewise, it can lead to collective ignorance of any sustainability considerations. Therefore, organizational culture can be considered an important internal performance lever for sustainability.

Governance: The type of firm governance may have important impact on the firm's sustainability performance. It makes a difference whether a firm is publicly owned and management is forced to focus on the (short-term) financial interests of anonymous investors or whether the firm is family-owned and management follows the beliefs and ethical guidelines of a strong patriarch who stays involved in the daily business and wants not only serve his own financial interests but (for instance) the local community. Hence, certain governance structures can foster sustainability performance while others can be an impediment. Johnson and Greening [19] examined the link between types of corporate governance and corporate social performance. They found that pension fund equity, for instance, was positively correlated with corporate social performance, whereas investment bank funds were not.

8.2.4 Outcome

This category is meant to cover the effective change of sustainability performance outcome. It represents indicators commonly associated with sustainability performance measurement, such as CO_2 emission, water usage, or spills of poisonous substances in terms of environmental performance, number of incidents with personal injury, stakeholder complaints or law suits, employee satisfaction, or compliance with ethical sourcing guidelines in terms of social performance, and common financial indicators to cover economic sustainability. As explained earlier, such indicators do not provide valuable information if not they are not set into a greater context with reference to underlying internal performance levers. It is necessary to have established the logical (cause and effect) links between the three categories and their indicators so that meaningful conclusions can be drawn from the assessment of the outcome [6, 22]. Systems are made up of several parts that do not function isolated but are closely interlinked; a random set of elements does not characterize a system. Hence, a set of indicators, regardless of its breadth, cannot represent the performance of a system without the indicators being interlinked with each other [40]. The character of the indicators in this category is lagging in nature, that is, they change after internal performance levers have changed. The category outcome can build on a broad variety of performance indicators developed for and employed by popular existing frameworks.

Figure 8.1 illustrates the framework and its principal components.

8.3 Summary

In this paper, we have introduced the principal components of a framework for sustainability performance assessment. The framework builds on the design employed by other frameworks such as EFQM and ISO 14031 and separates internal performance levers from outcome assessment. Since network relationships have an increasingly strong influence on the internal levers which firms can use to enhance their sustainability performance, the framework is complemented by a category comprising the network relationships of the firm. We argue that the performance of internal performance levers cannot be adequately explained without reference to the network within which the firm is embedded. Likewise, performance indicators measured in the category outcome cannot be adequately explained without reference to internal performance levers. The dimensions employed in the two performance categories network conditions and internal performance levers are based upon the identification of potential performance

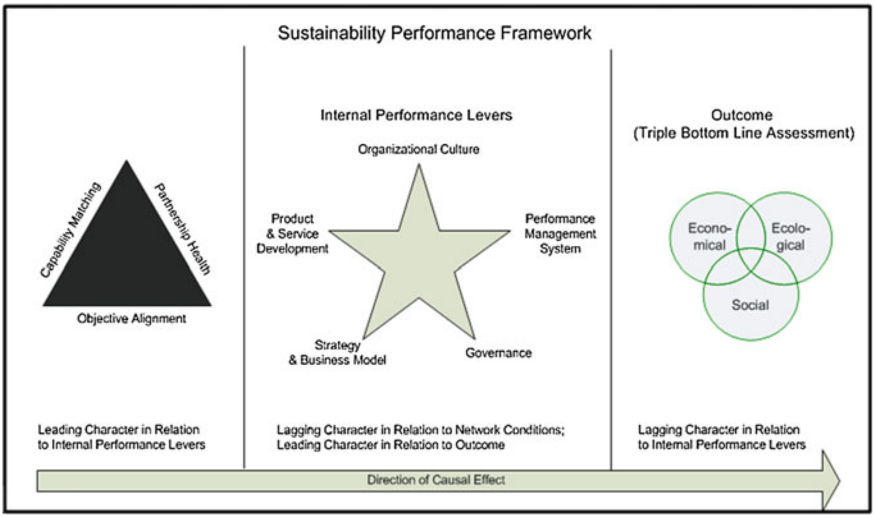


Fig. 8.1 Sustainability performance framework with its principal components

impact factors that have been identified in a first step before they have been categorized in a second step. The dimensions in network conditions include objective alignment, capability matching, and partnership health. The internal performance levers we found are governance, organizational culture, performance management system, product and service development, and strategy and business model. For the outcome, no categories have been proposed as the indicators employed will vary to a great extent depending on the particular industrial setting of the firm.

The framework introduced in this paper represents a novel approach that is currently being applied to industrial firms of various industries to test its validity and applicability. More research will be done to fill the categories and items of the framework with performance indicators and to create guidelines for the use of the framework.

Acknowledgments The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°262931.

References

1. Alexander ER (1995) How organizations act together: interorganizational coordination in theory and practice. Gordon & Breach, Luxembourg
2. Atkinson AA, Waterhouse JH, Wells RB (1997) A stakeholder approach to strategic performance measurement. Sloan Manage Rev, pp 25–37 (Spring)

3. Beer JE, Liyanage JP (2012) Sustainability performance in modern industrial context: from firm-centric to network-based assessment. In: Towards integration and interoperability in engineering asset management. Proceedings of the 7th world congress on engineering asset management. Springer
4. Crook C (2009) Complexity theory: making sense of network effects. In: Kleindorfer PR, Wind Y (eds) The network challenge: strategy, profit, and risk in an interlinked world. Wharton School Publishing, Philadelphia, pp 207–223
5. Dent SM (2004) Partnering intelligence: creating value for your business by building strong alliances. Davies-Black Publishing, Palo Alto, CA
6. Donabedian A (1988) The quality of care: how can it be assessed? J Am Med Assoc (JAMA) 260(12):1743–1748
7. Donabedian A (2005) Evaluating the quality of medical care. Milbank Q 83(4):691–729 [Reprinted from Milbank Memorial Fund Q 44(3, Pt. 2):166–203 (1966)]
8. Donabedian A, Bashshur R (2002) An introduction to quality assurance in health care. Oxford University Press, Oxford
9. Esty DC, Winston AS (2006) Green to gold: how smart companies use environmental strategy to innovate, create value, and build competitive advantage. Yale University Press, New Haven
10. Freeman RE (1984) Strategic management: a stakeholder approach. Pitman, Boston
11. Gereffi G, Humphrey J, Sturgeon TJ (2005) The governance of global value chains. Rev Int Polit Econ 12(1):78–104
12. Granovetter MS (1985) Economic action and social structure: the problem of embeddedness. Am J Sociol 91(3):481–510
13. Granovetter MS (1992) Problems of explanation in economic sociology. In: Eccles RG, Nohria N (eds) Networks and organizations: structure, form and action. Harvard Business School Press, Boston, pp 25–56
14. Gulati R (1998) Alliances and networks. Strateg Manag J 19:293–317
15. Gulati R, Dialdin DA, Wang L (2002) Organizational networks. In: Baum JAC (ed) Blackwell companion to organizations. Blackwell Publishers, Oxford, pp 281–303
16. Harland C, Brenchley R, Walker H (2003) Risk in supply networks. J Purchasing Supply Manag 9(2):51–62
17. Hitchcock D, Willard M (2006) The business guide to sustainability: practical strategies and tools for organizations. Earthscan, London, Sterling (Reprint 2007)
18. Johnsen TE, Lamming RC, Harland CM (2008) Inter-organizational relationships, chains, and networks: a supply perspective. In: Cropper S, Ebers M, Huxham C, Smith Ring P (eds) The Oxford handbook of inter-organizational relations. Oxford handbooks. Oxford University Press, Oxford, pp 61–89
19. Johnson RA, Greening DW (1999) The effects of corporate governance and institutional ownership types on corporate social performance. Acad Manag J 42(5):564–576
20. Jones C, Hesterly WS, Borgatti SP (1997) A general theory of network governance: exchange conditions and social mechanisms. Acad Manag Rev 22(4):911–945
21. Jones TM, Felps W, Bigley GA (2007) Ethical theory and stakeholder-related decisions: the role of stakeholder culture. Acad Manag Rev 32(1):137–155
22. Kaplan RS, Norton DP (2004) Strategy maps: converting intangible assets into tangible outcomes. Harvard Business School Press, Boston, Mass
23. Kenis P, Oerlemans L (2008) The social network perspective: understanding the structure of cooperation. In: Cropper S, Ebers M, Huxham C, Smith Ring P (eds) The Oxford handbook of inter-organizational relations. Oxford handbooks. Oxford University Press, Oxford, pp 289–312
24. Kilduff M, Tsai W (2003) Social networks and organizations. SAGE, London, Thousand Oaks
25. Liyanage JP, Beer JE, Valkokari K, Macchi M, Rana P, Short S, Evans S (2012) Multi-objective sustainability-performance framework. SustainValue Deliverable 4.1

26. Maskell BH (1991) Performance measurement for world class manufacturing: a model for American companies. Productivity Press, Portland
27. Mitchell RK, Agle BR, Wood DJ (1997) Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts. Acad Manag Rev 22(4):853–886
28. Murray-Webster R, Simon P (2006) Making sense of stakeholder mapping. Reprint from Lucid Thoughts, Lucidus Consulting. PM World Today VIII, 11
29. Narayanan VG, Raman A (2004) Aligning incentives in supply chains. Harvard Bus Rev 82(11):94–102
30. Neely A (2005) The evolution of performance measurement research: developments in the last decade and a research agenda for the next. Int J Oper Prod Manag 25(12):1264–1277
31. Nidumolu R, Prahalad C, Rangaswami M (2009) Why sustainability is now the key driver of innovation. Harvard Bus Rev, pp 57–64
32. Pfeffer J, Salancik GR (2003) The external control of organizations: a resource dependence perspective (Republication). Stanford business classics. Stanford Business Books, Stanford, CA
33. Prahalad CK, Krishnan MS (2008) The new age of innovation: driving cocreated value through global networks. McGraw-Hill, New York
34. Provan KG, Fish A, Sydow J (2007) Interorganizational networks at the network level: a review of the empirical literature on whole networks. J Manag 33(3):479–516
35. Provan KG, Kenis P (2007) Modes of network governance: structure, management, and effectiveness. J Public Adm Res Theor 18(2):229–252
36. Provan KG, Sydow J (2008) Evaluating inter-organizational relationships. In: Cropper S, Ebers M, Huxham C, Smith Ring P (eds) The Oxford handbook of inter-organizational relations. Oxford handbooks. Oxford University Press, Oxford, pp 691–716
37. Rebitzer G (2002) Integrating life cycle costing and life cycle assessment for managing costs and environmental impacts in supply chain. In: Seuring S, Goldbach M (eds) Cost management in supply chains. Physica, Heidelberg, New York, pp 127–146
38. Schein EH (1990) Organizational culture. Am Psychol 45(2):109–119
39. Schein EH (2009) The corporate culture survival guide. Jossey-Bass, San Francisco
40. Seiffert MEB, Loch C (2005) Systemic thinking in environmental management: support for sustainable development. J Clean Prod 13(12):1197–1202
41. Simons R (2000) Performance measurement and control systems for implementing strategy (Text and cases). Prentice Hall, Upper Saddle River
42. Sink SD (1993) Developing measurement systems for world class competition. In: Christopher WF, Thor CG (eds) Handbook for productivity measurement and improvement. Productivity Press, Portland, Oregon
43. Story L (2012) Senior executives at goldman had a role in mortgage unit. The New York Times (Apr 2012)
44. Supplier Business Ltd (2009) Supplying BMW. http://www.supplierbusiness.com
45. Uzzi B (1997) Social structure and competition in interfirm networks: the paradox of embeddedness. Adm Sci Q 42(1):35–67
46. Wimmer W (1999) The ECODESIGN checklist method: a redesign tool for environmental product improvements. In: Proceedings First international symposium on environmentally conscious design and inverse manufacturing. IEEE, pp 685–688

Chapter 9
Bearing Fault Diagnosis Using Attractor Comparison

Mehdi Behzad and Majid Karimi

Abstract The failure in rolling element bearings as an important part of rotary machine can lead to machine breakdown. Finding a good strategy for fault diagnosis can help to reduce maintenance cost and enhance reliability. The S-statistic based on the comparison of two reconstructed attractors (RA's) in the phase space was applied to diagnose different types of faults in bearings (inner, ball, and outer faults). To achieve the aim, measured signals of a normal bearing and several faulty bearings were decomposed to sub-signals by wavelet packet so that frequency of each defect is located in a unique packet. Comparing the result for normal and faulty sub-signals using the S-statistic indicate that the method can be applied for fault diagnosis.

9.1 Introduction

Bearings are one of the most important components of rotary machines. These components are main cause of malfunctions, which result in costly shutdown, poor product quality, and even human injury. Consequently, bearings can reduce reliability of these components, and therefore, a maintenance strategy is essential to keep reliability in a desired level. The first and primitive strategy is called breakdown or unplanned maintenance. The next is preventive or planned maintenance. In this scheme, maintenance is done after a specified time period irrespective of machine health status. Out of place maintenance is the disadvantage of preventive maintenance. The most effective generation of maintenance strategy is

M. Behzad · M. Karimi (✉)
Department of Mechanical Engineering, Sharif University of Technology,
Azadi Ave, Tehran, Iran
e-mail: majid.karimi1389@gmail.com

M. Behzad
e-mail: M_behzad@sharif.ut.ac.ir

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_9,

condition-based maintenance (CBM) which suggests maintenance decision based on the collected information through condition monitoring. Two elements of CBM are diagnosis and prognosis. Fault detection, isolation, and identification are related to the diagnosis field, and fault forecasting before happening is the subject of prognosis [1].

On account of the importance of bearings in industrial applications, many researchers have turned attention to the field of condition monitoring. Different techniques were proposed for fault diagnosis, including time, frequency, time–frequency domain, and soft computing methods. Successful results in time domain have been reported on vibration waveform of defected bearing by root mean square [2], kurtosis [3], skewness, peak value [4], and crest factor [5]. Researchers have presented the success in bearing defect detection by identifying the rotational frequency [6]. Because of fast computational algorithm and flexibility, the application of wavelet transform as a time–frequency method is common [7].

Rapid progress of computers and nonlinear theories in two past decades brought new methodologies which can be employed to identify and forecast complex nonlinear vibration behavior. Correlation dimension (CD) as a chaotic indicator was applied for bearing fault identification [8, 9]. A method is introduced to detect weak periodic signal and utilized this method for bearing and gear fault identification [10].

In this work, wavelet packet was used to decompose signals. These decomposed sub-signals were compared using a method (the S-statistic) introduced by Diks et al. [11] for bearing fault detection.

9.2 Experiment and Theory

9.2.1 Wavelet Analysis

Discrete wavelet is an appropriate signal processing tool which can be used for decomposition, compression, and feature selection. The time series $x(t)$ can be decomposed by the wavelet transform into various structures of the orthogonal signal components, the approximation sub-signals $A_J(t)$, and detail sub-signals $D_J(t)$. The original signal passes through two complementary filters and decomposes at the first level into two sub-signals, and decomposition process is repeated until the desired decomposition level J is reached. In this work, it is determined so that frequency of each defect falls in a unique packet. The formulae for calculating the fault frequencies were presented in [12]. The frequency band of each packet is as follows:

$$\left[n\left(\frac{f_s}{2^{J+1}}\right),\ (n+1)\left(\frac{f_s}{2^{J+1}}\right)\right] \tag{9.1}$$

where f_s are decomposition level and sample frequency, respectively.

9.2.2 Attractor Reconstruction

To compare defected bearing signals with healthy vibration signals, it is essential to first reconstruct two phase space attractors from both time series. The simplest method to reconstruct an attractor is the method of time delay (MOTD) which was applied by many researchers [13]. The state space was reconstructed from a scalar time series by copying time-delayed components of the original time series in the reconstructed state space. A sampled time series, $x(i), i = 1, 2, \ldots, N$, can be transferred into a set of $N - (d-1)\tau$ vectors, $\mathbf{s}(i)$, in the state space as follows:

$$\mathbf{s}(i) = [x(i),\ x(i+\tau),\ x(i+2\tau),\ \ldots,\ x(i+(d-1)\tau)] \tag{9.2}$$

where d is the number of elements of the reconstructed state vector which is equal to the number of coordinates in the reconstructed state space and is called the embedding dimension and τ is the embedding delay. It should be noted that the total number of state vectors (points) is equal to $N - (d-1)\tau$. Determining the embedding parameters, time delay and embedding dimension (τ, d) are the most important steps in the nonlinear time series which can be estimated by different methods including autocorrelation function (ACF), mutual information function (MIF) for time delay, and correlation dimension for embedding dimension [14].

9.2.3 The S-statistic

RAs of two independent time series were compared using a discriminating statistic proposed by Diks et al. [11] to reject or approve the null hypothesis. These signals were called reference signals and evaluation signals. The difference was expressed as the S-statistic which is defined by Disk et al. [11] as follows:

$$S = \frac{\hat{Q}}{\sqrt{V_c(\hat{Q})}} \tag{9.3}$$

where $\hat{Q}$ and $V_c(\hat{Q})$ are an unbiased estimator of squared distance of two delay vectors in the state space and variance of unbiased estimator of $\hat{Q}$, respectively. The S value is a random variable with zero mean and standard deviation equal to unity according to the null hypothesis. The null hypothesis is rejected with 95 % confidence level when the estimated value of S is larger than 3 which indicate that these two RA do not show similar distributions [11].

9.2.4 Experiment

Vibration data used in this work obtained from dataset of rolling element bearings [15]. Single-point faults were introduced to the test bearings using EDM with fault diameter of 0.007, 0.014, and 0.021 in, and fault depth is 0.011 in. Drive-end bearing type is 6205-2RS JEM SKF, deep groove ball bearing. Vibration signals were collected using a 16 channel DAT recorder with sample rate 12 kHz. A number assigned to every signal, which can be found in Table 9.1.

9.3 Result and Discussion

As mentioned above, *ACF* and *MIF* can be applied to specify the time delay parameter. The time delay for the attractor reconstruction, τ, is then taken at a specific threshold value of ACF, where a linear independence is found. The recommendation for this threshold is the first value of τ which ACF is equaled to zero. In case of MIF, the time delay parameter is known as the minimum mutual information criterion. While the ACF compares linear dependence, the value of MIF provides a measure of nonlinear dependence and a better estimate of time delay. Figure 9.1 shows ACF and MIF profiles versus time delay for vibration signals numbered in Table 9.1. The first pass of the ACF occurs from zero, and first minimum of the MIF occurs near 2. Therefore, Fig. 9.1 indicates that time delay can be taken 2 because of similar results of both ACF and MIF methods.

The correlation dimension can be calculated from the power law relation between the correlation integral, C, of an attractor and a neighborhood radius, r. Correlation dimension, which is the slope of the log plot of C versus r is defined as follows:

$$C_2(r) = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} H\left(r - \left\|\mathbf{x}_i - \mathbf{x}_j\right\|\right). \tag{9.4}$$

Correlation integral of the normal and faulty signals of bearing listed in Table 9.1 is plotted against hyper-sphere radius for a range of embedding dimensions in Fig. 9.2. The time delay τ was assumed to be 2 in this figure. Different curves in these figures are shown with embedding dimensions $d = 4$ to $d = 20$. As can be seen in Fig. 9.2a–d, when the embedding dimensions increase, the slopes of these curves reach to fixed values which defined as saturation state. These curves show a rigid saturation for higher embedding dimensions of about 15.

Comparison between experimental signals of normal and faulty bearings is not so simple because of noise existence in the measured signals. Therefore, wavelet packet as an appropriate tool for signal filtration can help to decompose measured signals into different frequency range so that frequency of each fault located in a unique packet. Then, obtained sub-signals from wavelet packet corresponding to

Table 9.1 Signal number

Defect type	Fault diameter (in)	Fault frequency (Hz)	Signal number
Inner	0.007	159.93	1
	0.014		2
	0.021		3
Ball	0.007	139.20	4
	0.014		5
	0.021		6
Outer	0.007	105.87	7
	0.014		8
	0.021		9

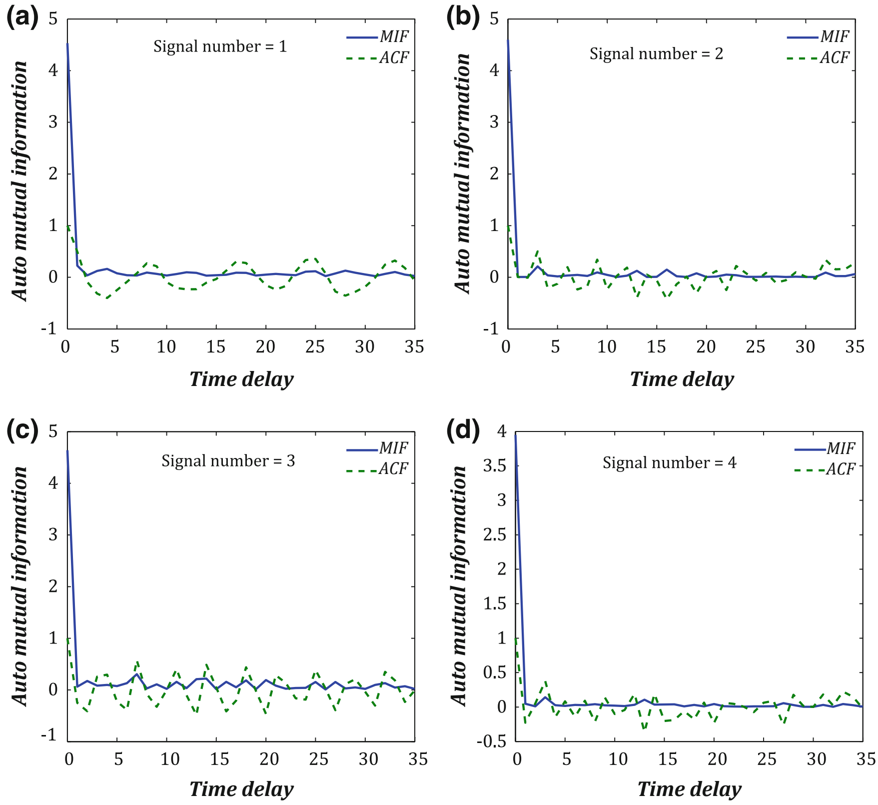


Fig. 9.1 ACF and MIF versus time delay for signals numbered in Table 9.1

the specified fault were compared to those of normal sub-signals. In order to decompose the signals into different levels, Symelts 8 wavelet was chosen as the optimal wavelet for decomposing signals. This mother wavelet showed the

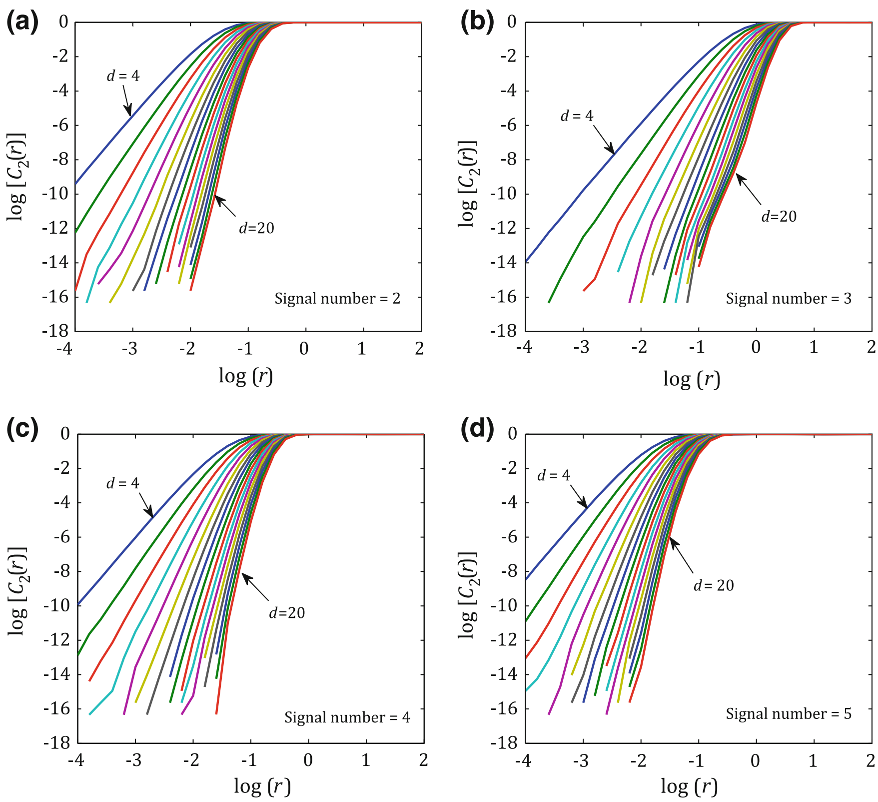


Fig. 9.2 Correlation integral versus r for signals numbered in Table 9.1

minimum reconstructed error after examining various types of Symelts and Daubechies wavelet. Number of level decomposition was assumed to be 9 to position each defect frequency in a unique level.

Figure 9.3a illustrates the result of the comparison between normal and faulty sub-signals (reference and evaluation time series, respectively) within the frequency range of 140.62 and 187.50 which contain inner defect frequency 139.20. As can be seen in Fig. 9.3a, the S values of signal numbers 1, 4, and 7 are greater than 3, and it can be concluded that these bearings are defected and the null hypothesis (i.e., similarity in the distribution of two sub-signals) is rejected. Furthermore, the estimated S value of the mentioned sub-signals increases and then decreases with defect diameter which can be attributed to variable path transmission. This result is in agreement with obtained result of [16]. In the case of other sub-signals numbered in x-axis, the S value is less than 3, the differences between these sub-signals and normal bearing are not significant, and they have not inner defect. Figure 9.3b shows the S values estimated using the comparison of normal sub-signals (considered as the reference time series) with the

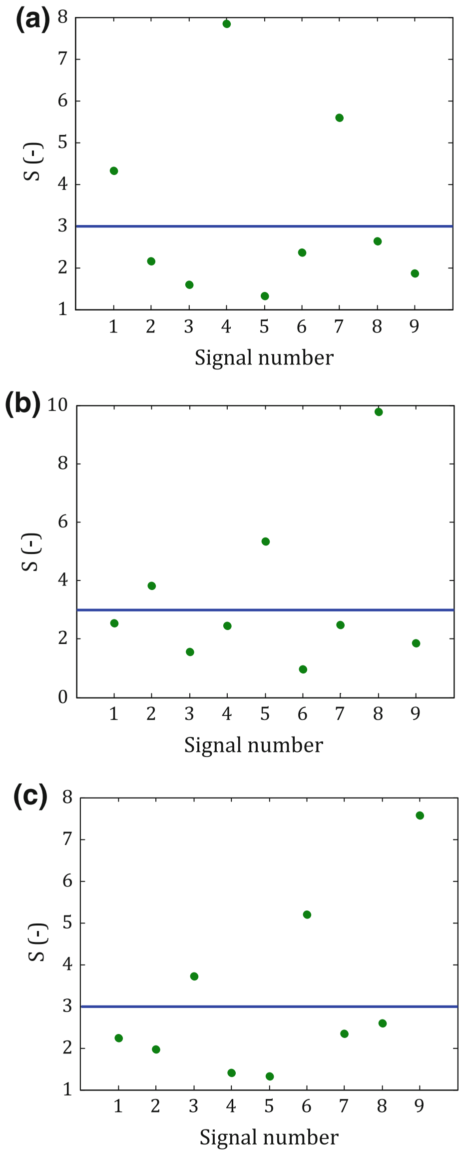


Fig. 9.3 The S value obtained from comparison of (**a**) ball, (**b**) inner, and (**c**) outer faults numbered in Table 9.1 with normal sub-signals

corresponding sub-signals of the defected ball bearings which contain defect frequency equal to 139.2. The frequency range of the compared sub-signals changes from 117.87 to 170.62. As demonstrated in Fig. 9.3b, the S value for faulty sub-signals numbers 2, 5, and 8 are greater than 3 which it means that they have different probability distributions, these bearings are defected, and the S value increases with defect diameter. The S values for the remaining sub-signals are smaller than 3 which can be concluded that the difference between them and normal bearing is not significant and there are not ball defect in these bearings. Figure 9.3c shows the result of the comparison between normal and faulty sub-signals within the frequency range of 93.75 and 117.18 which contain outer defect frequency 105.87. Results indicate that the S values of signal numbers 3, 6, and 9 are greater than 3 and raises with defect diameter. Since outer race does not rotate, changes in the S values of outer defect are more regular and uniform relative to ball and inner race cases. Other conclusions of the sub-signals with the S value smaller than 3 is the same results as Fig. 9.3a and b.

9.4 Conclusion

A combination of wavelet packet and the S-statistic can be used as an effective method for indicating defected state of a bearing. To investigate the state of a bearing, each signal was decomposed into three sub-signals. Sub-signals of normal and faulty bearings were selected to be reconstructed their attractor in the phase space and compared their distribution by the S-statistic. The S values indicated that can identify a defected bearing in comparison with a sub-signal of a normal bearing. When both sub-signals were related to undamaged bearings, the S values became less than 3. The S values have also shown to be sensitive to the increase of the fault diameter.

References

1. Jardine AKS, Banjevic D, Lin D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech Syst Signal Process 20:1483–1510
2. Tandon N (1994) A comparison of some vibration parameters for the condition monitoring of rolling element bearings. Measurement 12:285–289
3. Williams T, Ribadeneira X, Billington S, Kurfesss T (2001) Rolling element bearing diagnostics in run-to-failure lifetime testing. Mech Syst Signal Process 154:979–993
4. Mathew J, Alfredson RJ (1984) The condition monitoring of rolling element bearing using vibration analysis. Vibr Acoust Stress Reliab Des 106:447–453
5. Ingarashi T, Noda B, Matsushima E (1980) A study on the prediction of abnormalities in rolling bearing. JSLE 1:71–76
6. Igarashi T, Hamada H (1989) Studies on the vibration and sound of defective rolling bearings (first report: vibration of ball bearings with one defect). JSME 25:248–253
7. Ericsson S, Grip N, Johansson E, Persson LE, Sjoberg R, Stromberg JO (2005) Towards automatic detection of local bearing defects in rotating machines. Mech Syst Signal Process 19:509–535
8. Logan D, Mathew J (1996) Using the correlation dimension for vibration fault diagnosis of rolling element bearings—I. Basic concepts. Mech Syst Signal Process 10:241–250
9. Logan D, Mathew J (1996) Using the correlation dimension for vibration fault diagnosis of rolling element bearings—II. Selection of experimental parameters. Mech Syst Signal Process 10:251–264

10. Chongsheng L, Liangsheng Q (2007) Applications of chaotic oscillator in machinery fault diagnosis. Mech Syst Signal Process 21:257–269
11. Diks C, Zwet WRv, Takens F, DeGoede J (1996) Detecting differences between delay vector distributions. Phys Rev 53:2169–2174
12. Randall RB, Antoni J (2011) Rolling element bearing diagnostics-a tutorial. Mech Syst Signal Process 25:485–520
13. Takens F (1981) Detecting strange attractors in turbulence. Lecture Notes in mathematics, dynamical systems and turbulence, Springer, Berlin, Germany, p 898
14. Nayfeh AH, Balachandran B (2004) Applied nonlinear dynamics WILEY-VCH Verlag GmbH & Co. KGaA
15. Loparo KA (2003) Bearings vibration data set, in case Western Reserve University
16. Yiakopoulos CT, Gryllias KC, Antoniadis IA (2005) Spectral effects of the slip motion in defective rolling element bearings. Adv Vibr Eng 4:199–211

Chapter 10
From Mitigation to Adaptation in Asset Management for Climate Change: A Discussion

Srirama Bhamidipati, Telli van der Lei and Paulien Herder

Abstract With more than half of the world's population living in urban areas and the rate of urbanization on the increase, there is a continuous and increasing demand for infrastructure and services in these areas. In addition to this growing demand, climate change adds a whole new dimension that compels a relook at our asset management procedures for the critical urban infrastructure. For a considerable time now, mitigation strategies have played an important role in various policy decisions. There is a recent trend in the scientific, geographic, institutional, and business domains that mitigation alone cannot be seen as an effective strategy to limit our vulnerability to the causes and consequences of climate change. Adaptation strategies must also form a part of our efforts, policies, discussions, and planning procedures. In this paper, we present a discussion on how an equal importance to adaptation strategies is being advocated, in conjunction with mitigation strategies, for infrastructure. We review some relevant literature that promotes both these strategies either individually or collectively in relation to infrastructure and services.

10.1 Introduction

So how and where would asset management be different if we would consider climate change in our asset management procedures? Climate change and its impacts are somewhat similar to events of force majeure and disasters. Events such as these cause closure of services for few days till they are restored for normal, safe use. Asset management procedures do deal with such rare events. With respect to climate change, however, there is a very subtle difference from disaster

S. Bhamidipati (✉) · T. van der Lei · P. Herder
TPM, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands
e-mail: s.k.bhamidipati@tudelft.nl

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_10,

management. These events are now expected to happen more regularly and can no longer be treated as “force majeure” or rare events, and they need to be included in regular operation, maintenance, and restoration procedures of assets and their costing calculations in life-cycle assessments. In layman’s terms, asset management is nothing but an organized and informed management of assets. Asset management is a very recent science, which has built upon vast experiences of the past. Asset management systems and standards such as PAS 55 [1] distinguish different levels of decision making for making policies, strategies, and objectives. Maximizing value and minimizing risks are important drivers for optimization of the asset portfolio and systems [2]. Not surprisingly in the context of high stakeholder expectations, complex technological systems operating in the public realm, and decreasing governmental budgets, asset management is a fast-growing field [3]. In the context of climate change, however, the clock is ticking faster than expected, and we neither have time nor data to build upon from experience. There is an urgent need to organize our assets for an efficient use of resources. Because a climate event can impact a large geographic extent, the amount of assets that need to be managed and maintained is enormous compared to the limited funding made available to maintaining assets. If we can broadly classify the society from top to bottom, it would be: the government, the service providers, the end users, in that order, then mitigation is a bottom-up approach where individuals, manufacturers, or service providers, given an incentive, reduce the causes of climate change as they adhere to the policies formulated. Government policies on mitigation are directions/principles created for the greater interests and welfare of the society, and its role is merely to oversee that the policies are implemented effectively. On a positive note, these directive policies bring competition and innovation among the manufacturers (e.g., emission standards for cars, efficient home appliances) and service providers (e.g., efficiency standards for electricity providers). Adaptation on the other hand is a top-down approach, where the national and local governments have to both create and act on the adaptation policies although this trickles down to the organization responsible for infrastructure. Both mitigation and adaptation are like two sides of a coin, one incomplete without the other. In this paper, we discuss, with respect to three domains, how adaptation is gaining importance in actions against climate change and what it might mean for asset managers.

10.2 The Shift in the Scientific Domain

Climate change adaptation and mitigation are buzzwords that are spreading quickly in science and media alike. A simple search for literature on Google Scholar reveals a very unique picture. The main intention of this paper is to observe the trend, if any, in current research focus in the scientific community and to see if adaptation is gaining importance at an equal pace to mitigation. For this paper, three very simple criterion were used: (a) number of articles available for the phrases “climate change adaptation” or “climate change mitigation”; (b) same as case a but articles refer exclusively either to adaptation or mitigation; (c) same as b but articles also refer to asset management in addition to adaptation or mitigation. The literature was captured for a twenty-year period from year 1992 to year 2011. In Fig. 10.1, the x-axis indicates the year while the y-axis indicates the percentage of change in the number of articles from the previous period. As the representation of absolute number of articles for each year creates a spiky profile, we smoothed the results with a running sum of 5 years (a period), such that the value for the year 1992 refers to the total sum of articles published from the period 1988 till 1992. And year 1993 refers to the total sum of articles published from period 1989 till 1993.

From the selected three cases, we can see a trend in the rate of growth for literature on mitigation and on adaptation. In case a, where the exact phrase of "climate change mitigation" or "climate change adaptation" was searched for, we can clearly see that there is a shift and more growth in the articles published on adaptation than on mitigation from around the year 2001–2002. In case b, where the exact phrase of "climate change mitigation" was made stricter by including only those articles that do not mention anything about adaptation. Similarly, no reference to mitigation was allowed when articles were searched for adaptation. Again, we can clearly see that there is more growth in the articles published on adaptation than on mitigation from around the year 2001–2002. The trend can possibly be linked to the publication of IPCC assessment reports in year 1990, 1992, 1995, 2001, 2007, and particularly the formation of the UNFCC treaty in 1992. In and around these years of publication, there is an increase in the related articles. This can be seen as a correlation but not as a causal relation without further verification. In 2001, IPCC defined adaptation to as "adjustment in natural or human systems in response to actual or expected climatic stimuli or their effects, which moderates harm or exploits beneficial opportunities." This emphasis on adaptation in the third assessment report of IPCC can probably be seen as trigger for rapid growth of articles in favor of climate change adaptation compared to mitigation issues since 2001. From the Google results above, in case c, where we include articles that also refer to asset management in addition to climate change adaptation or mitigation, we get an aligning perspective with case a and case b. Though there is not a visible and conclusive trend in the literature, three important observations can be made from the figure for case c. Firstly, articles on asset management referring to climate change started with almost a lag of about 10 years in comparison to articles on climate change that do not address asset management. Secondly, the first trigger for articles on mitigation and asset management seems to be the first formal international agreement in the form of Kyoto Protocol that generated a great deal of interest in the scientific and manufacturing community. To adhere to the protocol and to be in competition, this is probably the first time that manufacturers and industries had to redesign, innovate, and revamp their process lines as noted in the OECD report [4]. Thirdly, the interest in asset management for adaptation started only in 2002, with a lag of 3–4 years compared to mitigation, and is growing steadily since the last five years.

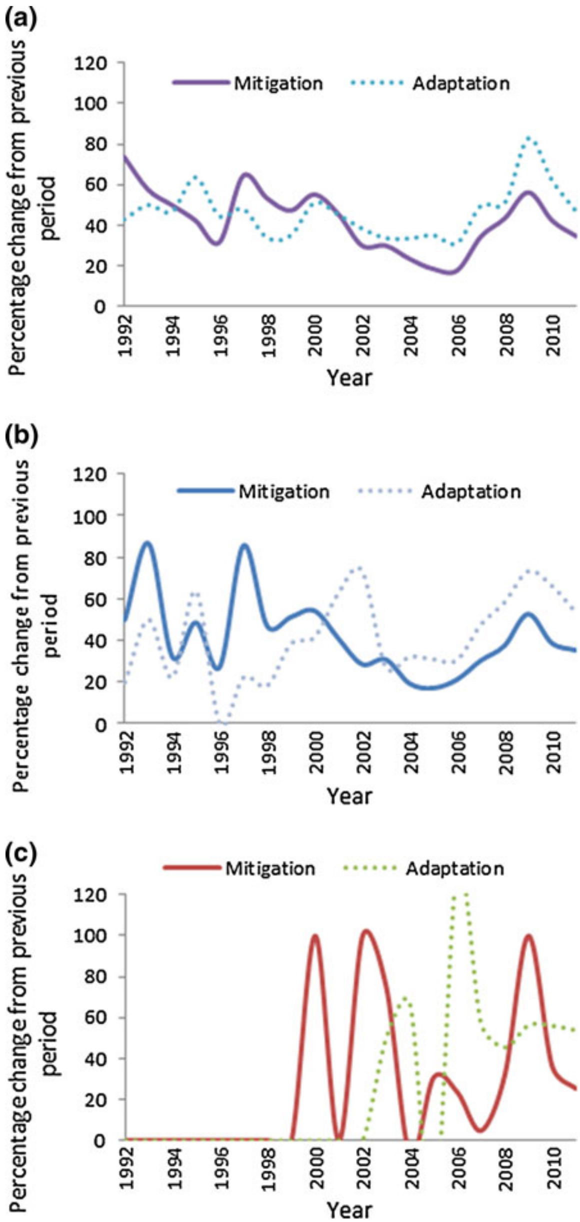


Fig. 10.1 Graphs showing the trends in articles published for **a** number of articles available for the phrases "climate change adaptation" or "climate change mitigation" **b** same as case (**a**) but articles refer exclusively either to adaptation or mitigation **c** same as (**b**) but articles also refer to asset management in addition to adaptation or mitigation

10.3 The Shift in the Geographic and Spatial Domain

In scientific literature and also in various global summits on climate change, the main dispute is about the responsibilities on the causes of emissions and rights of countries to development. That debate is for mitigation, while for adaptation, the discussion is focused on whether to spend on innovative techniques and methods for building resilient systems, adapting existing infrastructure, or whether to provide the basic infrastructure and maximize its accessibility to a wider population. This is a very clear and visible divide. This difference of opinion can be seen between the developed world and the developing world, and within a country, the same debate exists between urban areas and rural areas. These unresolved debates often lead to conflicting policies in urban and regional planning priorities. These differences are also among the main reasons for domestic and international migration causing intense urbanization across the world. The migration for reasons of climate change is being termed as climate mobility and the people as climate refugees, such as those from the Pacific Island of Tuvalu [5]. This migration is considered as an international security threat and also as a possible trigger for conflict for scarce resources in the future [6].

Satterthwaite [7, 8] presents an interesting case with empirical trends from both developing world and the developed world on how population growth cannot be labeled as a driver to increasing emissions, but it is the level of their consumption that is more causal. He argues that the developing countries should be allowed to raise their emissions to a fair-share level as it is not plausible to completely stop the process of urbanization or development, something that is also acknowledged in international agreements. This intense urbanization for reasons of prosperity, survival, safety, or climate puts a lot of stress on the provision, use, and maintenance of infrastructure and services, something that is the real core of asset management procedures. And urban infrastructures with their inherent nature of inter-connectedness and interdependencies make their management very complex and a real challenge for asset managers. And as if these interdependencies are not complex enough, there is also an increasing risk of multiple hazard scenarios because of climate change impacts as reviewed by Li et al. [9]. Corfee-Morlot et al. [10] and Hallegatte and Corfee-Morlot [11] identify that whether it is a debate of urban–rural divide, developed–developing regions divide, or areas within an urban agglomeration, most of the infrastructure management and governance is not integrated and works in silos. They highlight that its governance is often dependent on geographic scope of operations and responsibilities, origin of funding, involvement of different actors, and the channel of scientific information. But for global issues such as climate change, where infrastructure needs to be designed to withstand the expected variability in climate, there is an urgent need to integrate this disconnectedness in interdependencies. A possible multilevel risk governance framework suggested by Corfee-Morlot et al. [10] is a potential solution that can also streamline asset management responsibilities for interconnected assets.

While most of the focus in the scientific literature is on urban areas (both in terms of mitigation and adaption), a considerable part of the literature investigates the impacts of climate change on rural areas and rural economies. A very particular focus is on adaptation of agriculture. Agriculture needs to be sustainable for catering to the demands of urban areas. Agriculture is highly dependent on irrigation and in turn on water resources. Water is also very crucial for our survival, and we need to save it for our future. If we look at most of the literature on adaptation that is not concentrating on urban areas, it mostly deals with agriculture and water resources. For rural areas, the very little that is discussed in literature about mitigation is predominantly dealing with the use of inefficient equipment. Now, agriculture is sustainable as long as the land is irrigable by rivers, canals, or man-made channels. In the context of climate change, the long-term performance of an irrigation system depends on its geographic location, as some areas are predicted to have an increased rainfall while others a reduced rainfall, and therefore, the fertility of the land can change drastically. River deltas are also very fertile and prosperous regions in any country starting from the early civilizations till now. Delta areas run the dual risk of frequent flooding from increased rainfall and also the rising sea levels. Now, urban areas in delta regions, urban deltas, as such form the tip of the climate risk pyramid and demand the highest priority to reduce their vulnerability (adaptation). According to world demographics, 40 % of the population lives in coastal areas, and according to UN estimates, by 2050, 70 % of the world population will be living in urban deltas irrespective of being developed or developing countries.

Next to the debate of urbanization are the dilemmas of the developed and the developing countries. Kropp and Scholze [12] in their manual—"adaptation for developing countries"—summarize this dilemma. This importance is also realized in European Union White Paper on adapting to climate change along with compulsory inclusion of adaptation in all EU's external policies that deal with non-EU countries. Regardless of all the many possible citations, it is evident in almost each available article on adaptation that developing countries will be affected more by climate change than the developed countries. Firstly, due to lack of basic infrastructure, i.e., either in the arid regions of Africa that are facing draught and heat or in the tropical regions of Asia and America that are struggling with floods and rain. Second, their continuing high population growth rates and densities make them more vulnerable. Stephenson et al. [13] state that "large world populations will face greater challenges than smaller one in achieving climate-sustainable emissions" and that could trigger a mass migration as a legitimate adaptation response to climate change, an opinion also shared by Black et al. [14]. According to population statistics from a recent IMF document, about 15 % of the world population lives in the developed world while the remaining 85 % in the developing world, a direct indication of the extent of vulnerability to climate change. In additional to all the discussions on economics, finances, and population dynamics, the main issue catching attention is a new dimension in the fight against climate change, the dimension of human rights. Hall and Weiss [15] put forward a new term "adaptation apartheid" that refers to the differences in adaptation practices in the developed and developing worlds. The authors review and argue the importance of including a legal framework that also considers humans rights for the currently practiced and proposed adaptation practices. Such a framework could bring a balance in climate change investments between the two worlds. Whether it is the issue of urbanization, rural dependence, vulnerability of the developing world, security threat of resource conflicts, or the issue of human rights, adaptation is going to take a driver's seat in the very immediate future.

10.4 The Shift in the Financial and Business Domain

The discussions so far put the developed countries in a negative connotation of duplicity in their approach toward climate change, but that is certainly not true. Current CO_2 levels in the atmosphere are attributed to the emissions by developed countries since industrial revolution. But these countries are reining these emissions to bring it into control as per the international agreements. Secondly, for the first time in 2004, the emissions from the developing countries were higher than those from the developed countries. And as citied in Popp [16], these emissions are expected to exceed those of developed countries by nearly 70 % by 2030. Thirdly, developed countries because of their wealthy status are able to invest in R&D related to climate change. The developing countries on the other hand spend on building their basic infrastructure and mostly depend on the developed countries for the advances, the technology, and knowledge transfer. According to Dechezlepretre et al. [17], about 77 % percent of the climate change-related research is conducted in developed countries while the innovations in the developing countries are more of adaptations of the inventions from the developed world. In a stricter sense, only about 7 % of high-value research is conducted in the developing or emerging economies. Comparing articles that differ by 15 years highlights an interesting fact: Martinot et al. [18] discusses the importance of technology transfer irrespective of the time gap in that transfer, therefore allowing for a smooth flow of trade and an incentive for the inventors to capitalize and financially benefit from the time taken in this transfer. Popp [16] and Dechezlepretre et al. [17] also discuss the need of technology transfer but express the growing concern related to the time gap for diffusion of technology, and also the increasing demand for a completely different paradigm toward intellectual property rights concerned with climate change innovations that can speed up the knowledge transfer. While technology transfer (access to import of embodied equipment) is related to mitigation, the knowledge transfer is more important to the developing countries as it helps them in preparing for adaptation techniques.

Adaptation and mitigation are, however, very different. In mitigation, the purpose is to reduce the emissions, it has metrics and can have targets, directions can be focused, and processes can be streamlined and therefore a delight for asset managers. Mitigation for example with a target of reducing vehicular emissions can be achieved by innovations in redesigning the combustion engine. A car in the Netherlands can be the same as a car in India. Similarly, innovations made in coal-based power generation plants in the developed countries can be employed in the developing countries with a little or no delay. Adaptation is complex, and it is based on climate impacts that are varying, primarily unknown, and unpredictable. Dykes designed in the Netherlands cannot be used in tropical deltas of India. Similarly, rain-harvesting procedures in monsoon-ridden India cannot be practiced in arid regions of Africa. Every adaptation project is different; maintenance procedures are different and may have a very peculiar geographic characteristic and therefore are a real challenge for asset managers and a real opportunity for innovation and progress in asset management as a science.

As discussed in the previous sections: (1) 80 % of the world population is vulnerable to the climate impacts and needs knowledge transfer for adaptation, (2) 70 % of the high-value research is conducted in the developed world, and (3) by 2030, emissions from the developing countries will be 70 % higher and they need technology transfer for mitigation. Combining these even in a very pure business sense, just the economies of scale for knowledge and technology transfer highlight the future prospects of climate change-related activities, especially in adaptation considering the fact that it is a more localized activity compared to mitigation. In a recent report by Fulton [19] which specially focuses on investment opportunities and their risks for asset management practices related to both climate change adaptation and mitigation, they estimate the market to be around US$500b by 2050. This report and others by UNFCC that are focused on guidance to investors on climate change opportunities are from premarket crisis of 2008. The challenge of recovering from the financial crisis and also to take strong action on climate change issues has been a little too harsh on international institutions to make a quick exit from the recession. Lane and Maeland [20] with evidence from the crisis discuss the need of a modified and collective international action against climate change. This will take into account the emergence of changing global powers, therefore enhancing fair trade to empower the developing countries. The authors debate that the crisis has given an opportunity to cleanse the financial systems and to restructure them in favor of green economies. In an interesting investigation, Furrer et al. [21] find that banks have the ability to divert investments in favor of green industries, but they are not entirely involved, because so far climate change initiatives are not creating value for banks and most of the initiatives are being simply supported as any other bank products with a slight change in lending rates.

Most of the initiatives currently supported by banks are mitigation investments that they provide with incentivized lending rates, as mitigation proposals are easier to understand based on their reduction targets. We believe that this trend will change. As discussed previously, mitigation efforts are largely individual-based or company-based who are trying to adhere to the government policies, and risks involved in bank loans are of smaller scale. On the contrary, adaptation actions will largely involve redesign and retrofitting of large infrastructure and along with them is a large amount of uncertainly and unpredictability involved such that funding them is beyond the scope of individuals or companies. These actions have to be taken by government institutions. And for banks, lending to governments is much safer than to lend to individuals. Also considering the size of investments and the scale of infrastructures involved in climate change actions, the banks will start to see these as value creating opportunities. Haigh [22] reviews how climate policies are lacking to provide guidance on clear actionable steps for investors and how financial institutions understand and interpret the climate policies. The author, citing an example from the Carbon Disclosure Project (CDP), argues that the lack of clarity is resulting in misinterpreted climate change metrics (emissions per capita, emissions per annual revenue) and is therefore misguiding the investors further. Although metrics exist for mitigation policies, they are often being misused and for vested interests. On the other hand, adaptation because of its complexity is still lacking reproducible and comprehendible metrics for policy making. There is a huge gap and an immediate need to shift focus toward understanding adaptation, especially with the urgent need for adaptation policies and projects, considering the high vulnerability at stake.

For asset management in mitigation or adaptation, financing them is extremely critical for their success and adoption. Institutions and frameworks such as UNFCC, GEF, World Bank, EU, Kyoto Protocol, USAID, and Adaptation Fund (AF) play an important role in international efforts supporting and financing the anti-climate change policies. In a review by Ayers [23], estimates by these institutions on adaptation efforts for only the low- and middle-income countries range from US$10b to US$86b annually by 2015. The current funding comes via two main routes official development assistance (ODA) and UNFCC. The author discusses the pros and cons of these funding systems and also about the new channel, the adaptation fund (AF) that is exclusive for adaptation efforts. Ayers [23] also discuss two other recent proposals on creating funds for adaptation. One is to introduce a levy on international air travel, and the other proposal is to introduce a levy on maritime bunker fuel. The basic idea behind this levy is to raise revenue from individuals rather than countries and essentially to levy the rich using services that cause emissions. This is expected to discourage emission-causing activities and also to raise funds adaptation. These proposals are still under debate while the GEF and ODA channels being very slow because of the red tape involved. Therefore, with a sense of urgency for adaptation and lack of quick and sufficient funding channels, Durban COP 17 meeting in December 2011 launched a "Green Fund" that promised a transfer of US$100b annually from the developed countries by 2020. This is along the lines of AF, which allows for a local, country-based administration for the implementation of the funding and is also very popular among the recipients. It is therefore important for asset managers to understand various sources of funding for climate change actions and also to contribute by providing feedback to improve the sources as well as their channels of operation.

10.5 The Shift in Asset Management

Much of the research in asset management focuses on information management [2, 3, 24]. This makes sense since public infrastructure agencies engage in extensive data collection activities to support their decision processes at various levels. With the current major increase in the collection and use of data and open data, massive new sources of information can be tapped. These can be used to improve asset performance measurements, asset health monitoring, and asset maintenance planning (preventive as well as corrective). Planning and scheduling is also an important element in asset management. According to Neumann and

Markow [25], performance-based planning of transportation networks is systematic and analytic, building upon the policy in terms of quantifiable objectives, system performance, investment methods, decision-support tools, models for periodic system monitoring, and feedback mechanisms to assess performance trends and identify needed adjustments in investment priorities. Lambert et al. [26] propose a multicriteria analysis with climate change scenarios to reorder the priorities on transportation infrastructure assets based on these new sources (climate change based) of vulnerability. Even in the tight economic periods, experiences show that social and technical sustainability becomes more prominent. Scholars such as Barry and Campbell [27] and Van der Lei [28] show that social and technical sustainability becomes more prominent in asset management practice. Only a few scholars, however, focus on the political and social aspects of managing infrastructure assets. Feunekes et al. [29] provide a rare but excellent example of dealing with the political arena in developing long-term plans for managing a highway infrastructure and to secure commitment from decision makers and also support from the public for these plans. Meyer et al. [30] discuss transportation asset management with a focus on inclusion of risk appraisal for climate change that includes uncertainty, rate of climate change and extent, and severity of disruption. Adger et al. [31] present an excellent review of the softer side of adaptation to climate change. Also other scholars [32, 33] have started to analyze the effect of stakeholder values on technically dominated asset management decisions. We observe that the asset management research arena is shifting focus: from optimizing the (internal) decision-making processes within an organization that includes information management; risk management; and asset monitoring, to creating more value by properly including stakeholder values and greater societal trends such as climate change. Many scholars have not addressed the latter cause of including climate change, yet. To summarize from the above, the main institutional adaptation strategies that need to take place are training of asset managers to understand climate risks, to understand the available and to generate funding sources for climate change actions, and to understand the opportunities and scope of asset management that exists in climate change.

10.6 Outlook

When the three shifts that are discussed in this paper are superimposed, a daunting research endeavor unfolds itself. As argued, specifically on the adaptation side of the climate change discussion, there is and there should be a role for asset managers and owners of large-scale public infrastructures. However, adaptation of infrastructures will go hand in hand with mitigation measures within the infrastructures for example by means of new technologies. Both measures require that asset management approaches, tools, and measurement techniques are improved to manage new risks that come with climate change and with new technologies. And there is a need to develop more dynamic and localized indicators and metrics for the measurement of infrastructure performance under climate change and under various mitigation and adaptation programs. On top of that “internal” need for asset management improvement, the shift in geographic and spatial context of climate change calls for climate adaptation and mitigation approaches that are internationally transferrable and support the paramount value of human rights. This highly topical issue feeds into the knowledge gap in asset management research and practice concerning stakeholder involvement and value incorporation: These are increasingly gaining relevance and momentum in the scientific arena. The incorporation of values calls for different business models and different financing schemes. Such shifts in financing schemes, which were previously predominantly spent to support causes of mitigation, have now given way to exclusive funds for asset adaptation and introduction of resilient infrastructure. But a deep scientific knowledge gap exists on how to deal with such deep uncertainty in asset investments; operation and maintenance, which we think, can be attempted using more recent dynamic and adaptive models such as agent-based models. Some theoretical underpinnings for such models are starting to come up as in Markowitz and Shariff [34] who discuss the underlying reasons for individuals' behavior toward climate change and again as in Meyer and Weigel [35] who propose an adaptive systems approach toward managing transportation assets against climate change impacts. We feel that this paradigm of agent-based models with an inclusion of softer societal dimensions in addition to the traditional technical aspects of engineering is a potential direction that has vast scope of research for studying climate change uncertainties, and it would be a remiss for the asset management community not to devote time in this direction.

Acknowledgments This work is supported by the Knowledge for Climate program in the Netherlands, under the Project INCAH—Infrastructure Networks Climate Adaptation and Hotspots.

References

1. British Standards Institute (2004) PAS 55-1 and -2—Asset management: specification for the optimized management of physical infrastructure asset. The Institute of Asset Management, Bristol, UK
2. Moon FL, Aktan AE, Furuta H, Dogaki M (2009) Governing issues and alternate resolutions for a highway transportation agency's transition to asset management. Struct Infrastruct Eng 5:25–39
3. Brint A, Bridgeman J, Black M (2009) The rise, current position and future direction of asset management in utility industries. J Oper Res Soc 60:106–113
4. Jochem E, Madlener R (2003) The forgotten benefits of climate change mitigation: innovation, technological leapfrogging, employment, and sustainable development. In: OECD Workshop on the benefits of climate policy: improving information for policy makers. Last accessed on 11.04.2013 http://www.oecd.org/env/cc/19524534.pdf
5. Farbotko C, Lazrus H (2012) The first climate refugees? Contesting global narratives of climate change in Tuvalu. Glob Environ Change 22(2):382–390

6. Brown O, Hammill A, McLeman R (2007) Climate change as the ‘new’ security threat: implications for Africa. Int Aff 83(6):1141–1154
7. Satterthwaite D (2009) The implications of population growth and urbanization for climate change. Environ Urbanization 21(2):545–567
8. Satterthwaite D (2010) The contribution of cities to global warming and their potential contributions to solutions. Environ Urbanization Asia 1(1):1–12
9. Li Y, Ahuja A, Padgett JE (2012) Review of methods to assess, design for, and mitigate multiple hazards. J Perform Constr Facil 26(1):104
10. Corfee-Morlot J, Cochran I, Hallegatte S, Teasdale P-J (2010) Multilevel risk governance and urban adaptation policy. Clim Change 104(1):169–197
11. Hallegatte S, Corfee-Morlot J (2010) Understanding climate change impacts, vulnerability and adaptation at city scale: an introduction. Clim Change 104(1):1–12
12. Kropp J, Scholze M (2009) Climate change information for effective adaptation: a practitioner’s manual. Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ) GmbH
13. Stephenson J, Newman K, Mayhew S (2010) Population dynamics and climate change: what are the links? J Publ Health 32(2):150–156
14. Black R, Stephan R, Bennett G, Thomas SM, Beddington JR (2011) Climate change: migration as adaptation. Nature 478:447–449
15. Hall M, Weiss D (2012) Avoiding adaptation apartheid: climate change adaptation and human rights law. Yale J Int Law 3:309–366
16. Popp D (2011) International technology transfer, climate change, and the clean development mechanism. Rev Environ Econ Policy 5(1):131–152
17. Dechezlepretre A, Glachant M, Hascic I, Johnstone N, Meniere Y (2011) Invention and transfer of climate change-mitigation technologies: a global analysis. Rev Environ Econ Policy 5(1):109–130
18. Martinot E, Sinton JE, Haddad BM (1997) International technology transfer for climate change mitigation and the cases of Russia and China. Annu Rev Energy Env 22:357–401
19. Fulton M (2007) Investing in climate change: an asset management perspective. Deutsche Bank Germany. Last accessed on 31.08.2012 from http://cbey.research.yale.edu/uploads/Carbon%20Finance%20Speaker%20Series/Investing%20in%20Climate%20Change%20Yale%204.16.pdf
20. Lane J-E, Maeland R (2010) Global financial crisis and international institutions: challenges, opportunities and change. Publ Organ Rev 11(1):29–43
21. Furrer B, Hamprecht J, Hoffmann VH (2012) Much ado about nothing? How banks respond to climate change. Bus Soc 51(1):62–88
22. Haigh M (2011) Climate policy and financial institutions. Clim Policy 11(6):1367–1385
23. Ayers J (2009) International funding to support urban adaptation to climate change. Environ Urbanization 21:225–240
24. Pantelias A, Flintsch GW, Bryant JW Jr, Chen C (2009) Asset management data practices for supporting project selection decisions. Publ Works Manag Policy 13:239–252
25. Neumann LA, Markow MJ (2004) Performance-based planning and asset management. Publ Works Manag Policy 8:156–161
26. Lambert JH, Wu Y-J, You H, Clarens A, Smith B (2012) Climate change influence to priority setting for transportation infrastructure assets. J Infrastruct Syst. doi:10.1061/(ASCE)IS.1943-555X.0000094
27. Barry D, Campbell JD (2011) Achieving asset management excellence. In: Campbell JD, Jardine AKS, McGlynn J (eds) Asset management excellence—optimizing equipment life-cycle decisions. CRC Press, Boca Raton, pp 379–390
28. Van der Lei TE (2012) Towards a research agenda for strategic engineering asset management. In: van der Lei TE, Herder PM, Wijnia YC (eds) Asset management—the state of the art in Europe from a lifecycle perspective. Springer, Dordrecht, pp 169–172
29. Feunekes U, Palmer S, Feunekes A, MacNaughton J, Cunningham J, Mathisen K (2011) Taking the politics out of paving: achieving transportation asset management excellence through OR. Interfaces 41:51–65

30. Meyer MD, Amekudzi A, O'Har JP (2010) Transportation asset management systems and climate change. Transp Res Rec 2160(1):12–20
31. Adger WN, Dessai S, Goulden M, Hulme M, Lorenzoni I, Nelson DR (2008) Are there social limits to adaptation to climate change? Clim Change 93(3–4):335–354
32. Schraven D, Hartmann A, Dewulf G (2011) Effectiveness of infrastructure asset management: challenges for public agencies. Built Environ Project Asset Manag 1:61–74
33. Too EG (2012) Infrastructure asset: developing maintenance management capability. Facilities 30(5/6):234–253
34. Markowitz EM, Shariff AF (2012) Climate change and moral judgment. Nature 2(4):243–247
35. Meyer MD, Weigel B (2011) Climate change and transportation engineering: preparing for a sustainable future. J Transp Eng 137(6):393–403

Chapter 11
Examining Organizational Culture Within Professional Engineering Asset Management Firms: The Competing Values Framework

Yvonne Brunetto, Matthew J. Xerri and Silvia A. Nelson

Abstract This paper reviews the extant literature about engineering asset management (EAM), which highlights that organizational and cultural change may be required to develop effective asset management practices. As such, this study uses the competing value framework (CVF) and professional theories to reposition the role of professional and organizational culture and support changes in the workplace practices of engineers working in asset management organizations. In particular, a process is proposed for assessing the current organizational culture of asset management firms to determine whether the environment is conducive for supporting improvements in the overall asset management performance. The methods suggested in this paper include interviews and the development of a survey instrument, comprising the OCAI and other relevant questions about EAM employees' work practices. This tool is yet to be tested and applied to examine organizational culture within a professional asset management context. The evidence-based information from this study can be used by management to identify current and preferred work practices. In particular, the information is used to identify whether and what change is required to maximize employees' performance. The study expects to identify new information about best work practices required to support the successful implementation of asset management practices. This information can be used as a platform to provide evidence-based strategies for realigning organizational goals with the goals of the individual employees to improve asset management capability.

Y. Brunetto · M. J. Xerri (✉) · S. A. Nelson
Southern Cross University and CIEAM, Locked Bag No. 4, Coolangatta,
QLD 4225, Australia
e-mail: matt.xerri@scu.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_11,

11.1 Introduction

This study provides insight into the rationale informing the development and use of a diagnostic tool that diagnoses the current organizational culture in asset management firms to identify whether it is optimal for maximizing employee performance in changing organizational practices for engineers. In the engineering context, where skills and expertise are prized, effective asset management is likely to be more dependent upon organizational strategy and culture than the technical expertise of the individual engineers involved. That is to say, while an engineer or group of engineers may have a comprehensive grasp of the tasks and processes for which they are responsible, their performance may be limited and distorted by organizational cultures that do not support the strategic and comprehensive breadth of effective asset management strategies. This research is important because the implementation of organizational change fails more often than it succeeds. Specifically, the majority of organizations that have attempted to change or implement strategies that challenge their existing culture have either failed or found the process too long and laborious to achieve the desired outcomes of the initiative [1]. Beer and Nohria [2] estimate that approximately 70 % of all organizational change initiatives fail and such change often fails during the implementation stage. The implementation of change within an asset management context is particularly important, because failure to deal with the issue of asset sustainability and longevity may entail that the water, power, transport, or sewerage systems that we expect to maintain society are in jeopardy. In particular, asset management includes the management of high capital value assets and therefore, it is important to develop an understanding about how to maximize the efficiency and effectiveness of the management of these utilities and thus improve the overall value of the asset over its entire life. More importantly, developing a process to facilitate the effective management of assets will contribute to avoiding disasters related to the management of assets (e.g., nuclear reactor disasters).

Winter [3] suggests that at the macrolevel, there are four main factors that impact upon change in and implementation of policy, which can be separated into internal and external factors. The external factors include resources and external accountability, and the internal factors include organizational culture and leadership. To add to the argument, some past research suggests that one of the major reasons that most organizational change fails is because employees resist the change process [4–6]. Therefore, this study proposes that because the majority of organizational change fails, it is important to examine, among other things, whether leadership and organizational culture are supportive of the implementation of organizational change or whether they are thwarting the implementation process.

11.2 Engineering Asset Management Research

Turbulent business environments, competitive markets, and the need to maximize the return for shareholders have led to an increase in the need to maximize the value of assets over the life of the asset. In addition, it has been suggested that the automation of industries has increased the reliance of organizations on the maintenance and management of assets [7]. A greater reliance on assets suggests that safety issues and poor maintenance of assets could overall reduce performance or potentially put a stop to organizational operations. Furthermore, the relationship between organizational profit and the management of assets has seen an increase in the importance of strategic asset management within organizations. Strategic asset management is referred to as the alignment of the management of the asset with the demand for the product or service (profit maximization) [8].

Woodhouse [9] suggests that the major issue with current asset management practices and research are centered around human factors as opposed to the technical aspects of asset management. Specifically, Woodhouse outlines that the adjustments to the sophisticated requirements of contemporary asset management requires a change in the skills and capabilities (human factors) of employees. Usually, the problem is not with the technology or systems used to undertake the physical management of assets; instead, the weak link is the management of the employees undertaking the asset management function. However, research examining engineering asset management (EAM) predominantly focuses on the technology and information systems required to physically manage assets [10–13]. Consequently, there is a lack of research examining the impact of employee factors on physical asset management. This study proposes that a diagnosis of organizational culture can provide some very helpful indicators about the current state of the human factors associated with asset management.

The employees' characteristics associated with high-performing engineering asset cultures have become even more important recently, especially when considering that competition within markets is increasing worldwide and there has been several environmental/safety accidents that have had disastrous financial effects on some organizations. As such, this study plans to use the extant literature to develop an instrument to test the organizational culture of EAM organizations, so as to identify what organizational management strategies are required to modify the existing organizational culture to achieve "best-practice engineering culture." One such method for examining organizational culture is the competing values framework.

11.3 Examining Organizational Culture: The Competing Values Framework

The competing values framework (CVF) is a popular method for diagnosing organizational culture with a view to modifying a particular organizational culture. The CVF is the theoretical basis for the organizational culture assessment instrument (OCAI). The OCAI is one of the most widely used tools for assessing and diagnosing organizational culture, having been applied in approximately 10,000 organizations, concerning both public and private sectors including education, health care and NGOs [14, 15]. The instrument was originated by Campbell [16], who constructed 39 organizational effectiveness indicators to measure organizational culture and effectiveness. The 39 indicators were then expressed in the form of a two-dimensional framework developed by Quinn and Rohrbaugh [17]. Cameron and Quinn [18] then derived the OCAI from the earlier competing values approach developed by Quinn and Rohrbaugh [17, 19]. Cameron and Quinn's [18] OCAI is popular within the literature and has been applied and discussed within many studies (e.g., [14, 15, 20–29]). The OCAI developed by Cameron and Quinn [18] has also been applied within an Australian setting [15, 24, 26, 30, 31] and within an engineering context [15].

The OCAI includes a set of survey questions designed to examine organizational culture by categorizing four cultural forms (referred to in some literature as adhocracy, clan, market, and hierarchical) [25]. Cameron and Quinn [18] suggest that the questions within the OCAI are indicators of effectiveness, which represent what people value about an organization's performance. Hence, the indicators provide an organization with an understanding of what their employees perceive as being useful, right, and appropriate in the workplace. The four organizational cultural forms are examined in terms of six dimensions: dominant characteristics, organizational leadership, employee management, organizational glue, strategic emphases, and criteria for success [15, 20, 22, 23, 26, 31]. The OCAI measures organizational culture predominantly using a self-report survey where respondents are asked a series of questions referring to four scenarios within the organization.

The OCAI provides an examination of the competing demands within an organization, between their internal and external environments, and between control and flexibility [14, 23, 26]. Specifically, organizations with an internal focus are more concerned with integration, information management, and communication. On the other hand, organizations focused on the external are more concerned with growth, resource acquisition, and interaction with the external environment. However, organizations with a focus on control are more concerned with stability and cohesion, while organizations with a focus on flexibility are more concerned with adaptability and spontaneity. Additionally, the OCAI has been used to examine the current organizational culture and the preferred culture, using the same OCAI, but asking respondents to fill out the survey for what the current organizational culture is and what the respondent would prefer [15].

To simply assume that engineers are like all other workers is one of the reasons why there has been a failure to implement change within an EAM context, because we did not understand what is required for change within the professional context. There has been one study that has used the CVF to examine engineers within an engineering consultancy [15], which relied solely on the CVF and the OCAI. However, this study did not use the professional theoretical framework as the basis for informing the use or modification of the current CVF in an EAM context. It is important to contextualize the CVF for professionals, and a professional theoretical framework provides one way to achieve such a synergy. Although the OCAI is a very popular tool and is recognized as well developed, valid, and reliable, it does have limitations because it does not cover comprehensively all cultural phenomena (organizational and managerial problems) in organizations [18]. Because organizational culture is complex and some aspects of the overall organizational culture will remain hidden, regardless of what methods are employed, researchers should not rely solely on the CVF and the OCAI. While the CVF provides a useful tool to understand key aspects of the organizational culture and provides for replicability, the specific characteristics of any particular culture in an organization require an extended analysis comprising both qualitative and quantitative methods. This study proposes to develop a robust diagnosis of organizational culture based on mixed methods.

Due to the differences between some occupations and some professions, it may be that the competing values framework requires some additions to be better suited to examine engineers within an asset management context. In particular, there is mounting evidence that professionals have more discretionary power to make decisions in the workplace [32], although the extent to which this power is used to achieve organizational goals is a function of the quality of the relationship between the supervisor and the professional [33]. Hence, previous research about how professionals work and make decisions in EAM organizations must be used to add value to the competing values framework, so that it provides an evidence-based tool for diagnosing accurately the effectiveness of employees within EAM organizations.

11.4 Robust Diagnosis of Organizational Culture

The following mixed methods approach for diagnosis of organizational culture is developed in an attempt to understanding the organizational culture in EAM organizations. First, qualitative methods (e.g., interviews) are necessary in order to understand the key dynamics and processes that govern specific behaviors and performance in EAM organizations, given that there is limited understanding of these issues in the Australian literature of EAM companies.

Second, a soundly based survey of organizational members is necessary to better understand the major issues concerning engineering culture in asset management organizations. The survey will include questions related to demography, OCAI, and other critical questions regarding organizational culture, such as training, teamwork, relationship between supervisor and subordinate, organizational support, role ambiguity, autonomy, turnover intention, commitment, attitude toward change, well-being, and engagement. The survey instrument administered across the organization will provide a meaningful snapshot of the forces at work in the organization that govern key performance outcomes, in this instance, asset management. In this way, the link between organizational culture and performance outcomes can be established and evaluated.

In addition to this mixed methods approach, the analysis of the culture of engineering asset organizations also requires multiple theoretical/conceptual perspectives. The CVF provides one theoretical perspective for understanding the nature of organizational culture and offers a reliable method, using the OCAI, to organize organizational culture types [18]. Specifically, this study proposes that exploration is required into the culture of engineering asset organizations and therefore the behavior and attitudes of professionals. This study suggests that professional theory can be applied to better frame and analyze these issues.

11.5 Summary and Conclusion

The world of asset management is becoming increasingly competitive. An in-depth understanding of the state of the organization culture is an essential contributor to strategy development to achieve organizational outcomes and results. Without in-depth understanding of the real culture (as suppose to the assumed culture) of the organization, it is highly likely that organizational strategies will falter or even fail completely. Many organizations have well-conceived strategies, which they have been unable to implement and it is increasingly apparent that a major cause of failure has been less than robust organizational cultures. Asset management organizations which are most likely to succeed are those which strategically develop organizational cultures that support effective asset management practices and processes. Failure to understand and engage with the organizational culture is a certain recipe for failure, and it is those organizations that honestly and fearlessly seek to understand the values, beliefs, processes, and practices as well as the quality of leadership of the organization are those asset managers who will succeed. The results of this study will make a contribution to studies employing the CVF concept of organizational culture in a professional engineering and asset management context to the more general body of research on organizational culture.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established, and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Langan-Fox J, Tan P (1997) Images of a culture in transition: personal constructs of organizational stability and change. Brit Psy Soc 70:273–293
2. Beer M, Nohria N (2000) Cracking the code of change. Har Bus Rev 78:133–141
3. Winter S (1990) Integrating implementation research. In: Palumbo DJ, Calista DJ (eds) Implementation and the policy process opening up the black box. Greenwood Press, New York
4. Freeze T (2009) Transforming the cynic: recommendations for leaders implementing organizational change. Int J Knowl Cult Change Manage 9:141–154
5. Piderit SK (2000) Rethinking resistance and recognizing ambivalence: a multidimensional view of attitudes toward an organizational change. Acad Manag Rev 25:783–794
6. Vakola M, Nikolaou I (2005) Attitudes towards organizational change: what is the role of employees' stress and commitment? Empl Relat 27:160–174
7. Zuashkiani A, Rahmandad H, Jardine AKS (2011) Mapping the dynamics of overall equipment effectiveness to enhance asset management practices. J Qual Maintenance Eng 17:74–92
8. Barton R, Jones D, Gilbert D (2002) Strategic asset management incorporating ecologically sustainable development. J Facil Manag 1:70–84
9. Woodhouse J (2001) Asset management. John Woodhouse Partnership Ltd
10. Hipkin I (2001) Knowledge and IS implementation: case studies in physical asset management. Int J Oper Prod Manag 21:1358–1380
11. Hodkiewicz M, Pascual R (2006) Education in engineering asset management–current trends and challenges. In: 2nd international physical asset management conference, Tehran
12. Trevelyan J, Tilli S (2007) Published research on engineering work. J Prof Issues Eng Educ Pract 133:300–307
13. Tsang A (2002) Strategic dimensions of maintenance management. J Qual Maintenance Eng 8:7–39
14. Cameron KS (2006) Competing values leadership: creating value in organizations. Edward Elgar Publishing, Cheltnham
15. Igo T, Skitmore M (2006) Diagnosing the organizational culture of an Australian engineering consultancy using the competing values framework. Constr Innov: Inf Process Manag 6:121–139
16. Campbell JP (1977) On the nature of organizational effectiveness. New Perspect Organ Effectiveness: 13–55
17. Quinn RE, Rohrbaugh J (1981) A competing values approach to organizational effectiveness. Publ Prod Rev 5:122–140
18. Cameron KS, Quinn RE (1999) Diagnosing and changing organizational culture: based on the competing values framework. Addison-Wesley, Reading
19. Quinn RE, Rohrbaugh J (1983) A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis. Manag Sci 29:363–377
20. Belasen A, Frank N (2008) Competing values leadership: quadrant roles and personality traits. Leadersh Org Dev J 29:127–143
21. Dastmalchian A, Lee S, Ng I (2000) The interplay between organizational and national cultures: a comparison of organizational practices in Canada and South Korea using the competing values framework. Int J HRM 11:388–412
22. Goodman EA, Zammuto RF, Gifford BD (2001) The competing values framework: Understanding the impact of organizational culture on the quality of work life. Org Dev J 19:58–68
23. Hartnell CA, Ou AY, Kinicki A (2011) Organizational culture and organizational effectiveness: a meta-analytic investigation of the competing values framework's theoretical suppositions. J App Psy 96:677–694

24. Lamond D (2003) The value of Quinn's competing values model in an Australian context. J Manag Psycol 18:46–59
25. Panayotopoulou L, Bourantas D, Papalexandris N (2003) Strategic human resource management and its effects on firm performance: an implementation of the competing values framework. Int J HRM 14:680–699
26. Parker R, Bradley L (2000) Organisational culture in the public sector: evidence from six organisations. Int J Publ Sect Manag 13:125–141
27. Patterson MG, West MA, Shackleton VJ et al (2005) Validating the organizational climate measure: links to managerial practices, productivity and innovation. J Org Beh 26:379–408
28. Yu T, Wu NQ (2009) A review of study on the competing values framework. Int J Bus Manag 4:37–42
29. Zafft CR, Adams SG, Matkin GS (2009) Measuring leadership in self-managed teams using the competing values framework. J Eng Educ 98:273–282
30. Martin J, Simons R (2002) Managing competing values: leadership styles of mayors and CEOs. Aust J Publ Admin 61:65–75
31. Teo STT, Ahmad T, Rodwell JJ (2003) HR role effectiveness and organizational culture in Australian local government. Asia Pac J HR 41:298–315
32. Hoyle E, John PD (1995) Professional knowledge and professional practice. Continuum Int Pub Group, London
33. Farr-Wharton R, Brunetto Y, Shacklock K (2011) Professionals' supervisor–subordinate relationships, autonomy and commitment in Australia: a leader–member exchange theory perspective. Int J HRM 22:3496–3512

Chapter 12
Condition Monitoring of Naturally Damaged Slewing Bearing Based on EMD and EEMD Methods

Wahyu Caesarendra, Buyung Kosasih, Byeong-Keun Choi, Kiet Tieu and Craig A. S. Moodie

Abstract Vibration-based condition monitoring and prognostic of rotating element bearings have been studied. Satisfactory methods on how identify bearing fault and lifetime prediction have been shown in literature. There are two characteristics of the investigated bearings in the literature: (1) The bearings were run in moderate and high rotating speed and (2) damage was artificially introduced, e.g., artificial crack and seeded fault. This paper deals with slewing bearing with very low rotational speed (1 and 4.5 rpm) with natural fault development. Two real vibration data are discussed, namely data obtained from laboratory slewing bearing test-rig and data from a sheet metal company. In this study, the result of EMD and EEMD applied in two real cases were shown superior compared to FFT method.

12.1 Introduction

Slewing bearing is a subgroup of rolling element bearing commonly used in large industrial machinery such as turntable, steel mill cranes, offshore cranes, rotatable trolley, excavators, reclaimers, stackers, swing shovel, and ladle car [1]. The outer diameter of slewing bearing can vary from 200 to 8,000 mm, and they can weigh up to 11.3 tonnes [2]. A slewing bearing typically has three axes (double axial rows and radial row) to support high-axial and high-radial loads. Slewing bearings usually operate at slow speed range from 1 to 15 rpm with continuous or reversible direction.

W. Caesarendra · B. Kosasih · B.-K. Choi (✉) · K. Tieu · C. A. S. Moodie
Department of Energy and Mechanical Engineering, Institute of Marine Industry, Gyeongsang National University, 445 Inpyeong-dong, Tongyeong, Gyeongnam-do 650-160, South Korea
e-mail: bgchoi@gnu.ac.kr

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_12,

Condition monitoring of slewing bearings has been actively researched since the sudden failure of slewing bearing can cause costly breakdown maintenance and impact the productivity of systems utilizing slewing bearing. Slewing bearing is a special bearing where the average lifetime is typically long that range from 5 to 10 years. Due to the high load, dirt environment condition and variable application cause the bearing to fail unpredictably before the estimated lifetime. Up to date, there are few articles discussed the condition monitoring of slewing bearing. This is due to the long expected lifetime and some researcher thought that the slewing bearing is similar to common rolling element bearing which contain the outer race, inner race, and rolling element, but in practice, the vibration signal of slewing bearing is more difficult to analyze than that of common bearing.

Slewing bearing research has been studied from oil analysis [3, 4] and finite element method [5–10]. Earlier article about condition monitoring of slewing bearing based on vibration signal is presented by Rodger in 1979 [11]. This literature explained two methods using vibration signature analysis and acoustic emission source location to detect the incipient failure of heavily loaded anti-friction bearings. In 2010, Liu and Chen [12] briefly reviewed the condition monitoring and diagnosis technique on slewing bearing based on the signal of vibration, temperature, friction torque, acoustic emission, and stress wave. Three established vibration signal analysis methods such as time-domain, frequency-domain, and time–frequency domain were described. In time-domain signal analysis, dimensional features such as mean, peak, and RMS value; and dimensionless features such as crest factor, shape factor, and impulse factor were extracted. In frequency-domain analysis, FFT was used to identify the dominant frequency. In time–frequency-domain analysis, several potential methods were presented, i.e., Wigner distribution, wavelet transform (WT), and Hilbert-Huang transform (HHT). Temperature and friction torque measurement were carried out to enable an additional information on bearing condition monitoring. The temperature and friction torque signal will increase as the failure increases. Signal of acoustic emission (AE) and stress wave were also analyzed. As deformation or crack occur on material, the energy correspond to this phenomenon is released and captured by AE sensor. The shortcoming of this paper is the results of discussed methods were not presented. The latest exceptional published articles in condition monitoring of slewing bearing research were presented by Žvokelj et al. [13, 14]. Reference [13] proposed the combined ensemble empirical mode decomposition (EEMD) and multiscale principal component analysis (MSPCA), and Ref. [14] improved the combined method in previous publication in MSPCA part with kernel PCA. In the papers [13, 14], slewing bearing was run at 8 rpm and a single defect of inner raceway was introduced artificially. The fault signal of artificial single defect is relatively easier to identify than naturally multiple defects.

The aim of this study is to show the performance of EMD and EEMD employed in naturally multiple defects cases to identify the fault occurrences.

12.2 EMD and EEMD

Empirical mode decomposition (EMD) was first introduced by Huang et al. [15]. In the EMD method, the raw vibration signal $x(t)$ was decomposed in terms of intrinsic mode functions (IMFs), i.e.,

$$x(t) = \sum_{j=1}^{n} c_j + r_n \quad (12.1)$$

where n is the number of IMFs, c_j represents the IMFs, and r_n is the final residue of data. IMFs are oscillatory functions with varying amplitude and frequency. According to [15], an intrinsic mode function is a function that satisfies two conditions:

1. Throughput the whole length of a single IMF, the number of extrema and the number of zero-crossing must either be equal or differ at most by one.
2. At any data location, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The process of decomposition of a signal into IMFs is called the *shifting process*. The result of shifting process is IMFs that range from high frequency to low frequency. It has been generally known that the EMD has mode-mixing problem when deals with the intermittent signal. Mode-mixing is defined as a single intrinsic mode function either consisting of signals of widely disparate scales, or a signal if a similar scale residing in different IMF components [16]. To overcome this drawback, Wu and Huang [17] proposed a new method named ensemble EMD (EEMD), which significantly reduces the mode-mixing of the decomposition for any data. According to [17], the EEMD decomposition algorithm of the original signal can be defined in the following steps:

1. Add a white noise series to the original data;
2. Decompose the amalgamation data (original and white noise) into IMFs;
3. Repeat step 1 and step 2 again and again, but with different white noise series each time; and
4. Obtain the (ensemble) means of corresponding IMFs (Fig. 12.1).

E number of ensemble, t is data length of vibration signal; α is a ratio between standard deviation of white noise, $w(t)$, and the original signal, $x(t)$; d is number of decomposition or number of IMFs, and d can be calculated by $d = \log 2(t) - 1$.

The performance of the EEMD method has been applied to natural phenomena such as climate and voice [17]. In this paper, EEMD was examined for two real cases slewing bearing data with natural fault progression. The result and discussion are presented in Sect. 12.3.

Ensemble empirical mode decomposition

for $i = 1,\ldots,E$ **do**

White noise

$n(t)^{(i)} = w(t)^{(i)}\ \alpha \qquad , i = 1,\ldots,E$

Add white noise to the original data **(step 1)**

$y(t)^{(i)} = x(t)^{(i)}\ n(t)^{(i)} \qquad , i = 1,\ldots,E$

for $j = 1,\ldots,d$ **do**

Decompose $y(t)^{(i)}$ using EMD into d number of IMFs **(step 2)**

Result $imf_{(i)}^{(d)}$, $i = 1,\ldots,E$

end for

Repeat **step 1** and **step 2** for $i = 2,\ldots,E$ **(step 3)**

$\mathbf{IMF} = imf_{(i)}^{(d)} + imf_{(i+1)}^{(d)} + \ldots + imf_{(E)}^{(d)}$

end for

Obtain the ensemble means **(step 4)**

$result = mean(\mathbf{IMF})$

Fig. 12.1 Ensemble empirical mode decomposition algorithm

12.3 Real-Case Vibration Data

12.3.1 Laboratory Slewing Bearing Test-rig

The basic requirement of a slewing bearing test-rig is to run a new bearing under controlled conditions, continuously to the point where replacement is necessary. Figure 12.2 shows the schematic of a slewing bearing test-rig, demonstrating how the bearing is attached. The test-rig can be operated at speeds of 1–12 rpm. The test-rig was designed to simulate the real working conditions of a steelmaking company. In this project, the test-rig was operated in one direction of rotation at a speed of 1 rpm. The slewing bearing used was an axial/radial bearing supplied by Schaeffler (INA YRT260) with an inner and outer diameter of 260 and 385 mm. The bearing fault frequencies for rotation at a speed of 1 rpm in both directions (axial and radial) are shown in Table 12.1.

Ten sensors were installed on the bearing: four accelerometer sensors, four temperature sensors, and two acoustic emissions (AE) sensors. Accelerometer and temperature sensors were placed on the inner radial surface at 90° to each other, and the AE sensors were attached on the top axial surface. The illustration of the sensors arrangement is shown in Fig. 12.3. The type of accelerometer used is IMI608 A11 ICP type sensor. The accelerometers were connected to high-speed DAQ (PS3424) from Pico scope. To avoid noise corruption from laboratory environment, the vibration data collection was conducted at midnight outside of working hours with 4880-Hz sampling rates. The vibration data were acquired

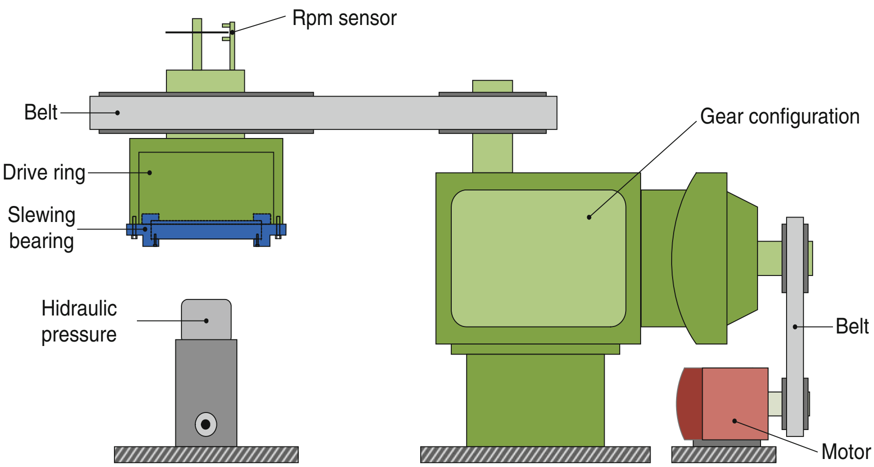


Fig. 12.2 Schematic of slewing bearing test-rig

Table 12.1 Slewing bearing fault frequencies

Defect mode	Fault frequencies (Hz)			
	Slewing bearing test-rig (run at 1 rpm)		Slewing bearing Bridge Reclaimer (run at 4.5 rpm)	
	Axial	Radial	Axial	Radial
Outer ring (BPFO)	1.32	0.55	13.41	11.38
Inner ring (BPFI)	1.37	0.55	13.58	11.56
Rolling element (BSF)	0.43	0.54	5.65	4.87

during the period of February to August 2007 (138 days). In order to accelerate the bearing defect, dust contamination was injected into the bearing middle of April 2007 (58 days from beginning). The constant load of 30 tonne was applied via hydraulic pressure as shown in Fig. 12.2.

In the slewing bearing test rig case study, three data groups were analyzed: (1) before dust contamination on February 24; (2) 1 month after dust injection on May 3; and (3) before bearing failure on August 30. FFT was calculated initially to find bearing fault frequencies which indicate the fault occurrences. Several signal or data analysis methods such as wavelet decomposition, short-time Fourier transform (STFT), and adaptive line enhancer (ALE) were also carried out. Unfortunately, the fault frequencies were also not identified using these methods (the results are not present in this paper due to limited space). According to Fig. 12.4d–f, the frequency of February 24 was localized in low-frequency area (approx. less than 100 Hz), while the frequencies after the injection of coal dust show moderate discrepancy and a couple of high frequencies; An identical phenomena also occurred on August 30. Low-pass filter (LPF) with cut-off frequency

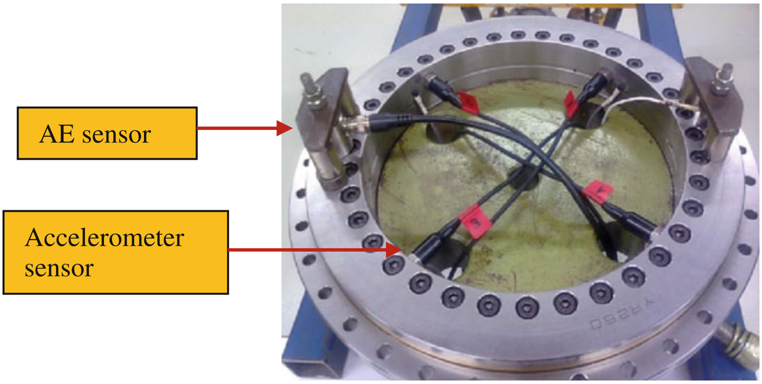


Fig. 12.3 Sensors arrangement

was applied. However, the FFT after LPF result was unsuccessful in identifying bearing fault frequencies (the result also not present in this paper). Therefore, as mentioned earlier, some established signal processing methods were used. A conclusion can be made from those methods, especially from wavelet decomposition and STFT. In case of wavelet decomposition, an excellent result solely depends on the option of mother wavelet, if the mother wavelet's shape and length are fit to the signal to be analyzed, the frequency of decomposition result close to bearing fault frequencies; and in case of STFT, the result varies based on the selection of moving window length. The EMD results for the three data groups are shown in Fig. 12.4a–c.

According to Figs. 12.4 and 12.6, the bearing fault frequencies were not identifiable on Feb 24. In contrast, the fault frequency of rolling element (axial) appeared on May 3 in 0.432 Hz and 0.442 Hz for EMD and EEMD, respectively. This frequency is associated with roller defect as shown in Fig. 12.5a. In addition, Fig. 12.5 shows the damage rollers in their axial plane and outer raceway after final failure upon dismantle on September 2, 2007. In the EEMD method, there are two parameters needed in order to achieve optimum results: (1) ratio of the standard deviation between the added noise and the inputted signal defined as α and (2) ensemble number for the EEMD denoted as E. Even though Wu and Huang [17] suggested the value of α of 0.1, 0.2, and 0.4, however, these values were not always suitable for all application. In this paper, the suggested α gave unsatisfied results. Therefore, the greater and lower α than the suggested range were tried. Finally, α of 0.01 is selected in this paper. For ensemble number, E, the recurrent value used in [17] and [18] is 100. Therefore, $E = 100$ was also selected for this study. We also tried the greater value, E, of 1,000 and lower value, E, of 100. Since the slewing bearing runs at a very low rotational speed of 1 rpm, the selection of E may result in two conditions. If the E is high (e.g., 1,000), then the calculation process of MATLAB is time consuming. On the contrary, if the ensemble number is low (e.g., 10), the result has a large diversity due to the effect

(a) EMD results February 24 (Hz)

Result 2	641.807	Result 10	4.849
Result 3	694.521	Result 11	2.474
Result 4	390.350	Result 12	1.439
Result 5	190.387	Result 13	0.687
Result 6	89.751	Result 14	0.327
Result 7	52.591	Result 15	1.990
Result 8	18.831	Result 16	0.106
Result 9	9.019		

(b) EMD results May 3 (Hz)

Result 2	702.880	Result 10	4.026
Result 3	684.340	Result 11	2.042
Result 4	346.227	Result 12	1.227
Result 5	1371.492	Result 13	0.638
Result 6	74.370	Result 14	**0.432**
Result 7	69.199	Result 15	0.194
Result 8	17.568	Result 16	0.069
Result 9	8.570		

(c) EMD results August 30 (Hz)

Result 2	651.717	Result 9	6.464
Result 3	679.862	Result 10	3.053
Result 4	245.923	Result 11	**1.374**
Result 5	121.319	Result 12	0.682
Result 6	63.554	Result 13	0.331
Result 7	28.891	Result 14	0.099
Result 8	13.682		

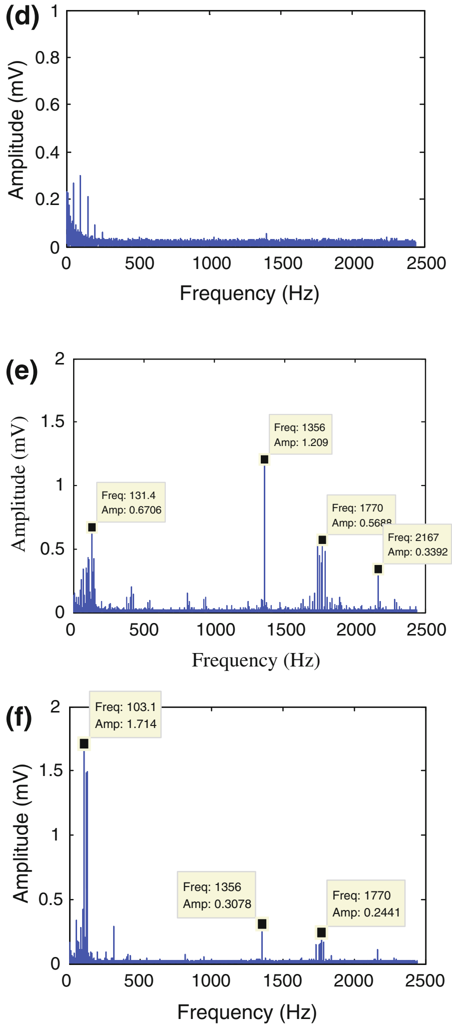


Fig. 12.4 EMD results; **a** Feb 24, **b** May 3, and **c** Aug 30; and FFT of slewing bearing test-rig; **d** Feb 24, **e** May 3, and **f** Aug 30

of random white noise. In 1960s of data acquisition, running at 1 rpm with a sampling frequency of 4,880 Hz produced 292,800 data points. Referring to fault frequencies presented in Table 12.1, the lowest frequency of 0.43 Hz (rolling element axial) was considered in calculation of input data length for EMD and EEMD methods. In simple calculation, 0.43 Hz will have one full cycle period in at least 2 s; if we choose 10 s (equal to 48,800 data points), it is enough to represent all fault frequencies. Therefore, the frequency calculation of decomposition result will be more accurate.

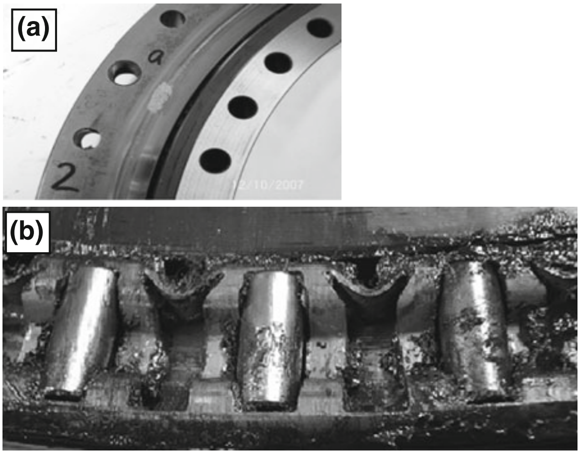


Fig. 12.5 A view of slewing bearing damage; **a** outer race damage and **b** damaged rollers in axial plane

Interesting results were obtained from data on August 30. In EMD, the fault frequency of the rolling element is disappeared and another fault frequency in the outer raceway of 1.37 Hz is emerged, while in EEMD, the fault frequency of rolling element in the axial plane of 0.443 Hz is still apparent. The outer raceway damage is shown in Fig. 12.5b. Earlier study in similar slewing bearing was presented by Moodie [2]. He proposed a new method called symmetric wave decomposition (SWD). Three fault frequencies of 0.45 Hz (BSF axial), 1.32 Hz (BPFO axial), and 1.37 Hz (BPFI axial) were identified (see Ref. [2], p. 94].

12.3.2 Bridge Reclaimer Data

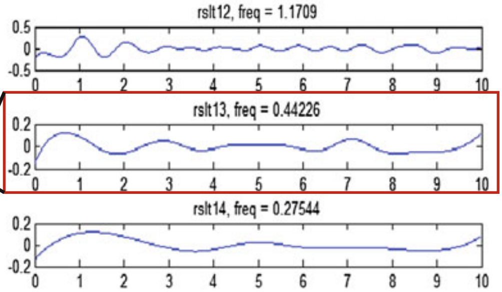
Current condition monitoring methods for the Bridge Reclaimer provide a short sample (4,098 max) of acceleration data at 240 samples/s. The slewing bearing usually rotates at a constant speed of approx. 4.5 rpm. The slewing bearings were mounted on a Coal Bridge Reclaimer in the vertical plane. The type of slewing bearing is RotheErde 4.3 diameter. The Coal Bridge Reclaimer located in the open air and subjected to all the elements including all the various forms of dust particles that are part of a steelmaking complex. The bearing fault frequencies are presented in Table 12.1. Historical time waveform data were made available from the Bridge Reclaimer database system employed by the industrial partner. The accelerometers employed are IMI512, 500 mV/g ICP type piezos, magnetically mounted. The sensors are not permanently installed. Data were collected from 2003 to 2006 at each sample point approximately once per month and was captured via an industrialized portable DAQ unit. The DAQ equipment employed during this project was based on a laptop PC utilizing the NI 5102 PCM-CIA digitizer. For signal pre-amplification, a Behringer UB802 unit was employed. For

(a) EEMD results February 24 (Hz)

Result 2	640.367	Result 6	95.277	Result 10	16.438	Result 14	3.647
Result 3	655.461	Result 7	49.166	Result 11	2.251	Result 15	0.106
Result 4	375.814	Result 8	31.141	Result 12	1.290	Result 16	0.046
Result 5	192.573	Result 9	8.973	Result 13	0.724		

(b) EEMD results May 3 (Hz)

Result 2	706.538	Result 10	23.347
Result 3	675.126	Result 11	2.122
Result 4	334.005	Result 12	1.170
Result 5	139.962	Result 13	**0.442**
Result 6	73.153	Result 14	0.275
Result 7	38.522	Result 15	0.101
Result 8	24.166	Result 16	0.050
Result 9	24.035		



(c) EEMD results August 30 (Hz)

Result 2	656.950	Result 10	3.197
Result 3	663.255	Result 11	1.674
Result 4	219.531	Result 12	0.851
Result 5	103.659	Result 13	**0.443**
Result 6	57.064	Result 14	0.219
Result 7	28.081	Result 15	0.101
Result 8	21.660	Result 16	0.046
Result 9	27.284		

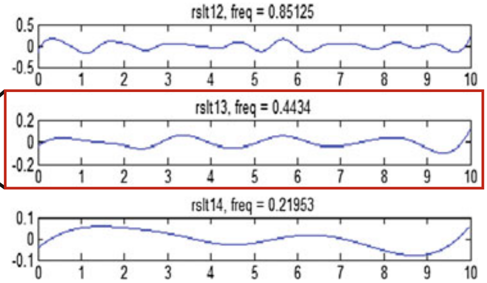


Fig. 12.6 Ensemble EMD result for slewing bearing test-rig; **a** February 24, **b** May 3, and **c** August 30

the ICP accelerometers, a portable Wilcoxon charge amplifier (internal 9 V batteries) was used to convert the charge signal to a voltage signal.

In second real case, FFT of the Bridge Reclaimer data from December 22, 2005, was examined and is shown in Fig. 12.7b. These data were collected after running the bearing for two years and supposedly in fault mode. It can be seen from FFT that the frequency dominant was not correspond to the bearing fault frequencies. Therefore, the similar EMD and EEMD methods employed in slewing bearing test-rig data were carried out. The sampling rate of the data acquisition is 240 Hz. The duration of data input is considered from the lowest fault frequency of 4.87 Hz (rolling element radial). Concisely, the duration of signal input for EMD

(a) EMD results	
Result 2	35.354 Hz
Result 3	27.476 Hz
Result 4	14.494 Hz
Result 5	7.486 Hz
Result 6	3.153 Hz
Result 7	1.687 Hz
Result 8	0.934 Hz
Result 9	0.199 Hz

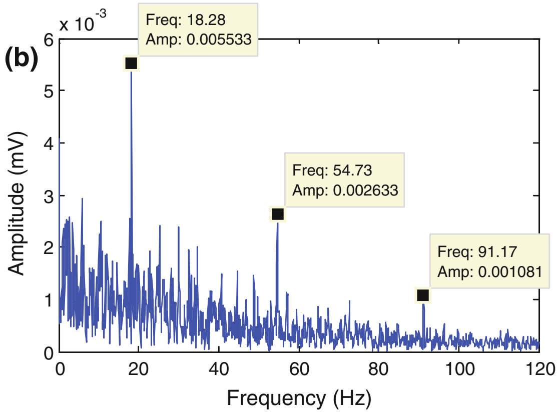


Fig. 12.7 **a** EMD result; and **b** FFT plot of Bridge Reclaimer data

EEMD results	
Result 2	35.736 Hz
Result 3	25.109 Hz
Result 4	**13.417** Hz
Result 5	6.844 Hz
Result 6	**5.674** Hz
Result 7	1.519 Hz
Result 8	0.878 Hz
Result 9	0.361 Hz
Result 10	0.198 Hz

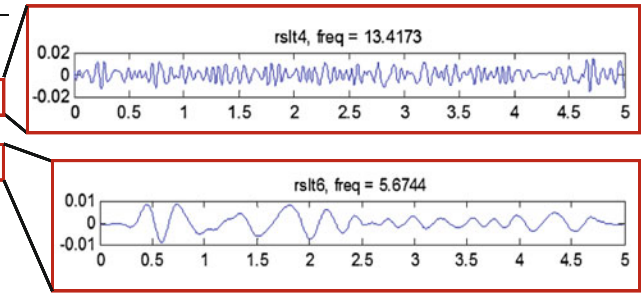


Fig. 12.8 Ensemble EMD results for Bridge Reclaimer data

and EEMD is selected in 5 s (equal to 1,200 number of data). The selected EEMD parameter is α of 0.05 and E of 100. The result of EMD was not as good as in the case of laboratory slewing bearing test-rig. According to Fig. 12.7a, none of the fault frequencies of the Bridge Reclaimer can be identified. Surprisingly, the EEMD results shown in Fig. 12.8 showed two fault frequencies 13.417 and 5.674 Hz which are associated with the fault frequency of the bearing outer raceway (axial) and rolling element (axial) as shown in Table 12.1.

12.4 Summary and Conclusion

The EMD and EEMD methods have been employed in two naturally slewing bearing damage data. In the case of laboratory slewing bearing test-rig, the performance of EMD is slightly better than EEMD, where two bearing fault frequencies were identified using EMD and only one fault frequency was identified using EEMD. In Bridge Reclaimer data, the EEMD presents more superior than EMD, where two fault frequencies can be identified and no fault frequency is identified using EMD. These results provide the contribution of EMD and EEMD application in real bearing case, where the faults are developed naturally whereas the most discussed literature were used EMD and EEMD for artificial bearing fault data.

Earlier study which reported the application of EMD and EEMD for slewing bearing naturally damage data is presented in [19]. The rotational speed of slewing bearing was 15 rpm. The slewing bearing data were acquired from Korean industrial company. Two fault frequencies were identified using EEMD method.

Acknowledgments The first author gratefully acknowledges the University of Wollongong financial support through University Postgraduate Award (UPA) and International Postgraduate Tuition Award (IPTA). Also this research was partially supported by the Basic Science Research Program through NRF (No. 2011-0013652).

References

1. Slewing Bearings [Product Catalogue] (2007) RotheErde GmbH, Dortmund
2. Moodie CAS (2009) An investigation into the condition monitoring of large slow speed slew bearings. PhD thesis in faculty of engineering, University of Wollongong, Wollongong
3. Bai X, Xiao H, Zhang L (2011) The condition monitoring of large slewing bearing based on oil analysis method. Key Eng Mater 474–476:716–719
4. Liu R (2007) Condition monitoring of low-speed and heavily loaded rolling element bearing. Ind Lubr Tribol 59(6):297–300
5. Marciniec A, Torstenfelt B (1994) Load distribution in flexibly supported three-row roller slew bearings. Tribol Trans 37(4):757–762
6. Kania L (2006) Modelling of rollers in calculation of slewing bearing with the use of finite elements. Mech Mach Theory 41(11):1359–1376
7. Göncz P, Potočnik R, Glodež S (2011) Load capacity of a three-row roller slewing bearing raceway. Eng Procedia 10:1196–1201
8. Gang Z, Xue Z, Kaifeng Z, Juan R, Dede J, Qingzhen Y, Mingyan L, Yan Z (2011) Optimization design of large-scale cross-roller slewing bearing used in special propeller. Appl Mech Mater 48–49:787–792
9. Glodež S, Potočnik R, Flašker J (2012) Computational model for calculation of static capacity and lifetime of large slewing bearing's raceway. Mech Mach Theory 47:16–30
10. Aguirrebeitia J, Abasolo M, Avilés R, Fernández de Bustos I (2012) Theoretical calculation of general static load-carrying capacity for the design and selection of three row roller slewing bearings. Mech Mach Theory 48:52–61
11. Rodger LM (1979) The application of vibration signature analysis and acoustic emission source location to on-line condition monitoring of anti-friction bearings. Tribol Int 12:51–59
12. Liu Z, Chen J (2011) Monitoring and diagnosis technique on slewing bearing. Adv Mater Res 156–157:1478–1483
13. Žvokelj M, Zupan S, Prebil I (2010) Multivariate and multiscale monitoring of large-size low-speed bearings using ensemble empirical mode decomposition method combined with principal component analysis. Mech Sys Sign Proc 24:1049–1067

14. Žvokelj M, Zupan S, Prebil I (2011) Non-linear multivariate and multiscale monitoring and signal denoising strategy using kernel principal component analysis combined with ensemble empirical mode decomposition method. Mech Sys Sign Proc 25:2631–2653
15. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond A 454:903–995
16. Guanlei X, Xiaotong W, Xiaogan X (2009) Time-varying frequency-shifting signal-assisted empirical mode decomposition method for AM-FM signals. Mech Sys Sign Proc 23:2458–2469
17. Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 1(1):1–41
18. Wu Z, Huang NE (2005) Ensemble empirical mode decomposition: a noise-assisted data analysis method, centre for ocean-land-atmosphere studies. Technical Report 193:51, http://www.iges.org/pubs/tech.html
19. Caesarendra W, Park JH, Kosasih PB, Choi BK (2013) Condition monitoring of low speed slewing bearings based on ensemble empirical mode decomposition method. Trans Korean Soc Noise Vib Eng 23(2):131–143

Chapter 13
Understanding Wear Mechanisms and Their Implication to Service Life of Pneumatic Conveying Pipelines

A. A. Cenna, S. Biswas, K. C. Williams and M. G. Jones

Abstract Pneumatic conveying is a process of transporting particulate material through pipelines using compressed gas. As material is conveyed through pipeline and bends, the pipeline especially after bends suffers severe wear due to particles' interactions with the surfaces. Removal of material from solid surfaces by action of impinging particles is known as erosion. It is well known that particle velocity and impact angle play a major role in determining the material removal rate from the surface. In a recent study, it was demonstrated that materials' response to deformation during impacts dictates how the material is removed from the surface. This paper presents the surface characteristics of ductile materials due to single-particle impacts as well as standard erosion using micro-sand blaster. Surface and subsurface damage characteristics with respect to the impact parameters as well as particles' angularity have been presented. Aluminum and mild steel surfaces impacted by spherical zirconia and angular alumina particles have been analyzed using scanning electron microscope (SEM). Finally, the material removal mechanisms have been discussed with respect to the service life of pneumatic conveying pipelines.

13.1 Introduction

Wear is surface damage that generally involves progressive material loss due to relative motion between contacting substance or substances. There are a number of theories on the mechanisms of material removal in wear, but the wear of pneumatic conveying pipelines caused by erosion and abrasion is in common agreement among researchers. Based on the interactions between the surface and the

A. A. Cenna (✉) · S. Biswas · K. C. Williams · M. G. Jones
School of Engineering and Built Environment, University of Newcastle,
Callaghan, NSW 2308, Australia
e-mail: ahmed.cenna@newcastle.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_13,

erodent, material removal mechanism can be defined as cutting and deformation [1]. Although material can be removed from the surface by a single impact in cutting, in deformation, multiple impacts as well as a secondary process such as fatigue may be involved.

Pneumatic conveying is a complex problem of two-phase flow, and at present, there is a limited understanding of the particles' interactions with the pipeline surfaces. Understanding of the surface phenomena at impacts would play a crucial role in developing wear models to predict the wear in pneumatic conveying pipelines and hence the service life.

Due to the characteristics of the conveying process, wear is an inherent problem in pneumatic conveying pipelines. As the air conveying the material expands, air velocity as well as particle velocity increases toward the end of the pipeline. As a result, wear characteristics and wear rates change along the pipeline. A better understanding of the particle flow characteristics as well as surface characteristics at impacts is needed to develop predictive models for pneumatic conveying pipelines.

Erosion is a complex phenomenon of material removal from surface which is usually explained through cutting and deformation by the impacting material or particles. When particle impinges on a surface, particle may only deform the surface without removing any material or material removal can take place with accompanied deformation. In the case of no material removal, particles' energy is transferred into the surface and gradually work-hardened through accumulated strain energy. Effects of surface work hardening have been studied in Refs. [2, 3].

Surface deformation characteristics of ductile materials have been presented as two distinct zones based on the plastic flow of materials in Ref. [4]. Zone 1 represents an unconstrained layer in which higher dynamic shear intensified due to high-strain-rate deformation that causes localized heating and softening of material. Zone 2 represents the material undergoing bulk deformation at relatively low strain rates, similar to the deformation zone produced in static indentation. Figure 13.1 presents the schematics of the surface characteristic at impacts for (a) normal impacts and (b) oblique impacts on ductile materials.

Surface deformation characteristics at high-strain-rate impacts have been studied by the authors as presented in Ref. [5]. For a better understanding of the phenomena, surface analysis of single-particle impacts has been analyzed in conjunction with surface analysis of ductile materials from micro-sand blaster. The experimental setup is explained in the following section.

13.2 Experimental Investigations

The experimental results presented here have been obtained from two different test facilities. Single-particle impact tests conducted on a high-pressure gas gun and standard erosion tests were conducted using micro-sand blaster. The single-particle impact tests were conducted using spherical zirconia and angular alumina particles

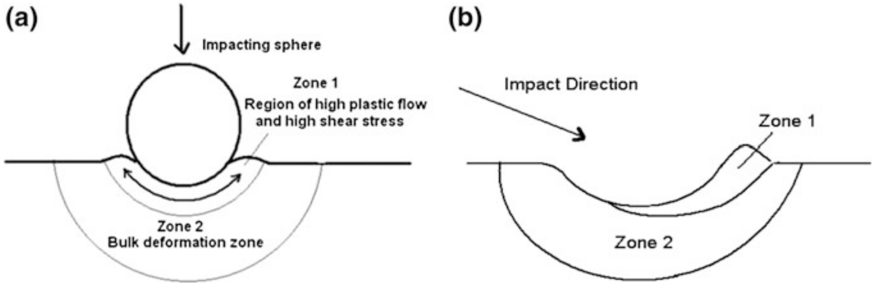


Fig. 13.1 Characteristics of impact crater on ductile materials produced by **a** normal impact and **b** oblique impact by a spherical particle

(95–100 μm) on mild steel and aluminum surfaces. Figure 13.2 presents the schematic of the gas gun used for the tests. Detailed description of the gas gun and the experimental procedure can be found elsewhere [5]. The erosion tests were conducted according to the ASTM G76-07 standard using a micro-sand blaster.

The objective of these tests was to study the surface characteristics of mild steel by single-particle impacts and relate those surface effects with standard erosion tests surfaces. It is expected that this will help to develop the understanding of wear mechanisms in pneumatic conveying processes and develop predictive model for service life.

13.3 Surface Analysis

13.3.1 Surface Analysis in Single-Particle Impacts

The characteristics of surface craters from single-particle impacts are presented in this section. As both the particles are extremely hard compared to the impacted surface, the particle hardness effect can be ignored in this analysis. Figure 13.3 presents the surface craters in aluminum by spherical zirconia particles at different particle velocities and impact angles.

It is obvious that at normal impacts, surface craters are characterized by deformation only, with material piled up around the impact crater. Figure 13.3b shows the characteristics of a surface crater consistent with particle fracture on impact. The inward deviation of the crater indicates less deformation associated with particle's fracture on impact.

Images of impact craters at 30° and 60° (with the surface), respectively, are presented in Fig. 13.3c and d. The lip is formed at the exit of the particle in Fig. 13.3c that is consistent with surface heating on impact. The thin extended lip attached to the surface apparently re-solidified after being drawn outward by surface adhesion with the departing particle.

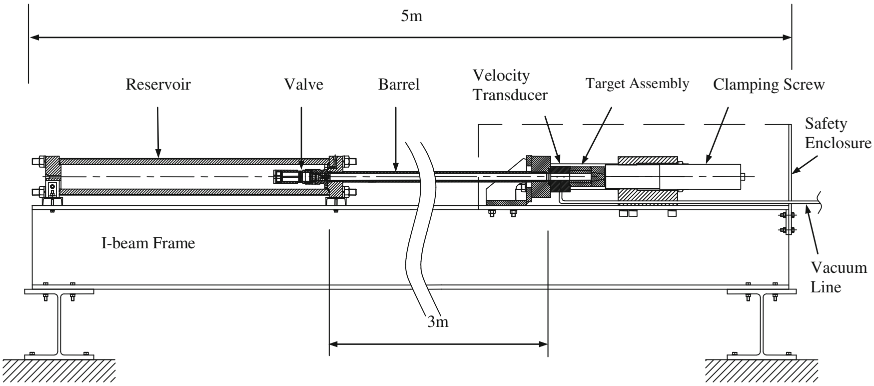


Fig. 13.2 Schematic diagram of the gas gun

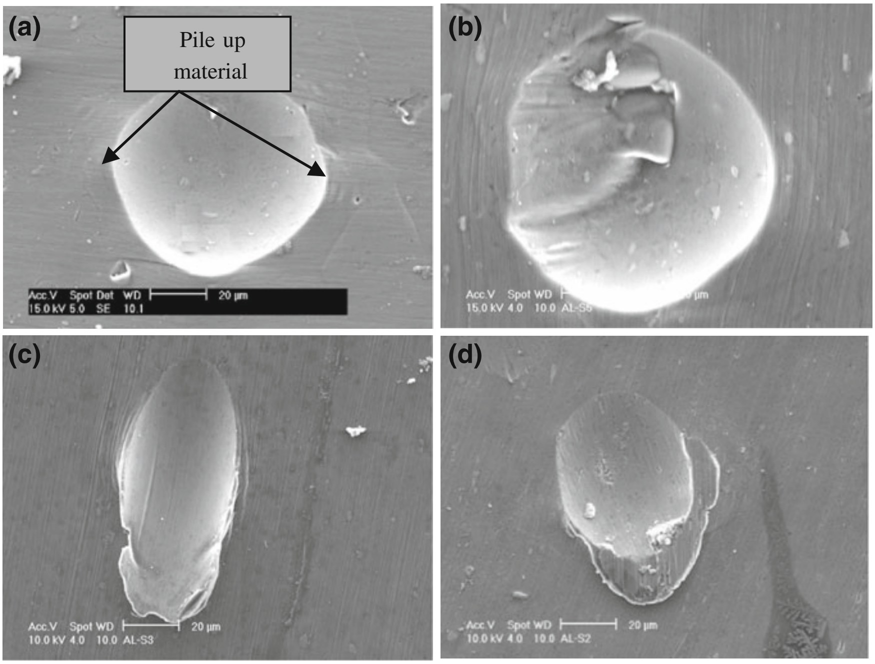


Fig. 13.3 Characteristics of surface craters for normal impacts by particles traveling at **a** 74 m/s and **b** 84 m/s and **c** 30^0 and **d** 60^0 for oblique impacts

Figure 13.3d shows the lip attached to the surface with undamaged surface texture presenting a different mechanism of material removal from the surface. The white edge of the lip is due to charging in the electron beam—a strong indication that this edge of lip has been detached from the original surface. Shear band generated in front of the impact particle reached the surface to remove material from the surface as a microchip [5]. The material along the shear band is subjected to high temperature that can soften or possibly melt as shown in Fig. 13.3d.

13.3.2 Surface Analysis for Angular Particles

Surface craters for angular particles are presented in Fig. 13.4. Although it is expected that cutting is the dominant mechanism for material removal by angular particles, significant amount of deformation also occurs as demonstrated by these experiments.

Figure 13.4a and b present the surface craters at normal impacts. Even though the particles are clearly cut into the surface, the figures clearly demonstrated different surface characteristics based on the particles' orientation at impact. High deformation can be observed in Fig. 13.4a, while cutting and formation of chip through formation of shear band is the primary mechanism in Fig. 13.4b.

Similar complexity has been observed in oblique impacts as presented in Fig. 13.4c and d for angular particles. Figure 13.4c presents an impact crater from which material has been removed through cutting as a chip without any apparent deformation of the surface. This is the expected dominant mechanism of material removal for low impact angles. Figure 13.4d presents an impact crater which shows substantial deformation on the surface without any apparent material removal.

Analysis presented for single-particle impact tests demonstrates the complexity of material removal process from surface in erosion by spherical and angular particles. It is expected that the understanding of particle energy dissipated into the surface can be utilized in developing predictive models for erosion in pneumatic conveying pipelines.

13.3.3 Surface Analysis in Particle Erosion

Surface characteristics of erosion tests samples have been analyzed in this section. The aim of this section is to extend the understanding of the surface characteristics of ductile materials in single-particle impacts into wear mechanisms in erosion.

Figure 13.5 presents the SEM micrographs of mild steel surface eroded by spherical zirconia particles. Figure 13.5a shows the surface eroded at normal angle, while Fig. 13.5b presents the surface eroded at oblique angle. Both surfaces clearly demonstrate that material removal from the surface occurs as thin platelets. Ripples pattern can be seen at oblique impacts (Fig. 13.5b). Surface flaking at the edge of the ripple seems to the primary mechanism of material removal in this case. Clearly, this is a very high energy extensive process for material removal. As

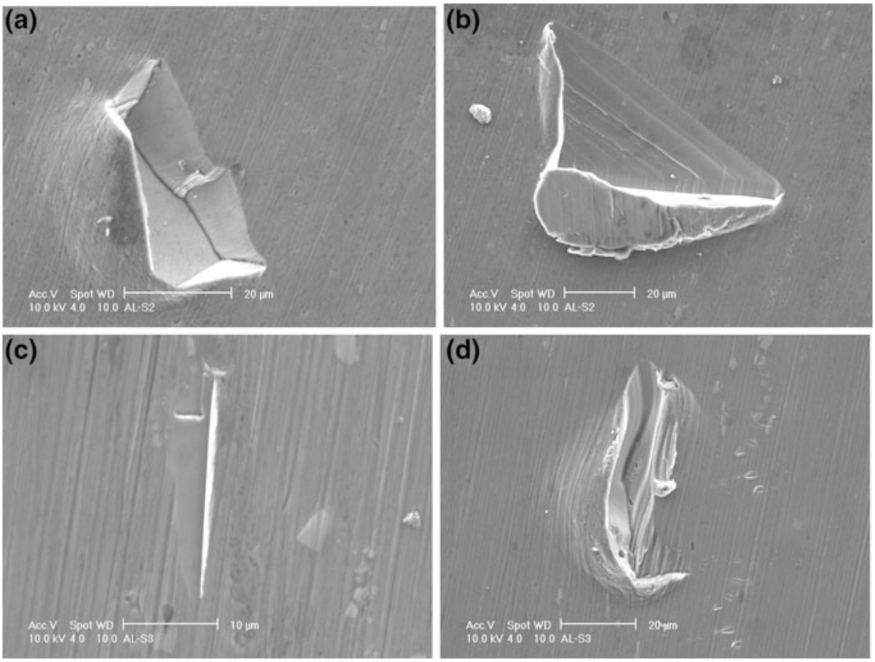


Fig. 13.4 Characteristics of surface craters by angular alumina particles at normal impacts: **a** deformation, **b** cutting and at oblique impacts, **c** cutting, and **d** deformation only

observed in the single particle, energy dissipated into the surface is clearly very high at normal impact which is associated with surface hardening and scale formation.

Micrographs of wear surface from eroded by angular particles are presented in Fig. 13.6. Figure 13.6a presents the micrograph of mild steel surface eroded by alumina particles at normal impact, and Fig. 13.6b presents the surface characteristics at oblique impacts.

Although cutting is clearly the dominant mechanisms of material removal for angular particles in oblique angle (Fig. 13.6b), surface characteristics at normal impacts (Fig. 13.6a) present different phenomena. Surface cracks observed in Fig. 13.6a are consistent with brittle fracture which is the result of the work hardening of surface layer through loading cycles on the wear surface. Material removal through surface hardening and brittle fracture has been studied extensively in Ref. [3].

Surface heating and surface cracking in erosion have been further studied through subsurface analysis of the wear samples and are presented in the following section.

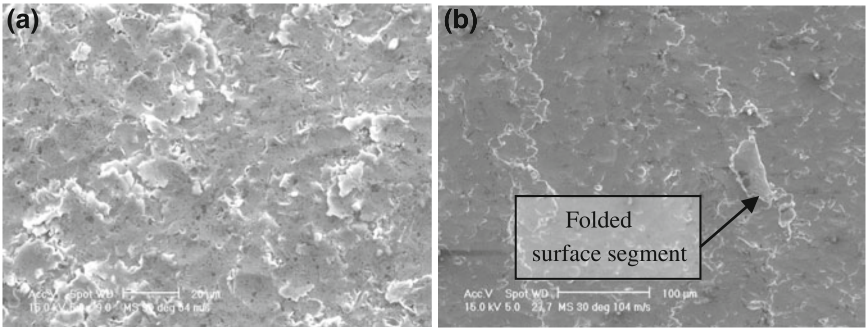


Fig. 13.5 Surface characteristics of mild steel eroded by spherical zirconia particles at **a** normal impacts and **b** oblique impacts

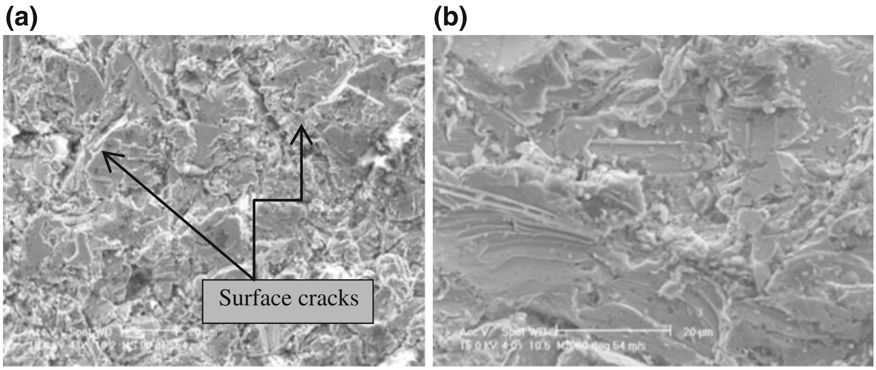


Fig. 13.6 Surface characteristics of mild steel eroded by alumina particles at **a** normal impacts and **b** oblique impacts

13.3.4 Subsurface Damage

From the surface analysis of the wear samples and single-particle impact tests, it became obvious that other than direct effects of the particles' interactions with the surface, energy dissipated into the surface also plays an important role in material removal from surface. The effects of particles' energy dissipated into the surface have been analyzed through analysis of the subsurface damage of the samples.

Samples were sectioned through the impact crater and polished using diamond cloth. The surfaces were then etched to observe the subsurface grain boundaries as well as the depth of surface damage. Figure 13.7 presents the micrographs of subsurface and reveals the subsurface damage.

Figure 13.7a shows the subsurface of a mild steel sample impacted at oblique angle by angular alumina particles. No microstructural damage has been observed other than a subsurface crack beneath the surface that extends into the surface

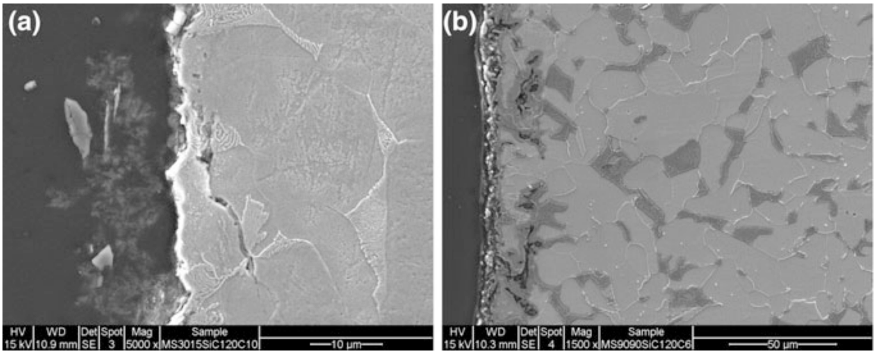


Fig. 13.7 Subsurface damage by alumina particles: **a** surface eroded at oblique impacts and **b** surface eroded at normal impacts

extended almost parallel to the surface. It appears that the crack has been extended along the grain boundary. The subsurface micrograph of wear sample eroded at normal impact is presented in Fig. 13.7b. Extensive subsurface damage is observed 10–15 μm below the surface which is consistent with heat generation due to high-strain-rate deformation. The analysis demonstrates that the heat generation and accumulation are not at the surface rather somewhat below the surface possibly dictated by the particle size, impact velocity, or a combination of different factors. It is also apparent that the subsurface damage due to heat may form a surface layer that can be eventually removed from the surface with ease, increasing the wear rate substantially. This analysis recognizes the existence of a third dominant mechanism of material removal at normal impact.

13.4 Conclusions

A different concept for understanding the wear mechanisms in ductile materials has been presented. Even though cutting and deformation mechanisms dictate the material removal from surface of ductile materials, energy absorbed into the surface also plays a critical role in determining the extent of subsurface damage which eventually adds to the overall material removal from surfaces.

Particles' impact energy absorbed into the surface and subsurface damage and its effects on wear mechanisms have been discussed in the case of single-particle impacts and erosion. It is evident that subsurface damage accumulation has a substantial effect on the wear characteristics of ductile materials and this has a profound effect on service life of pneumatic conveying pipelines. To develop predictive models for service life of pneumatic conveying pipelines, both the surface mechanisms such as cutting and deformation and subsurface damage characteristics need to be accommodated. This can be achieved through a clear understanding of the particles' interactions with the surface and the subsequent subsurface behavior of the materials.

Acknowledgments This work forms part of the research program for the CRC for Integrated Engineering Asset Management. The authors also acknowledge the EM-X-ray unit of the University of Newcastle for their continual assistance in SEM/EDS analyses.

References

1. Hutchings IM (1992) Tribology: friction and wear of engineering materials, Chap. 6. Wear by hard particles. Edward Arnold, London
2. Cenna AA, Williams KC, Jones MG, Page NW (2008) Wear mechanisms in dense phase pneumatic conveying of alumina. Wear mechanisms in pneumatic conveying pipelines. Wear 264:905–913
3. Cenna AA, Williams KC, Jones MG, Page NW (2009) Generation of transfer film and its effects on wear mechanisms in alumina conveying pipeline of mild steel. Wear 267:362–367
4. Levy AV (1981) The solid particle erosion behavior of steel as a function of microstructure. Wear 68:269–287
5. Cenna AA, Page NW, Kisi E, Jones MG (2011) Single particle impact tests using gas gun and analysis of high strain-rate impact events in ductile materials. Wear 271:1497–1503
6. Cenna AA, Williams KC, Jones MG (2010) Wear mechanisms in pneumatic conveying of sand and analysis of predictive model for pipeline thickness loss. WCEAM, Brisbane Queensland, Australia from 25–27 Oct 2010

Chapter 14
Effects of Surface Modifications on Wear Mechanisms in Fly Ash Conveying Pipelines

A. A. Cenna, K. C. Williams, M. G. Jones and W. Robinson

Abstract Pneumatic conveying is a frequently used method of material transport particularly for in-plant transportation over relatively short distances. This is primarily to exploit the degree of flexibility it offers in terms of pipeline routing as well as dust minimization. Approximately 80 % of industrial systems are traditionally dilute phase system which uses relatively large amount of air to stay away from trouble, such as blocking the pipeline. Wear in pneumatic conveying is a very complex problem, and at present, there is limited understanding of the wear mechanisms responsible for the severe wear in certain areas of the pipeline. One of the recent studies [1] showed that the surface modifications by conveying materials can alter surface characteristics which change the wear mechanisms of the pipeline. Better understanding of the surface modifications and their effects on wear mechanisms can play a significant role in developing predictive models for the service life of pneumatic conveying pipelines. In this paper, wear surfaces from fly ash conveying pipeline have been studied for a better understanding of the surface modification and its effects of wear mechanisms. Wear samples were analyzed using SEM Energy Dispersive X-ray Spectroscopy (SEM-EDS). Analysis of surface elements discovered high level of silicon (Si) and aluminum (Al) present in the modified surface areas which apparently responsible for brittle failure of the surface layer. Although the actual form of the chemical compounds has not been analyzed, it is evident that the surface modification by the constituents of the conveying material is one of the major contributors to the severity of wear in fly ash conveying pipelines.

A. A. Cenna (✉) · K. C. Williams · M. G. Jones · W. Robinson
School of Engineering and Built Environment, University of Newcastle, Callaghan, NSW 2308, Australia
e-mail: Ahmed.Cenna@newcastle.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_14,

14.1 Introduction

Pneumatic conveying involves the transportation of a wide variety of dry powdered and granular solids through pipeline and bends using high-pressure gas. It is a frequently used method of material transport particularly for in-plant transportation over relatively short distances. This is primarily to exploit the degree of flexibility it offers in terms of pipeline routing as well as dust minimization. Approximately 80 % of industrial systems are traditionally dilute phase system which uses relatively large amount of air to achieve high particle velocities to stay away from trouble, such as blocking the pipeline. However, for many applications, higher velocities lead to excessive levels of particle attrition as well as wear of pipelines, bends, and fittings. To combat these problems, there are systems designed to operate at relatively low-velocity regimes. Yet one problem remains as a major issue with these conveying systems which is wear.

Silicon dioxide (SiO_2) or silica and aluminum oxide (Al_2O_3) or alumina are the two major constituents of fly ash. The percentage of silica can be as high as 65 %, and alumina can vary between 15 and 30 %. Both alumina and silica are very hard materials, with silica having a hardness value of about 6 on the mohs scale of hardness and alumina being close to 8 [2]. Due to the high concentration of these two hard constituents, fly ash is very abrasive and causes serious damage to the pipeline surfaces.

Wear is a complex phenomenon of material removal due to particle's interaction with the surface. It is well known that in ductile materials, the material removal occurs primarily through deformation and cutting by free-moving particles. When a particle impinges on a ductile surface, the effects can be represented by one of the three situations presented in Fig. 14.1a, and depending on the particle velocity and impact angle [3], (a) particle may only deform the surface (plowing), (b) particle may dig into the surface and stop while cutting (cutting and deformation), and (c) particle may leave the surface while cutting removing a small chip in front of it.

Material removal in brittle materials occurs in different mechanism than that of ductile materials. Fracture and crack propagation are considered the dominant mechanism of material removal in brittle materials. As a particle slides over the surface forming a plastic groove, lateral cracks grow upwards to the free surface from the base of the subsurface deformed region as shown in Fig. 14.1b. Material removal due to cracking is the dominant mechanism in brittle materials.

In the case of wear of pneumatic conveying pipelines, it is expected that cutting and deformation by the conveying materials dominate the material removal rates especially for mild steel pipelines. On the contrary, it was demonstrated that for pipeline conveying alumina, dominant wear mechanism was crack formation and removal of material through spalling of material in layers [1]. In that study, through extensive analysis of wear surfaces from alumina conveying pipelines, the authors have shown that a transfer film generated by the conveying material alters the surface characteristics from ductile to brittle. As the transfer film is harder than

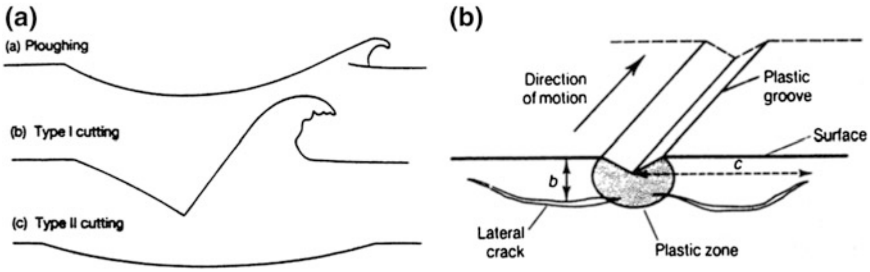


Fig. 14.1 Material removal characteristics: **a** in ductile materials and **b** in brittle materials [4]

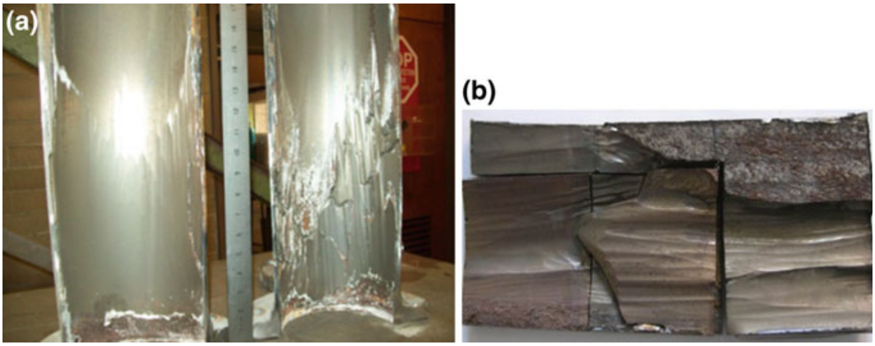


Fig. 14.2 Wear sections from fly ash conveying pipeline for SEM analysis. **a** Pipe section spit in halves and **b** carbon-coated samples for SEM/EDS analysis

the substrate, it get delaminated from the surface due to fluctuation of pressure and removed through cracking, a brittle mechanism of material removal from surface.

The aim of this paper is to analyze the wear mechanisms in fly ash conveying pipelines and the effects of surface modifications on wear mechanisms. It has been observed that similar to alumina conveying pipelines, surface characteristics in fly ash conveying pipelines also altered which dominate the service life of these pipelines.

14.2 Wear in Fly Ash Conveying Pipelines

Analysis of wear surfaces to determine the dominant wear mechanisms in pneumatic conveying pipelines is presented in this section. Wear samples collected from industrial pipeline have been prepared for microscopic analysis. Figure 14.2 shows the wear sections from fly ash conveying pipeline and samples cut for surface analysis using scanning electron microscope (SEM). Figure 14.2a shows the severity of wear after the bends.

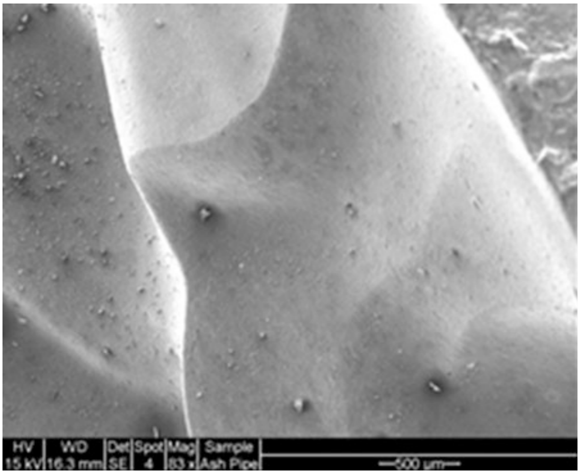


Fig. 14.3 Surface features representing the erosive wear mechanism

Wear mechanisms in fly ash conveying pipelines have been studied extensively using the scanning electron microscope and presented at the WCEAM 2011 [5]. Figure 14.3 showed the complex surface profile created on exit pipeline after a bend. This is due to the high turbulent flow generated in the bend which affects the pipeline after the bends. Smooth appearance of the surface represents surface wear by fine particles in erosion. Particle flow characteristics in bends were studied through flow visualization and were presented earlier in [6]. These studies demonstrated that the turbulent flow continues through the reacceleration zone after the bend.

Figure 14.4 presents the surface characteristics consistent with abrasive wear mechanisms. Abrasive wear occurs when particles remain in contact with the surface while moving. Larger angular particles are free to roll on the surface while sliding. The abrasive wear situation is created after the bends due to the well-known rope-like structure formed by the particles. As this structure flows over the surface, surface does not recover from individual particle impacts, rather the particle slides along the surface under pressure from upcoming particles.

Figure 14.4a shows a surface area which is consistent with fatigue damage due to particle rolling on the surface. Rolling and sliding of angular particles generates fatigue damage that creates irregular pitting on the surface (Fig. 14.4b). Observation of the surface damage clearly reveals brittle failure of the surface layer which is believed to be due the modification of the ductile pipeline surface. Analysis of the surface modification is presented in the following section.

Figure 14.5 presents the micrographs of modified surface area and ensuing wear mechanisms in fly ash conveying pipelines. The area (Fig. 14.5a) is characterized by severe cracking of a surface layer which could be removed easily through spalling. Figure 14.5b presented an area from which the cracked surface layer has been removed leaving a small segment still attached to the surface.

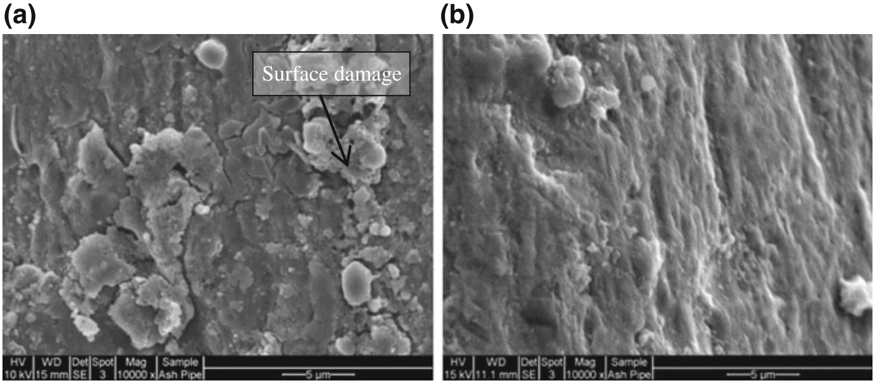


Fig. 14.4 Surface characteristics consistent with abrasive wear

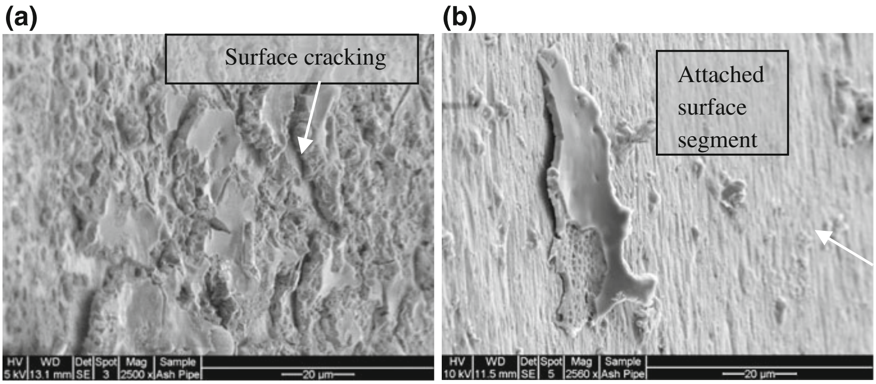


Fig. 14.5 Surface modification and material removal through brittle mechanism

Material removal in this process is extremely high compared to those due to cutting and deformation by direct impacts of particles. Clearly, the material characteristics in this surface layer are brittle, different from the pipeline material, mild steel. Crack pattern is consistent with temperature cycling on the surface. This is possible in fly ash conveying pipelines as the surface undergoes temperature cycles due to the conveying cycles of hot ash.

14.3 Analysis of Surface Modification

Surface modifications in fly ash conveying pipeline have been analyzed using SEM-EDS. EDS provides elemental profile from the surface using backscatter electron (BSE) detector. BSE provides high-resolution compositional map of

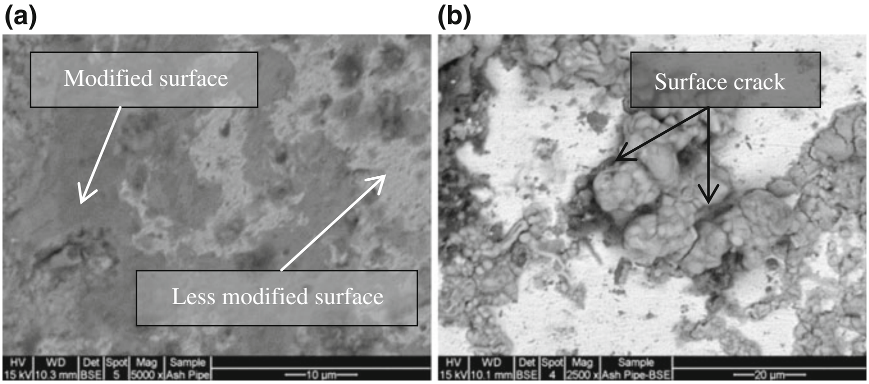


Fig. 14.6 BSE images of typical surface areas of fly ash conveying pipeline. **a** Initial stage of modification and **b** final stage of modification

elements and helps to distinguish different elements present on the surface. A brighter surface area represents the presence of higher atomic number elements on the surface, and darker areas represent the presence of lower atomic number elements.

Figure 14.6 presents BSE images from two different areas of wear surface. Figure 14.6a shows a smooth surface area with lighter and darker appearances. As discussed earlier, darker surface areas represent lower molecular weight elements than the brighter areas. It is understood that the mild steel (Fe) surface indicated by the brighter appearance, whereas the surface modified by lower molecular weight elements such as silicon (Si) or aluminum (Al) or oxygen (O_2) appears darker. Figure 14.6b shows an area with darker appearance with cracks and seems to be attached loosely with the surface. It is understood that the modified surface seen in Fig. 14.6a has been further modified and transformed into texture observed in the Fig. 14.6b.

Figure 14.7 shows the elemental spectrum from the representative brighter and darker areas of the surface shown in Fig. 14.6a, b. Both Fig. 14.7a, b reveal the presence of a number of foreign elements in the spectrum other than the parent material mild steel (Fe). Some of which might be present in any mild steel samples such as sulfur (S), titanium (Ti), and manganese (Mn). Strong presence of aluminum (Al) and silicon (Si) in (Fig. 14.7b) in the gray area of Fig. 14.6b was believed to be the modifying agent of the surface.

It appears that the brittle characteristic of the surface is related to the surface modifications by the element from the conveying material. Although the exact chemical structure of the compounds is not known, it is apparent that surface modification by alumina and silica is responsible for the change in wear mechanisms in these areas.

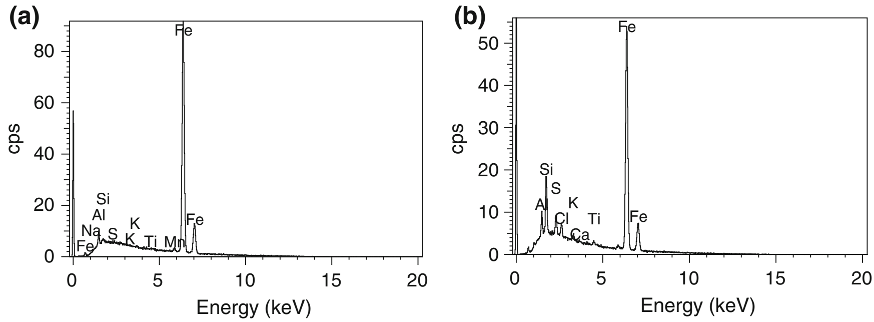


Fig. 14.7 EDS elemental spectrum from pipeline wear surface, **a** spectrum from surface *brighter area* and **b** spectrum from *darker surface area*

14.4 Conclusions

Wear is complex phenomena of material removal in pneumatic conveying pipelines. Although the interactions between surface material and conveying particles and ensuing wear mechanisms are better understood over the years, materials' behavior due to surface modification and related wear mechanisms are less understood.

Surface modification and its effects on wear mechanisms in mild steel pipeline conveying fly ash have been presented in this paper. From the analysis of wear mechanisms, it is evident that the pipeline surface characteristics have been altered from ductile to brittle due to surface modification by the constituents of the conveying material. It is believed that this modification plays an important role in controlling the service life of conveying pipeline. Surface modifications involved transfer of element from the conveying particles into the surface material. Actual mechanism(s) of surface modification and chemical structures of the modified surface remained to be investigated.

Acknowledgments This work forms part of the research program for the CRC for Integrated Engineering Asset Management. The authors also acknowledge the EM-Xray unit of the University of Newcastle for their continual assistance in SEM/EDS analyses.

References

1. Cenna AA, Williams KC, Jones MG, Page NW (2009) Generation of transfer film and its effect on wear mechanisms in alumina conveying pipeline of mild steel. Wear 267:362–367
2. Mills D, Jones M, Agarwal V (2004) Handbook of pneumatic conveying engineering. Marcel Dekker Inc, New York
3. Field JE, Hutchings IM (1987) Surface response to impact in materials at high strain rates, Blazynski TZB (ed) Elsevier Applied Science
4. Moore MA, King FS (1980) Abrasive wear of brittle solids. Wear 60:123–140

5. Cenna AA, Williams KC, Jones MG, Robinson W (2011) Analysis of wear mechanisms in pneumatic conveying pipelines of fly ash. Presented at WCEAM, 2011
6. Cenna AA, Williams KC, Jones MG, Page NW (2006) Flow visualisation in dense phase pneumatic conveying of alumina. Presented at the inaugural world congress on engineering asset management, Gold Coast, Australia, 11–14 July 2006

Chapter 15
Infrastructure Asset Management for Different Types of Facilities Using Normalized Level of Service

Myung Jin Chae

Abstract Infrastructure is understood as a lifeline network for the society. As facilities are connected with each other, failure or poor performance of a facility in the network affects the rest of the network resulting in the overall performance deterioration. Social infrastructures such as roads, bridges, water supply, wastewater treatment, parks, and public buildings are all connected each other. Failure and or poor performance of an infrastructure would result the poor-performance infrastructure network. In order to keep the overall performance of the infrastructure network at certain level, the performance level of different facilities should be maintained in balance, i.e., performance leveling. Managing different facility in same level of performance is like comparing the qualities of oranges and apples. However, the facility managers must compare the performance of roads, bridges, water supplies, public buildings and parks in order to allocate appropriate budget to maintain those facilities at almost same level of performance. Authors have developed a method to normalize the level of service of infrastructure to compare different type of infrastructure performing at same level of service. First, levels of service indicators have been identified; then, they have been converted into normalized values, which are used as cross-asset management. Normalization factors were derived from surveying of users and statistically calibrated to remove the biases. This paper describes the development of level of service indicators and normalization methods for cross-asset management.

M. J. Chae (✉)
Construction Engineering and Management Research Division, Korea Institute of Construction Technology, Goyang-si, South Korea
e-mail: myungjinchae@gmail.com

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_15,

15.1 Introduction

The definition of infrastructure asset-related service levels has a comprehensive meaning. Level of Service (LoS) is a defined service quality for a particular activity or service area against which service performance may be measured [1]. Service levels usually relate to quality, quantity, reliability, responsiveness, environmental acceptability, and cost.

The application of customer-oriented approach is also found in the transportation area. "Customer Focus in Highway Transportation Asset Management (HTAM)" included economical, political, environmental, social, and technical considerations. Orndoff [2] presented a methodology that provides a tool for decision makers in other engineering areas to increase customer focus prior to enacting policies or projects. Customer asset value was determined through contingent value methodology (CVM).

INGENIUM and the National Asset management Steering (NAMS) Group of New Zealand (2007) suggested a systematic method to develop a level of service based on the customers' perspective. The purpose of the manual aims to assist organizations deliver better customer value through reviewing the level of service provision in consultation with their customers and monitoring performance to ensure they deliver the agreed levels of service.

Utility theory is one of the widely used methods to prioritize and make difficult decisions systematically. In relation to the risk management strategy, utility values are used to measure the risk values as a subjective risk management strategy. This paper describes the method to measure the level of services converted into the utility values that assess the level of service of different kinds of facilities.

15.2 Utility Theory

Utility is a term to represent the use of goods or service. It is usually possible to rank the alternatives in their order of preferences. While the dollar value is the direct measure of value the goods and service, the utility values are different from different persons and entities because customers' needs for goods or services for a certain dollar values can be different. When utility is applied in risk management and decision-making process, it represents the value of the risk or decisions made.

A utility function measures an investor's relative preference for different levels of total wealth. For example, usually the wealth value of a property increases with the utility increases, i.e., more wealth is preferred to less wealth. However, the investor does not always have so much utility for investing certain amount of the wealth. Figure 15.1 shows the risk averse attitude of the investor; relationship between wealth and utility is not linear.

The similar situation can be found in everyday lives of common people. When a customer purchases a car, the more he pays, the higher performance and more

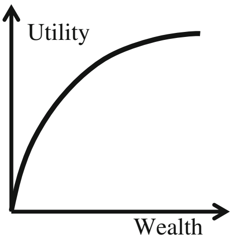


Fig. 15.1 Utility function of risk aversion

luxury car he can get. But he has to pay much more wealth to get relatively not so much performance improvement and luxury options. Everyone has different financial situations and preferences, and he/she would make decision at certain point that he/she can afford and maximum utility.

The utility theory can be used for investment decision for this reason. All the investors are in different financial situations. Investors would use a utility function that translates from dollar to utility, instead of maximizing expected value, maximizing expected utility.

15.3 Assessment of LOS and Conversion to Utility Value

Due to the different interests and situations of communities, there is not common standard method to assess the level of service of public facilities. For example, snow removal and good road condition maintenance are the most important issue for the communities in mountain area with heavy snow, while safety of levees and banks is considered as the most critical facilities in coastal areas where the hurricanes are frequent. Traffic congestions and air pollutions are usually considered as important issues in most populated cities.

Decision makers of public facilities have to determine how they invest maintenance budget. Investing more on roads and bridges or public parks or water reservoirs is often determined by non-expert managerial people rather than facility managers and engineers. In fact, expert engineers can only tell which facilities need immediate maintenance activities for safety issue but they cannot tell how to make a balanced investment for different kinds of facilities in order to improve the overall satisfaction of the community.

In this paper, it is discussed how to make the different levels of services of facilities can be assessed using utility functions. The community can decide what facilities are more important than the others, and the weight values are assigned. Using utility functions, performance levels are converted into the utility values.

For example, each levels of service is rated from A, B, C, D, and E, where A is the best and E is the worst. Preference levels from E to D; D to C; C to B; and B to A are to be measured. Required energy or preference level to improve one step

Table 15.1 Utility function converted from performance level

Performance level indicator	A	B	C	D	E
Performance level description	Pavement in perfect condition	Minor cracks on pavement	Medium cracks minor potholes	Large potholes and cracks. drivers must reduce speed	Must reduce speed to avoid accident
Order of the importance	4	3		2	1
Importance weight factors	x_1	x_2		x_3	x_4
Sum of the importance weights	$\Sigma = 10$				
Utility values	10/Σ	9/Σ	7/Σ	3/Σ	3/Σ

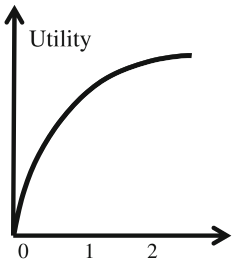


Fig. 15.2 Utility function for performance level

from lower level to the higher level is determined. Then, the levels of preferences are collected, and the line is drawn to construct the utility function.

$$
\begin{aligned}
u_i(x_{iA}) - u_i(x_{iB}) &= \alpha[u_i(x_{iD}) - u_i(x_{iE})] \\
u_i(x_{iB}) - u_i(x_{iC}) &= \beta[u_i(x_{iD}) - u_i(x_{iE})] \\
u_i(x_{iC}) - u_i(x_{iD}) &= \gamma[u_i(x_{iD}) - u_i(x_{iE})]
\end{aligned}
\tag{15.1}
$$

where x_{iA} is ith performance level *and* u_i is the utility values for ith performance indicator at certain performance level from A to E. α, β, and γ are the relative importance of changes of the performance levels in regards as lowest changes in levels, which is $u_i(x_{iD}) - u_i(x_{iE})$ in Eq. 15.1.

Table 15.1 is an example of performance scoring table for road pavement condition. Importance levels are the preference levels to improve from one level to the other. Utility values are determined in the same way. Once this table is constructed and utility values are found, utility function is drawn as Fig. 15.2 where horizontal axis is performance level and vertical axis is utility.

This performance level to utility value conversion can also be used in cross-asset management, i.e., comparing different kinds of the assets in order to allocate maintenance budget based on the performance data, level of service, and community needs through utility values. Because the utility values are in non-dimensional unit-free number, utility values can be set in any number ranges such as 0–1 or 0–100 scale. The utility values driven from different kinds of facilities such as bridges, roads, parks, and public buildings are in same scale levels; thus, they can be automatically normalized and can be compared without further normalization process.

In the same way, different facilities can be ordered and scored in utility values. For example, in certain community, road condition maintenance is more important than levees and banks. Environmental protection and water resource management can be considered more important than reducing traffic jams. This is often determined by the city mayor and city council or engineers of facility maintenance department and/or public works department. With the help of utility theory and level of service maintenance method, systematic and scientific approach to allocate the maintenance budge is possible.

15.4 Summary and Conclusion

This paper discussed the possible use of utility theory for the normalization of level of services in public infrastructures. Performance levels are converted into utility values so the performance indicators are in a common scale of utility values. As the same way, different facilities are compared using utility value scoring and ordering, and thus, the decision makers can tell which asset is more important than the others. Systematically allocating maintenance budget over various assets can be done by the method presented in this paper.

However, there should be more research done to prove the result of the cross-asset management using utility theory. Level of service should be developed, and utility values should be driven using the real data to prove the methodology presented in this paper.

References

1. InfraGuide (2004) Managing infrastructure assets, decision making and investment planning (DMIP) best practice. National Research Council of Canada (NRC), Ottawa, Canada
2. Orndoff CJW (2001) Customer focus in highway transportation asset management. Ph.D. thesis, University of Illinois at Urbana-Champaign, Urbana, IL

Chapter 16
A Standards-Based Approach for Interoperability Between Engineering Design and Asset Management

Dukki Chung, Dan L. Carnahan and Em G. delaHostria

Abstract There is a wealth of asset information in the design environment that is necessary in the operational environment to effectively manage the asset. However, the asset design information model is typically a complex structure that requires in-depth knowledge of the model to access the necessary information. Currently, there are no well-defined methodologies to extract the relevant design information that is needed in the asset operational domain. In addition, there are no well-defined methods to define the relationships between the terms and definitions common to the design environment and the operational environment. In this paper, we will propose the use of an open technical dictionary and an information exchange approach based on ISO 18435 to organize the relevant information to achieve interoperability of applications in both design and operational environments.

16.1 Introduction

Within a manufacturing enterprise, there exists a need to increase the visibility of asset management operations regarding information about not only the current state of the assets, but also the capability of those assets intended by design to meet changing manufacturing requirements. These changing requirements can be due to new customer demands, process upsets, or changes in either asset or process capabilities. The response of an asset management system, in response to the new requirements, depends on having current information about the status and capability of the assets comprising the automation system configuration that enables the manufacturing applications. These asset information structures can be available if

D. Chung (✉) · D. L. Carnahan · E. G. delaHostria
Rockwell Automation, 1 Allen-Bradley Dr., Mayfield Heights, OH 44124, USA
e-mail: dchung@ra.rockwell.com

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_16,

appropriate interfaces are enabled among the resources deployed in a manufacturing system.

A new effort to integrate the design information with operational information was recently initiated as an ISO project (ISO 18101). This effort will use existing industry data sets and use cases to delineate the use of existing standards, such as ISO 8000 (data quality), ISO 13374 (conditioned-based maintenance), ISO 15926 (reference design information), ISO 18435 (diagnostics and production information integration), ISO 22745 (open technical dictionary), and IEC 62264 (enterprise–control integration).

Many terms and definitions associated with an asset commonly used in both the design context and the operational context are often ambiguous or inconsistent. Their interpretation or meaning depends on the particular context. The open technical dictionary (OTD) based on ISO 22745 facilitates the resolution in the ambiguity or inconsistency of these terms and definitions.

ISO 18435 provides a framework for harmonized use of selected industry and international standards in order to integrate control, diagnostics, prognostics, capability assessment, and maintenance applications. By using a standards-based application integration modeling approach, information exchange requirements for key interoperability interfaces can be identified and concisely documented in terms of application-specific information exchange matrices. These matrices enumerate a set of particular standards for enabling context-, content-, and conveyance-oriented exchanges to enable the asset interoperability and integration of applications dealing with asset information structures. In the oil and gas industries, ISO 15926 standards provide as-designed and as-built information of the engineering design models of the assets used in the operation domain. The information models based on ISO 15926 can be utilized in conjunction with the ISO 18435 framework to enable information exchange among control, diagnostics, and maintenance applications during operation of the deployed assets. The use of ISO 22745-based OTD concept identifiers and definitions to enumerate the ISO 15926 asset information structures in the ISO 18435 information exchange matrices can facilitate the proper handling of the design and operation contexts to achieve interoperability of the applications.

16.2 Related Standards for Integration and Interoperability

16.2.1 ISO 15745—Application Integration Framework

Within each level of an application hierarchy in an enterprise, a set of applications are deployed to provide the functions required at that level. For each application, a set of resources are used to conduct the related processes and to perform the information exchanges. ISO 15745 standard [3] defines a method to construct an application integration model identifying the interfaces required to integrate the resources. An integration model is constructed to represent several component aspects of an application—a set of processes with associated sets of resources that perform a set of information exchanges to support the execution and coordination of the processes. In Fig. 16.1, these manufacturing aspects (process, resource, and information exchange) can be profiled to describe the interoperability aspects.

For the manufacturing application, an application interoperability profile can be constructed that delineates the information exchange supported by the resources used in the application. A resource profile can also be constructed to enumerate the capability requirements for the information exchanges.

16.2.2 IEC 62264—Enterprise–Control System Integration Standard

A joint working group of the IEC SC65E and the ISO TC184/SC5 technical subcommittees recently developed a multi-part IEC 62264 standard [1]. IEC 62264 is based on the ANSI/ISA-95 specification [2], which defines an information exchange framework to facilitate the integration of business applications and manufacturing control applications, within an enterprise. In the ANSI/ISA-95 standard, the applications within a manufacturing enterprise are structured as a hierarchy of activity domains. At each level in the hierarchy, a group of functions are performed to support a specific operational level of an enterprise. In addition, the standard also defines an equipment hierarchy that distinguishes in which physical grouping and at which organizational level a piece of equipment is being used. Figure 16.2 shows the functional hierarchy defined in IEC 62264.

16.2.3 ISO 18435—Industrial Automation Systems Integration

ISO 18435 [5] is intended to provide a framework for harmonized use of selected industry and international standards in order to integrate control, diagnostics, prognostics, capability assessment, and maintenance applications. By using an ISO 15745 application integration modeling approach, key interoperability interfaces can be identified and concisely documented in terms of profiles. These application interoperability profiles can be used when evaluating whether applications can readily integrate with each other. In Fig. 16.2, an Activity Domain Integration Diagram (ADID) identifies those activities associated with the applications of interest. The R-levels in this diagram relate to a resource hierarchy deployed among the applications within a manufacturing facility.

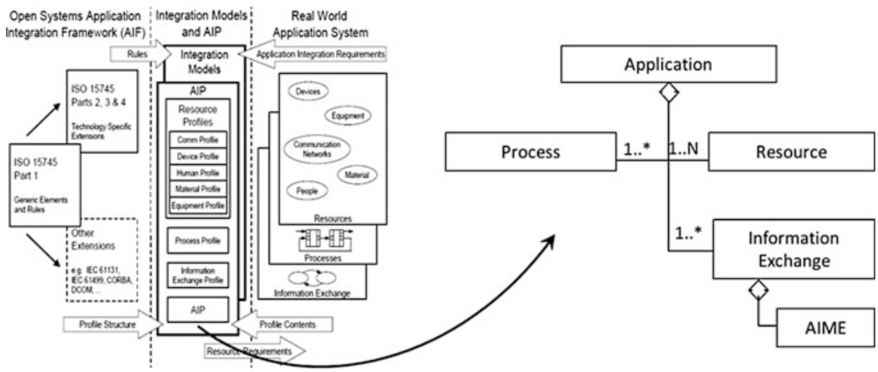
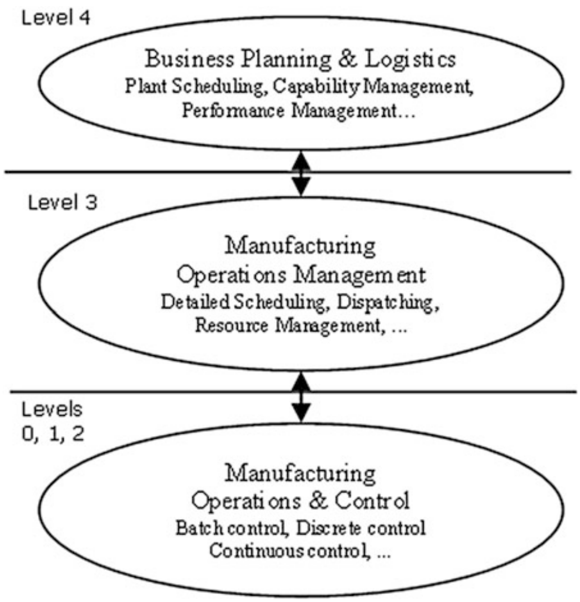


Fig. 16.1 Integration model class diagram derived from ISO 15745

Fig. 16.2 Functional hierarchy per IEC 62264



The purpose of the ADME is to describe the information exchange requirements of the applications. For each application, interfaces used for information exchange are described using the Application Interaction Matrix Element (AIME). The AIME details the resource capabilities that meet the information exchange requirements to support the interoperability of two applications. A set of AIMEs represents the interface profiles supported by the applications and the corresponding resources, and these AIMEs comprise an ADME. An ADME and AIMEs that qualify interoperability relationship between two applications are shown in Figs. 16.3 and 16.4.

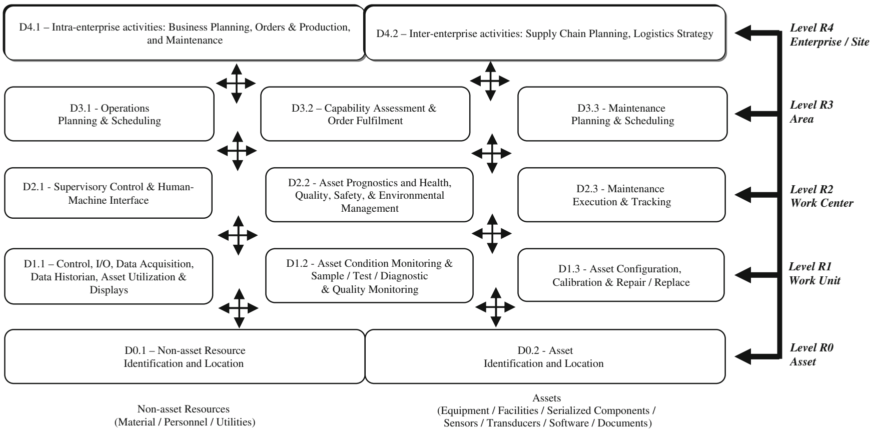


Fig. 16.3 Application domain integration diagram from ISO 18435

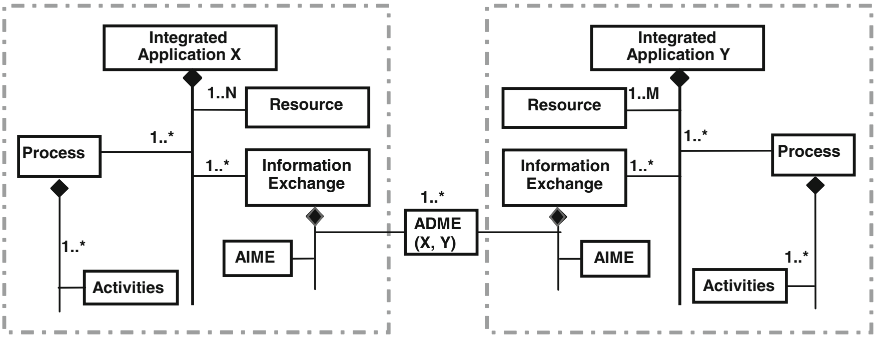


Fig. 16.4 Interoperability of applications using ADME and AIMEs

16.2.4 ISO 22745—Open Technical Dictionaries and Their Application to Master Data

The ISO 22745 standard [6] provides the framework needed for any organization to conduct business with internationally recognized data quality (ISO 8000). ISO 22745 standard provides a means to realize the benefits of ISO 8000 (data quality) by specifying the data requirements for messages containing master data that is exchanged between organizations, specifically requirements for syntax, semantic encoding, and portability.

ISO 22745's primary facilitator is the open technical dictionary (OTD), a database of concept identifiers and associated descriptive words used to define individual data elements. Once each element is described with the concept identifier from the OTD, the descriptive elements can be stored, sent, received, and displayed by different organizations without losing any meaning.

ISO 22745 also includes guidelines for the use of identification guides (IG). An identification guide is a statement of requirements describing what data is needed about an item. If all elements are included in the description, this IG facilitates the machine-aided analysis of data quality because we have a clear understanding of what data is required without a person having to review the data.

16.2.5 ISO 15926—Integration of Life-cycle Data for Process Plants Including Oil and Gas Production Facilities

The ISO 15926 [4] is a standard for the electronic sharing, exchange, and integration of plant design and life-cycle data. Information concerning engineering, construction, and operation of production facilities is created, used, and modified by many different organizations throughout a facility's lifetime. The purpose of ISO 15926 is to facilitate integration of asset information to support the life-cycle activities and processes of production facilities. ISO 15926 provides a model and library classes and templates for representing life-cycle information about technical installation and their components. The standard avoids fixed schemas, to accommodate change and development of industry data.

16.3 Proposed Approach for Interoperability

The increasing number of applications is accompanied by an increase of proprietary information exchange specifications between applications. One of the challenges for providers of asset management solutions is to achieve, at the lowest cost possible, the interoperability of application for the design, operation, and maintenance of the assets, without sacrificing the freedom to choose from the entire spectrum of available applications. This is best achieved when correct, complete, and unambiguous information exchange standards are implemented.

ISO 18435 specifies provisions that enterprise applications are expected to satisfy, in terms of a set of interoperability profiles. For example, if a diagnostic application requires flow information from the control application controlling a pump to assess the overall asset health, these two applications must have compatible profiles for this particular information exchange.

ISO 15926 provides design information and installation information of the engineering assets of the production facility. Figure 16.5 shows an example representation of a pump model using ISO 15926.

The pump model and installation which identify the physical pump "Pump 101" have flow information available since the model and the particular installation have a flowmeter attached to the motor starter for the pump. Furthermore,

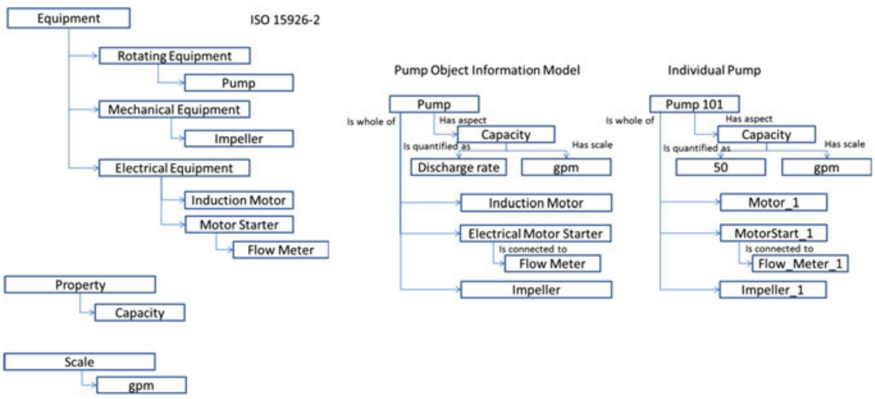


Fig. 16.5 Pump model in ISO 15926

"Capacity" information of the pump in the actual installation can be used to support the pump diagnostics. Interoperability requires the integration of information over multiple levels of representation. In this example, the context around "Capacity" needs to be resolved to convey the meaning accurately. By specifying the right context in the ADMEs and AIMEs using OTD and an associated identification guide, one can reduce the inconsistency and ambiguity.

By examining one of the AIMEs representing resource capability of the control application, it can be identified that one of the resources of the control application "MotorStart_1" acquires flow information from the flowmeter. The conveyance section of the same AIME also reveals an industrial network protocol, e.g., Ethernet/IP, available for conveying flow information, as shown in Fig. 16.6.

A partial AIME from one of the resources of the pump diagnostics application is shown in Fig. 16.7. By examining these AIMEs further, it can be determined that these two applications can exchange information on flow using the same protocol.

Even though the flow information appears in both profiles, the contexts used in these applications could be different. For example, the pump control application could use the context from IEC 61131 standard. On the other hand, the diagnostics application could use the context driven from ISO 13374 standards. The differences in context need to be resolved to fulfill the accurate information exchange. Utilizing OTD based on ISO 22745 resolves this ambiguity. Many terms and definitions can be identified without ambiguity using concept identifiers from the OTD. Figure 16.8 shows that all the information types are replaced with the concept identifiers in AIMEs. Once the AIMEs are established, an ADME can be constructed between the applications and interoperability can be ensured.

```
<Conveyance_Section>
…
     <description>PumpControl Example</description>
     <informationType name="DischargeRateReqType" type="tDischargeRateRequest">
            <description>Pump Discharge Rate Request Information Type</description>
     </informationType>
     <informationType name="DischargeRateResType" type="tDischargeRateResponse">
            <description>Pump Discharge Rate Response Information Type</description>
     </informationType>
…
<channelType name=”ControlChannel” type=”ISO15745_ENet_CommNet_Profile”>
        <description>Ethernet/IP based on ISO15745-4 Profile</description>
</channelType>
…
```

Fig. 16.6 Example of conveyance section of pump control AIME

```
<Conveyance_Section>
…
     <description>Pump Monitor Example</description>
     <informationType name="FlowRateReqType" type="tFlowRateRequest">
            <description>Pump Flow Rate Request Information Type</description>
     </informationType>
     <informationType name="FlowRateResType" type="tFlowRateResponse">
            <description>Pump Flow Rate Response Information Type</description>
     </informationType>
…
<channelType name=”MonitorChannel” type=”ISO15745_ENet_CommNet_Profile”>
        <description>Ethernet/IP based on ISO15745-4 Profile</description>
</channelType>
…
```

Fig. 16.7 Example of conveyance section of pump diagnostics AIME

```
<Conveyance_Section>
  …
  <description>Control Example</description>
  <informationType name=”DischargeRateReqType”
       type=”0161-1#01-080769#1”>
  <description> Discharge Rate Req.</description>
  </informationType>
  <informationType name=”DischargeRateResType”
       type=”0161-1#01-018900#1”>
  <description>Discharge Rate Res.</description>
  </informationType>
  …
  <channelType name=”ControlChannel”
       type=”0161-1#01-1074539#1
  <description>Ethernet/IP ISO15745-4 Profile
    </description>
  </channelType>
  …
```

```
<Conveyance_Section>
  …
  <description>Monitor Example</description>
  <informationType name="FlowRateReqType"
       type="0161-1#01-080769#1">
  <description> Flow Rate Req.</description>
  </informationType>
  <informationType name="FlowRateResType"
       type="0161-1#01-018900#1">
  <description>Flow Rate Res.</description>
  </informationType>
  …
  <channelType name=”MonitorChannel”
       type=”0161-1#01-1074539#1”>
  <description>Ethernet/IP ISO15745-4 Profile
    </description>
  </channelType>
  …
```

Fig. 16.8 ISO 18435 AIMEs with concept identifiers from OTD based on ISO 22745

16.4 Conclusion

The integration of applications for asset operation and management can be described using the combined aspects of ISO 18435, ISO 22745, and ISO 15926 standards. The resource design information models, following the provisions of ISO 15926, are mapped to the open technical dictionary (OTD). ISO 22745(OTD) provides for a method to unambiguously establish relationships between common terms and definitions used in asset design, and operations, and maintenance phases. The information associated with the production and maintenance operations involving the application role-based resources and physical assets is designated and enumerated according to the hierarchies and activity models of IEC 62264. These operational information models are also mapped to the ISO 22745-based OTD. Using the application integration concepts of ISO 15745, the interoperability profiles for the resources, information exchanges, and processes of the design, operations, and maintenance applications are constructed. These ISO 15745 profiles delineate the interoperability interfaces needed to enable the exchanges of information that are organized by using ISO 18435. The use of the ISO 18435 interoperability methodology based on AIMEs and ADMEs enables asset management systems to effectively harness the capabilities of manufacturing assets which have been detailed at design time and deployed at operational phase in providing better visibility into the "Design–Operate–Maintain" life cycle of automation assets.

References

1. IEC (2003) IEC 62264: Enterprise-control systems and integration. Geneva, Switzerland
2. ISA (2010) ANSI/ISA-95: Enterprise-Control System Integration. Raleigh, North Carolina, USA
3. ISO (2003) ISO 15745: Industrial automation systems and integration—open systems application integration framework. Geneva, Switzerland
4. ISO (2004) ISO 15926: Industrial automation systems and integration—integration of life-cycle data for process plants including oil and gas production facilities. Geneva, Switzerland
5. ISO (2009) ISO 18435: Industrial automation systems and integration—diagnostics, capability assessment and maintenance applications integration. Geneva, Switzerland
6. ISO (2010) ISO 22745: Industrial automation systems and integration—open technical dictionaries and their application to master data, Geneva, Switzerland

Chapter 17
Structural Complexity in Structural Health Monitoring: Design of Laboratory Model and Test Plan

Craig J. L. Cowled, David P. Thambiratnam, Tommy H. T. Chan and Andy C. C. Tan

Abstract Many researchers in the field of civil structural health monitoring (SHM) have developed and tested their methods on simple to moderately complex laboratory structures such as beams, plates, frames, and trusses. Fieldwork has also been conducted by many researchers and practitioners on more complex operating bridges. Most laboratory structures do not adequately replicate the complexity of truss bridges. Informed by a brief review of the literature, this paper documents the design and proposed test plan of a structurally complex laboratory bridge model that has been specifically designed for the purpose of SHM research. Preliminary results have been presented in the companion paper.

17.1 Introduction

The structural complexity of operational bridges poses a significant challenge for structural health monitoring (SHM) researchers. Reference [1] highlights that a '*primary source of epistemic uncertainty* [in structural identification] *is related to the relatively high level of structural complexity typical of constructed systems*,' (p. 406). Ciloglu [2] found that structural complexity contributes significantly to the uncertainty of structural identification by operational modal analysis (OMA). Aktan et al. [3] argue that the basic assumptions that enable system identification of structures do not hold true for more complex constructed systems, particularly in a climate with daily temperature fluctuations of more than 10 °C where temperature and humidity can have a significant effect on the vibration characteristics of an operational structure.

C. J. L. Cowled (✉) · D. P. Thambiratnam · T. H. T. Chan · A. C. C. Tan
Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: craig.cowled@qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_17,

Considering that one of the key goals of the concerted worldwide research efforts into SHM and damage detection (DD) is to be able to monitor complex operational structures and diagnose their health state, it stands to reason that experimental testing should likewise focus on similarly complex laboratory structures. Tee et al. [4] have noted a trend in research toward increasingly '*large systems with many unknown parameters*,' (p. 398). Sami Masri (as cited in [5]) has called for multiple benchmark problems with increasing complexity. The evidence in the literature therefore supports a move toward using more complex structures in the laboratory for SHM research.

This paper reviews some of the structures in the SHM literature and presents a structurally complex model of a bridge structure that has been designed by the primary author, fabricated by Taringa Steel Pty. Ltd. and assembled at Queensland University of Technology (QUT). This bridge model has been designed as a benchmark structure for SHM researchers at QUT to test DD methods and conduct SHM research. This paper will also outline the proposed test plan of the QUT Benchmark Structure. Preliminary results will be presented in a companion paper [6].

17.2 Structural Complexity in the Literature

17.2.1 Introduction

The structural complexity in the SHM literature is first reviewed in this section. The review is divided into three types of structures, namely numerical models, laboratory structures, and operational structures. The QUT Benchmark Structure is then presented as the most complex laboratory structure of its kind. The term 'structural complexity' is used in this paper to encompass the complexities of geometry, connections, and boundary conditions.

17.2.2 SHM of Numerical Models

Rather than spending time and money on laboratory experiments or fieldwork, most SHM researchers employ numerical models to test their DD methods. These numerical simulations often attempt to replicate real-world environmental and operational variations and measurement noise by introducing random noise of some kind into the data.

When testing their DD methods, these researchers tend to use simple models, often in two dimensions only. Common examples include cantilevered beams, plates, simple frames, and simple trusses. This approach is understandably necessary for the initial development of a new method of DD. As the method is refined and developed further, however, it can be argued that it ought to be tested against increasingly complex models as a test of its robustness. Given that the complexity of numerical models is limited only by one's imagination and the processing power of one's computer, it is somewhat surprising that relatively few researchers attempt to test their methods on structurally complex models.

Following is a review of the literature focusing on the degree of complexity in *numerical models*.

Yan et al. [7] used a modal strain energy sensitivity-based method on three numerical models: a simply supported beam (15 elements), a continuous beam (30 elements), and a three-storey planar frame (33 elements). Mehrjoo et al. [8] conducted a numerical analysis using artificial neural networks (ANN) to detect damage at the connections of two truss structures. Both structures were planar trusses with 7 members in the first structure and 29 members in the second structure. Yan et al. [9] tested their wavelet-based DD method on a numerical model of a three-dimensional, five-storey structure with 55 members. Rama Mohan Rao et al. [10] developed a DD technique combining proper orthogonal decomposition with a self-adaptive differential evolution algorithm which was tested on numerical models of a cantilevered beam (20 elements), a simply supported plate (342 elements), and a planar truss having 55 members.

All of the aforementioned examples have used relatively simple numerical models, often limited to two dimensions, to test their DD methods.

17.2.3 SHM of Laboratory Structures

When testing their DD methods in the laboratory, most researchers tend to use structures that are simple to moderately complex. There are very few examples of structurally complex laboratory structures in the SHM literature.

Following is a review of the literature focusing on the degree of complexity in *laboratory structures*.

Samali et al. [11] conducted a detailed study of a laboratory timber structure that consisted of four F11 radiata pine girders and an F11 structural plywood deck (4.5 m × 2.4 m × 0.1 m). Papatheou et al. [12] investigated the feasibility of using known masses to guide feature extraction on an aluminum structure consisting of two C section ribs, two angle section stringers and a 3-mm-thick plate (0.75 m × 0.5 m × 0.05 m). Quek et al. [13] tested the damage locating vector (DLV) method on an aluminum space truss with 24 members (3 m × 1 m × 0.7 m). Weber and Paultre [14] used a sensitivity-based damage identification method to detect damage on a 112 member tower truss (0.3 m × 0.3 m × 2.8 m). Gao et al. [15] conducted experimental verification of the DLV method on a 156 member space truss. Meruane and Heylen [16] investigated the use of antiresonances in DD on a three-dimensional aluminum space truss with 43 members (3 m × 0.5 m × 0.5 m). Dackermann et al. [17] used neural network ensembles and the conventional damage index method in their numerical (200 elements) and experimental study of simply supported steel beams (2.4 m × 0.032 m × 0.012 m). Catbas et al. [18] presented a benchmark problem intended to represent a large proportion of the bridge stock, namely short- to medium-span highway bridges. The 15-member steel structure, located at the University of Central Florida, consists of two parallel longitudinal beams connected by seven transverse beams and supported on six columns (5.48 m × 1.83 m × 1.07 m). The finite element (FE) model of this moderately complex structure consists of 181 elements and 1,056 DOFs. Johnson et al. [19] introduced a benchmark problem based on the complex four-storey steel frame structure (2.5 m × 2.5 m × 3.6 m) located at the University of British Columbia. This structure has 112 members. Ciloglu [2] conducted detailed structural identification of a simply supported steel grid with composite sandwich plate deck (6.1 m × 2.75 m) supported by six steel pedestal frames. The steel grid consisted of three longitudinal beams and 14 transverse members made from rectangular hollow steel sections bolted together with gusset plates. The composite deck consisted of a sandwich plate with a balsa wood core and E-glass sheets top and bottom bonded together with a vinyl ester resin. The boundary conditions of the fabricated steel and composite deck consisted of roller bearings fixed to the pedestal frames. The structure is located in the Drexel Intelligent Infrastructure Laboratory in Philadelphia, PA. Caicedo [20] undertook an experimental study to verify the parameter identification SHM technique, using a stainless steel model of a cable-stayed bridge structure (2 m × 0.29 m × 0.5 m). The main tower of the structure is an H frame and holds 60 cables that support the deck. The deck consists of 30 transverse members, 60 longitudinal members, and 28 lead masses. Although the structure has 379 DOFs, only the last six elements of the structure, totaling just 12 DOFs, were studied.

All of the aforementioned examples demonstrate that researchers have used laboratory structures that are relatively simple to moderately complex in their research.

17.2.4 SHM of Operational Structures

There are a wide variety of operational bridge structures that have been instrumented, analyzed and reported in the SHM literature, ranging from relatively simple short-span slab on girder bridges to extremely complex truss and suspension structures.

Following is a review of the literature focusing on the degree of complexity in *operational structures*.

Thambiratnam [21] conducted a study of the vibration characteristics of an approach span of the Story Bridge in Brisbane, a deck on simply supported steel truss structure, and compared the results with an FE model to show that the structure was operating within limits for strength and serviceability. James et al. [22] detail tests and present some results of DD on the I40 Bridge over the Rio Grande in Albuquerque, New Mexico, USA. This three-span steel plate girder bridge was subjected to various damage scenarios. Krämer et al. [23] presented DD tests on the widely documented Z24 Bridge in Switzerland. This three-span prestressed concrete box girder bridge became a benchmark structure [24]. Heywood et al. [25] used strain and deflection data along with a code-based analysis to assess the fitness-for-purpose of four steel bridges, of various configurations, in Australia and New Zealand. Senthilvasan et al. [26] tested the dynamic response of a multi-span-curved pre-stressed concrete box girder bridge under a truck load. Results were used to validate a FE model and to compare the experimentally determined dynamic strains with code predictions. It was evident that the code predictions underestimated the dynamic strains. Catbas et al. [27] conducted structural identification studies on the Commodore Barry Bridge spanning the Delaware River in the USA. Agrawal et al. [28] introduced the 91/5 highway bridge as a benchmark structure. This structure is a prestressed concrete box girder bridge with two spans. The FE model of the actual structure is very simple; however, the boundary conditions are quite complex. Chan et al. [29] describe the extensive wind and structural health monitoring system (WASHMS) installed on the famous Tsing Ma Bridge in Hong Kong, a long-spanning suspension bridge. The WASHMS is described as the most comprehensive SHM system installed on any operational structure in the world.

All these examples demonstrate that operational structures in the SHM literature tend to be far more structurally complex than numerical models or experimental structures in the literature.

17.2.5 Structural Complexity of the QUT Benchmark Structure

The most structurally complex laboratory structure in the SHM literature is the one that has been designed, fabricated, and assembled for this study (8.55 m × 0.9 m × 2.6 m, see Fig. 17.1 overleaf). This structure is esthetically similar to Brisbane's Story Bridge, which is a long-span cantilevered steel through truss bridge. Not counting gusset plates, the deck plate or support members, the superstructure of this bridge model has 314 members in a configuration with 100 nodes. There are 600 DOFs in the superstructure plus additional DOFs in the supports. The first mention of this bridge model in the literature can be found in [30] where a FE model of the structure was developed to test a correlation-based DD method using modal strain energy and optimized via a multilayer genetic algorithm.

17.2.6 Conclusion

This section of the paper constituted a review of structural complexity in the literature of numerical models, laboratory structures, and operational structures. This review confirmed the gap in structural complexity between examples of

Fig. 17.1 The QUT Benchmark structure

numerical models / laboratory structures in the literature and operational structures in the literature.

Finally, this review presented the QUT Benchmark Structure as the most structurally complex steel through truss bridge model in the SHM literature, having complexities of geometry, connection detailing, and boundary conditions.

17.3 QUT Benchmark Structure

Benchmark problems in SHM allow researchers to objectively compare and contrast DD methods and have also been shown to highlight the limitations of current techniques [12]. The previous section of this paper advocated the need for increasing structural complexity in laboratory structures in the field of SHM research.

This section of the paper outlines and discusses the design, specifications, and assembly of the QUT Benchmark Structure (see Fig. 17.1). The 600+ DOF bridge model is structurally complex and has been designed for the purpose of studying SHM and DD. The structure is raised approximately 800 mm off the concrete slab to allow room for static load tests and modal tests using a shaker.

17.3.1 Design

The six main criteria that guided the concept design of the structure were as follows: (1) complex structural form; (2) dynamic characteristics that can be captured by standard sensors and data acquisition equipment at reasonable sampling rates; (3) structural steel materials; (4) bolted connection details; (5) modular design; and (6) detailing that facilitates simulation of damage scenarios.

The geometry of this structure was chosen to resemble the main bridge span of Brisbane's iconic Story Bridge, a cantilevered steel through truss bridge with suspended span. The geometry is also similar to that of the Commodore Barry Bridge, spanning the Delaware River in the USA, which has been extensively studied (e.g., [26]). The structure was designed in accordance with Australian Standard AS4100-1998 [32].

17.3.2 Specifications

A breakdown of the materials that were used in the structure can be found in Table 17.1.

17.4 Proposed Test Plan

Considering the intent to present the structurally complex laboratory structure as a benchmark structure for SHM research, it is prudent to conduct parallel tests using a variety of different methods. Both experimental modal analysis (EMA) and OMA will be examined using two different kinds of excitation within each method. The EMA tests will employ the use of an impact hammer and a shaker as alternative excitation tools. The OMA tests will employ the use of fans and simulated traffic as alternative excitation tools.

In order to reduce the effects of noise and nonlinearities in the results, the tests will be repeated and averaged. The number of repeat tests is expected to be in the order of 10–30; however, the final number will be determined at the time of testing by analyzing the averaged frequency response functions and applying a cost/benefit criteria.

The modal information will be used, in conjunction with a suitable FE model updating method, to validate the FE model of the structure for the purpose of establishing a baseline model of the undamaged structure. A static load test of the undamaged structure, within serviceability limits, will also be conducted to help validate the FE model.

Once the structural identification process is complete, progressive damage scenarios will be applied to the structure and the cycle of testing will proceed.

Table 17.1 Materials

Member type	Qty	Material	Size (mm section)	Comments
Deck	–	MS grade 250	850 × 5.0 PL	Future installation
Deck beams	21	MS grade 350	50 × 25 × 2.0 RHS	
Chords, webs, and struts	215	MS grade 350	20 × 1.6 SHS	
Tower frames	2	MS grade 350	2/30 × 3.0 SHS 2/20 × 1.6 SHS	Fully welded
Bracing	70	MS grade 250	20 × 3.0 FL	
Static loading points	6	MS grade 300+ MS grade 250	Half 250UB31 2/100 × 6.0 FL	Welded to detail +6/M20 pins
Supports	8	MS grade 300+ MS grade 250 MS grade 250	Half 250UB31 4/100 × 6.0 FL 200 × 200 × 10 PL	Welded to detail +8/M20 pins
Gusset plates, cleats, and capping plates	~900	MS grade 250	3.0 PL/FL	Various details
Cleats for high load	20	MS grade 250	5.0 FL	Near tower frames
Bolts	~800	MS 4.6/S	M6	+ nuts and split washers
Bolts for high load	32	MS 4.6/S	M8	+ nuts and split washers
Screws	80	Zinc-Plated Teks	6 g	
Plinth beams	4	MS grade 300+	310UC96.8	10 PL stiffeners
		Neoprene	2 layers × 3.0 mm	For damping
		MS 4.6/S	6/M16	Hand tightened with extension bar
		Hilti chemical anchors	6/HIS-N M16x170 ferrules	Anchored to RC substrate by Hilti HIT-HY 150 Max

Notes MS mild steel, *PL* plate, *FL* flat bar, *RHS* rectangular hollow section, *SHS* square hollow section, *UB* universal beam, *UC* universal column, *RC* reinforced concrete

Some proposed damage scenarios include the following: progressive loss of cross-sectional area on a deck beam or truss member to mimic corrosion; buckling of a compression truss member; progressive loosening of bolts at connections; growing crack in a cleat or gusset plate; fatigue failure of a weld; complete failure of a bracing member; impact damage; differential settlement at supports; and altered stiffness at supports to mimic bearing seizure or failure.

The data from the various damage scenarios will be analyzed using a number of different DD methods. Modal flexibility and modal strain energy methods show some promise for application to steel truss structures such as this one [31, 33].

17.5 Conclusion

This paper has presented a review of structural complexity in the SHM literature that identifies a gap between the structural complexity of numerical models and laboratory structures compared with the structural complexity of operational structures, particularly steel truss bridges. The QUT Benchmark Structure was presented as the most complex purpose-built laboratory structure of its kind, designed for research into DD and SHM. The proposed test plan was also outlined in brief.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established, and supported under the Australian Government's Cooperative Research Centres Programme. The primary author is a postgraduate student, studying at Queensland University of Technology, Brisbane. The primary author wishes to acknowledge the Australian Research Council for providing a living allowance scholarship, and the Cooperative Research Centre for Integrated Engineering Asset Management, Queensland Department of Transport and Main Roads and Brisbane City Council for providing top-up scholarships. Taringa Steel Pty. Ltd. is also acknowledged for their high-quality craftsmanship in the fabrication of the bridge model.

References

1. Moon FL, Aktan AE (2006) Impacts of epistemic (bias) uncertainty on structural identification of constructed (civil) systems. Shock Vib Digest 38(5):399–420
2. Ciloglu SK (2006) The impact of uncertainty in operational modal analysis for structural identification of constructed systems. Dissertation, Drexel University
3. Aktan AE, Ciloglu SK, Grimmelsman KA et al (2005) Opportunities and challenges in health monitoring of constructed systems by modal analysis. In: Proceedings of the international conference on experimental vibration analysis for civil engineering structures, Bordeaux, France, 26–28 Oct 2005
4. Tee KF, Koh CG, Quek ST (2009) Numerical and experimental studies of a substructural identification strategy. Struct Health Monit 8(5):397–410
5. Catbas FN (2004) Meeting of the bridge health monitoring committee of international association of bridge maintenance and safety. Meeting Minutes, Kyoto, Japan, 21 Oct 2004
6. Cowled CJL, Thambiratnam DP, Chan THT, Tan ACC (2012) Structural complexity in structural health monitoring: preliminary experimental modal testing and analysis. Paper presented at the 7th world congress on engineering asset management, Daejeon, Korea, 8–10 Oct 2012
7. Yan W-J, Huang T-L, Ren W-X (2010) Damage detection method based on element modal strain energy sensitivity. Adv Struct Eng 13(6):1075–1088
8. Mehrjoo M, Khaji N, Moharrami H, Bahreininejad A (2008) Damage detection of truss bridge joints using artificial neural networks. Expert Syst Appl 35(3):1122–1131
9. Yan G, Duan Z, Ou J, De Stefano A (2010) Structural damage detection using residual forces based on wavelet transform. Mech Syst Signal Process 24(1):224–239
10. Rama Mohan Rao A, Lakshmi K, Venkatachalam D (2012) Damage diagnostic technique for structural health monitoring using POD and self adaptive differential evolution algorithm. Comput Struct 106–107:228–244

11. Samali B, Li J, Choi FC, Crews K (2010) Application of the damage index method for plate-like structures to timber bridges. Struct Control Health Monit 17(8):849–871
12. Papatheou E, Manson G, Barthorpe RJ, Worden K (2010) The use of pseudo-faults for novelty detection in SHM. J Sound Vib 329(12):2349–2366
13. Quek ST, Tran VA, Hou XY, Duan WH (2009) Structural damage detection using enhanced damage locating vector method with limited wireless sensors. J Sound Vib 328(4–5):411–427
14. Weber B, Paultre P (2010) Damage identification in a truss tower by regularized model updating. J Struct Eng-ASCE 136(3):307–316
15. Gao Y, Spencer BF Jr, Bernal D (2004) Experimental verification of the damage locating vector method. In: Chang F-K, Yum CB, Spencer BF Jr (eds) Proceedings of the first international workshop on advanced smart materials and smart structures technology, Honolulu, Hawaii
16. Meruane V, Heylen W (2012) Structural damage assessment with antiresonances versus mode shapes using parallel genetic algorithms. Struct Control Health Monit 11(3):345–357
17. Dackermann U, Li J, Samali B (2010) Dynamic-based damage identification using neural network ensembles and damage index method. Adv Struct Eng 13(6):1001–1016
18. Catbas FN, Caicedo JM, Dyke SJ (2006) Development of a benchmark problem for bridge health monitoring. In: da Sousa Cruz PJ, Frangopol DM, Canhoto Neves LC (eds) Advances in bridge maintenance, safety management, and life-cycle performance: proceedings of the third international conference on bridge maintenance, safety and management, Porto, Portugal, 16–19 July 2006
19. Johnson EA, Lam HF, Katafygiotis LS, Beck JL (2004) Phase I IASC-ASCE structural health monitoring benchmark problem using simulated data. J Eng Mech-ASCE 130(1):3–15
20. Caicedo JM (2003) Structural health monitoring of flexible civil structures. Dissertation, Washington University
21. Thambiratnam D (1995) Vibration analysis of Storey Bridge. Aust Civ Eng 37(2):91–97
22. James G, Mayes R, Carne T et al (1995) Health monitoring of operational structures: initial results. In: Proceedings of the 36th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference, New Orleans, Louisiana, 10–12 Apr 1995
23. Krämer C, De Smet CAM, De Roeck G (1999) Z24 bridge damage detection tests. In: Proceedings of the 17th international modal analysis conference, vol 3727, Kissimmee, Florida, Feb 1999. SPIE, pp 1023–1029
24. Maeck J, De Roeck G (2003) Description of Z24 benchmark. Mech Syst Signal Process 17(1):127–131
25. Heywood R, Roberts W, Taylor R, Andersen R (2000) Fitness-for-purpose evaluation of bridges using health monitoring technology. Transport Res Rec 1696:193–201
26. Senthilvasan J, Thambiratnam DP, Brameld GH (2002) Dynamic response of a curved bridge under moving truck load. Eng Struct 24(10):1283–1293
27. Catbas FN, Grimmelsman KA, Aktan AE (2000) Structural identification of the Commodore Barry Bridge. In: Aktan AE, Gosselin SR (eds) Proceedings of SPIE: nondestructive evaluation of highways, utilities and pipelines IV, vol 3995, Newport Beach, California, Mar 2000. SPIE, pp 84–97
28. Agrawal A, Tan P, Nagarajaiah S, Zhang J (2009) Benchmark structural control problem for a seismically excited highway bridge—part I: phase I problem definition. Struct Control Health Monit 16(5):509–529
29. Chan THT, Wong KY, Li ZX, Ni Y-Q (2011) Structural health monitoring for long span bridges: Hong Kong experience and continuing onto Australia. In: Chan THT, Thambiratnam DP (eds) Structural health monitoring in Australia. Nova Publishers, Hauppage, New York, pp 1–32
30. Wang L, Chan THT, Thambiratnam DP et al (2012) Correlation-based damage detection for complicated truss bridges using multi-layer genetic algorithm. Adv Struct Eng 15(5):693–706

31. Wang L (2012) Innovative damage assessment of steel truss bridges using modal strain energy correlation. Dissertation, Queensland University of Technology
32. Standards Association Australia (1998) AS4100-1998: steel structures. Standards Association Australia, Sydney
33. Shih HW, Thambiratnam DP, Chan THT (2009) Vibration based structural damage detection in flexural members using multi-criteria approach. J Sound Vib 323(3–5):645–661

Chapter 18
Structural Complexity in Structural Health Monitoring: Preliminary Experimental Modal Testing and Analysis

Craig J. L. Cowled, David P. Thambiratnam, Tommy H. T. Chan and Andy C. C. Tan

Abstract Many researchers in the field of civil structural health monitoring have developed and tested their methods on simple to moderately complex laboratory structures such as beams, plates, frames, and trusses. Field work has also been conducted by many researchers and practitioners on more complex operating bridges. Most laboratory structures do not adequately replicate the complexity of truss bridges. This paper presents some preliminary results of experimental modal testing and analysis of the bridge model presented in the companion paper, using the peak picking method, and compares these results with those of a simple numerical model of the structure. Three dominant modes of vibration were experimentally identified under 15 Hz. The mode shapes and order of the modes matched those of the numerical model; however, the frequencies did not match.

18.1 Introduction

This paper presents the preliminary modal testing of the structurally complex QUT Benchmark Structure (see Fig. 18.1), which was designed and constructed in response to an identified need to conduct structural health monitoring research on more complex structures, as argued in the companion paper [1].

The numerical model of the structure is described. The aim of the experiment is presented. The equipment is listed, and the test method is detailed. Data are analyzed, and results are presented and discussed. The paper concludes with a discussion of the results that ties in with the aim of the experiment.

C. J. L. Cowled (✉) · D. P. Thambiratnam · T. H. T. Chan · A. C. C. Tan
Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: craig.cowled@qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_18,

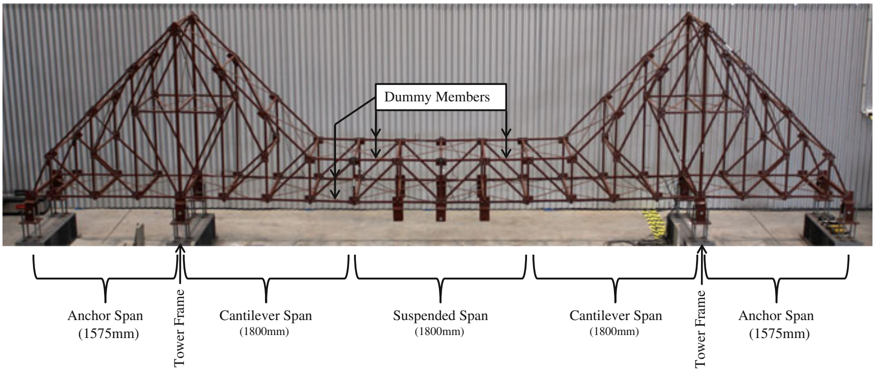


Fig. 18.1 The QUT Benchmark structure

18.2 Preliminary Numerical Model

Two numerical models have been developed for this structure. The first model was developed in Microstran [2] for the purpose of determining the design actions. A second model of the structure has been developed for model updating and damage detection [3]; however, the simple Microstran model is the focus of this paper.

The superstructure was modeled in Microstran as a space frame using the material specifications detailed by Cowled et al. [1]; however, the substructure was not modeled for this preliminary numerical model. The boundary conditions at the base of the tower frames were treated as simple pins, whereas the boundary conditions at the anchor tie down points were treated as rollers. All cross-bracing members were treated as tension only members with pinned joints. Likewise, struts and deck beams were given pinned joints. The fully welded 'H'-shaped tower frames were modeled as members with fully fixed joints. The joints of the chords and webs of the two main trusses were given semi-rigid connections (for in-plane bending), with a rotational stiffness of 10 kNm/rad (value determined by trial and error). The dummy chord members (shown in Fig. 18.1) were given a translational spring of zero stiffness at one end to mimic the slotted-hole structural detail.

The results of the dynamic analysis of the numerical model are presented in Table 18.2 along with those of the preliminary experimental modal analysis.

18.3 Preliminary Test

18.3.1 Aim

The aim of this experiment is to identify, from impact hammer tests, the frequencies, damping, and mode shapes of the first few modes of vibration and compare these results with those from the numerical model of the structure.

18.3.2 Equipment

A list of the equipment used in this experiment is shown in Table 18.1. The data acquisition software used in this experiment was LabVIEW Signal Express 2011 v.5.0.0 [4]. The software used for analysis was MATLAB R2012a [5].

18.3.3 Method

Experimental modal analysis with impact as the excitation force [6] has been adopted as the test method.

Eighteen of the sensors were fixed to the structure. These sensors were located at (or near) twelve node points on the structure with four sensors aligned in the global x direction (longitudinal), eight sensors aligned in the global y direction (vertical), and six sensors aligned in the global z direction (lateral). One sensor was positioned on the bottom flange of one of the 310UC plinth beams to monitor ambient vibrations transmitted to the structure from the concrete slab. Data from the reference sensor were qualitatively scrutinized prior to saving the results of a test. Tests with too much ambient vibration from the reference sensor were rejected and retested.

In order to capture sufficient spatial resolution in the modal model, 65 node points, out of a possible 100, were selected as excitation locations on the structure. A total of 102 impact hammer tests were conducted, representing 102 degrees of freedom (DOFs), of which 20 were aligned in the longitudinal direction, 46 were aligned in the vertical direction, and 36 were aligned in the lateral direction.

In order to capture sufficient resolution in the impact signal, which typically had a duration of 10 ms (based on a few test runs), a sampling frequency of 5.12 kHz was chosen.

Prior to conducting the tests, all machinery in the building, such as the air compressor, water cooler, and fridge, was turned off in an attempt to reduce the influence of ambient vibrations on the test data. In addition to the DOF of the impact, the date, time, air temperature, and relative humidity were recorded.

Each of the 102 tests was repeated at least twelve times. Prior to saving test data, the signals were examined for evidence of double impacts, poorly executed impacts, overshocked sensors, and 'out of range' accelerometer readings. Individual repeat tests which failed any one of these criteria were rejected and, if necessary, repeated. The minimum number of successful repeat tests in an impact test was seven, and the maximum number was sixteen.

The raw data from LabVIEW Signal Express [4] were exported to .txt files for transfer to MATLAB [5], where the results were analyzed.

Table 18.1 Experimental equipment

Item description	Qty	Specifications/comments
PCB Piezotronics ICB 086C03 impulse force hammer	1	Range ±2.2 kN, sensitivity 2.20 mV/N
Coaxial cable with BNC plugs	1	15 m
PCB Piezotronics 393B05 single-axis ceramic shear accelerometer	15	Range ±0.5 g, sensitivity ~10 V/g, Mass ≈ 120 g incl. MS base and magnet
MS flat bar	15	25 × 25 × 4.0
N42 rare earth disk magnets	15	20 dia. × 4.0
Coaxial cable with 10–32 sockets and BNC plugs	15	15 m
PCB Piezotronics 393B12 single-axis ceramic shear accelerometer	4	Range ±0.5 g, sensitivity ~10 V/g, Mass ≈ 340 g incl. MS base and magnet
MS flat bar	4	40 × 40 × 6.0
N42 rare earth ring magnets	4	25 dia. × 8.0
Coaxial cable with MILC-5015 sockets and BNC plugs	4	15 m
National instruments compact data acquisition system NI cDAQ-9172	1	Houses up to 8 modules
National instruments module NI 9234	5	Accepts up to 4 channels

Notes MS mild steel

18.3.4 Analysis and Results

The raw input data (impact hammer) were trimmed to 20 s time windows (102,400 data points) and truncated to reduce the influence of noise (see Fig. 18.2).

The raw response data were also trimmed to 20 s and de-trended. Minor adjustments were made to the data to account for sensors that were out of plumb and/or offset from a node point in order to create a 'virtual' signal at that node point.

Figure 18.3 shows ten repeated 'virtual' signals collected from one sensor during an impact test. The signals are well correlated, demonstrating the repeatability of the test method, which is one of the three key assumptions of modal analysis [6].

The auto spectra of the truncated impact force signal, $\boldsymbol{S_{ff}}(\boldsymbol{\omega})$, and normalized response signals, $\boldsymbol{S_{aa}}(\boldsymbol{\omega})$, were calculated using the *pwelch* function in MATLAB [5]. The cross-spectra, $\boldsymbol{S_{af}}(\boldsymbol{\omega})$ and $\boldsymbol{S_{fa}}(\boldsymbol{\omega})$, were calculated using the *cpsd* function in MATLAB [5]. The frequency response function (FRF) estimates, $\boldsymbol{H}_1(\boldsymbol{\omega})$ and $\boldsymbol{H}_2(\boldsymbol{\omega})$, were calculated using Eqs. 18.1 and 18.2 below:

$$\boldsymbol{H}_1(\boldsymbol{\omega}) = \frac{\boldsymbol{S_{fa}}(\boldsymbol{\omega})}{\boldsymbol{S_{ff}}(\boldsymbol{\omega})} \quad (18.1)$$

$$\boldsymbol{H}_2(\boldsymbol{\omega}) = \frac{\boldsymbol{S_{aa}}(\boldsymbol{\omega})}{\boldsymbol{S_{af}}(\boldsymbol{\omega})} \quad (18.2)$$

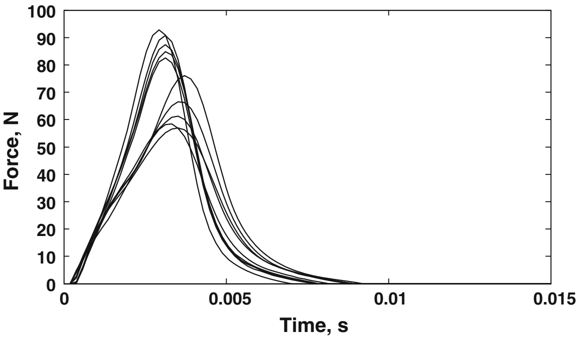


Fig. 18.2 Ten repeat truncated impact signals for one impact test

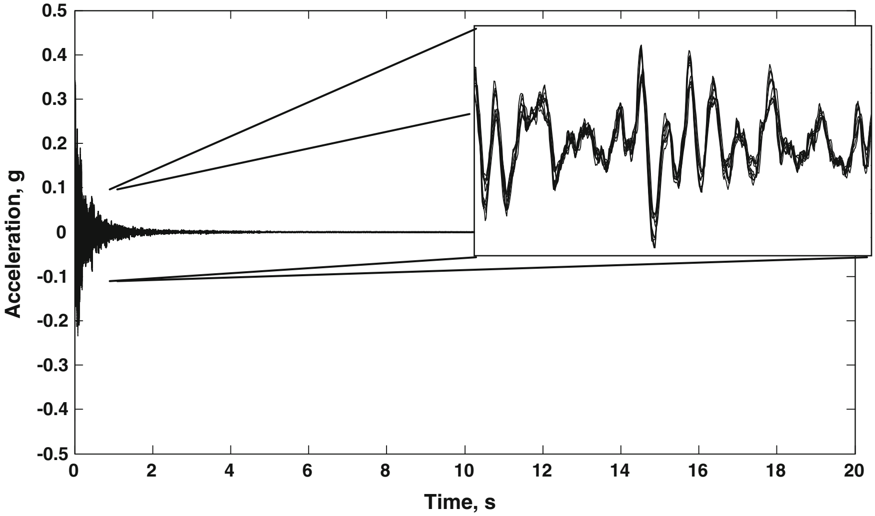


Fig. 18.3 Ten repeat acceleration signals from one sensor for one impact test (zoomed insert shows 0.1 s of data)

The coherence functions, γ^2, were also calculated using Eq. 18.3 below:

$$\gamma^2(\boldsymbol{\omega}) = \frac{\boldsymbol{H}_1(\boldsymbol{\omega})}{\boldsymbol{H}_2(\boldsymbol{\omega})} \tag{18.3}$$

Figure 18.4 overleaf shows the log modulus (power, dB) plot of ten repeat $\boldsymbol{H}_2$ FRF estimates representing one impact test with one input location and one output location. These estimates are very well correlated, indicating a high degree of linearity in the structure, which is the most important of the three key assumptions for modal analysis [6].

The mean value of the repeat FRF estimates from each impact test was calculated and passed into a 300 × 300 FRF matrix. Each of the rows of the FRF matrix represents an output signal at a particular translational DOF and each of the columns represents an input signal at a particular translational DOF. For example, $\boldsymbol{H}_{jk}$ would denote the FRF for the $\boldsymbol{j}$th row and the $\boldsymbol{k}$th column. The diagonal FRFs of this matrix represent the driving point FRFs (i.e., where $\boldsymbol{j} = \boldsymbol{k}$). The remaining FRFs each have a reciprocal FRF (i.e., $\boldsymbol{H}_{jk}$ and $\boldsymbol{H}_{kj}$). Figure 18.5 on the next page shows a reciprocal pair of averaged $\boldsymbol{H}_1$ FRF estimates (DOF 145 and DOF 298). The two FRFs are well correlated for most of the resonance peaks which shows that the key assumption of reciprocity [6] holds true for this structure.

One of the simplest methods of identifying modes is to use a simple mode indicator function (MIF). The simplest MIF is found by summing all the absolute values of the FRFs in the FRF matrix. The premise of this method is that global resonance peaks will become more prominent in the MIF, whereas noise and local resonances will become less prominent. A variation of this MIF was adopted for this study whereby FRFs were grouped into three categories and averaged, not summed. The three categories are namely those where the inputs and outputs were both aligned in the: (1) global x direction (longitudinal), (2) global y direction (vertical), and (3) global z direction (lateral). Not only does this MIF highlight global resonance peaks, it also provides more insight into the mode shapes (i.e., whether they might be lateral, vertical, longitudinal, or a mixture). Figure 18.6 shows the directional MIF for this structure, representing the averages of 584 FRFs out of a possible 1,836. It can be observed that the first three modes are dominant, well separated, and have more power in the global z direction (lateral).

The frequencies and damping values of the first three modes were determined by applying the peak picking and half-power methods [6] to each FRF individually. Not all FRFs were useful in estimating these values, particularly at DOFs where the mode shape of interest was small in magnitude. Subsequently, the results from some FRFs were rejected. Results are presented in Table 18.2.

The experimental results for natural frequency showed very little variation, particularly for modes 1 and 2. This provides a great deal of confidence in the accuracy of frequency results.

There was more variation in the results for damping; however, it should be noted that other methods of modal analysis have been applied to the same structure with "*almost no agreement between the identified damping ratios from the two techniques*" of enhanced frequency-domain decomposition (EFDD) and data-driven stochastic subspace identification (SSI-DATA) [3] (p. 155). This suggests that it may not be possible to obtain accurate results for damping. Notwithstanding the lack of accuracy in damping results, it can be observed that modes 1 and 2 are lightly damped (approximately 0.7 %) and mode 3 is moderately damped (approximately 1.5 %).

The dynamic analysis of the numerical model produced mode shapes similar to those obtained experimentally, and in the same order as the experimentally obtained modes, however, the natural frequencies of the numerical model are significantly higher than those obtained experimentally.

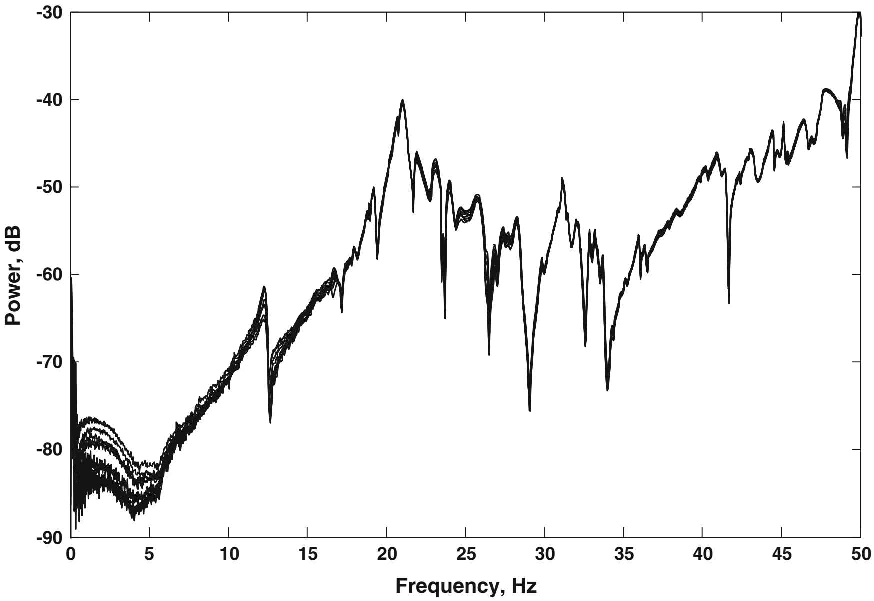


Fig. 18.4 Ten repeat $\boldsymbol{H}_2$ estimates from one sensor for one impact test

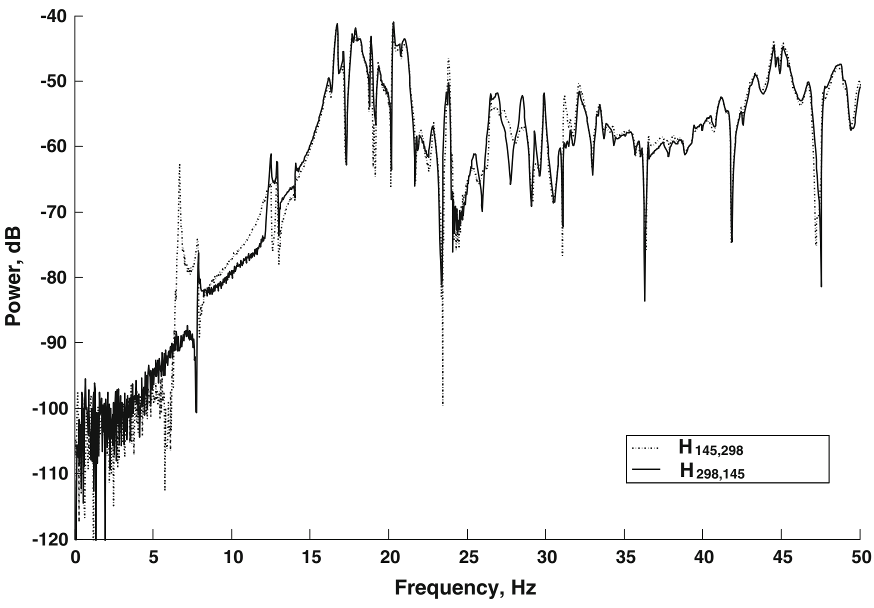


Fig. 18.5 Reciprocal pair of averaged $\boldsymbol{H}_1$ estimates for DOF 145 and DOF 298

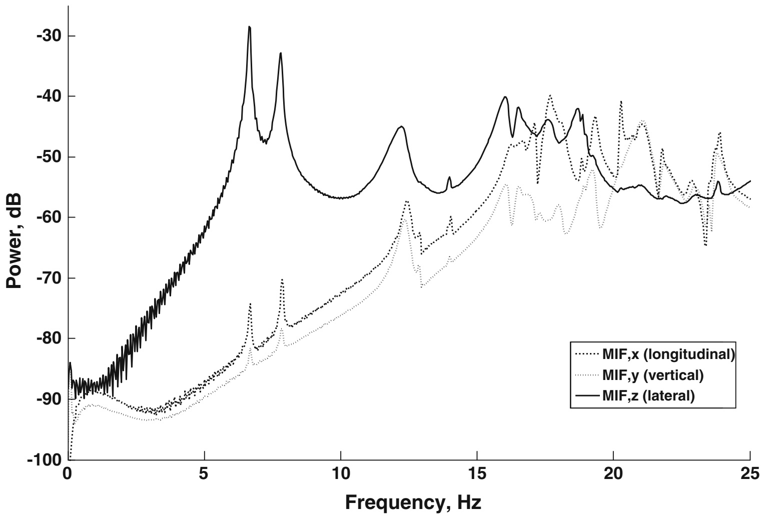


Fig. 18.6 Directional MIF

Table 18.2 Natural frequencies and damping of the first three modes

Mode no.	Frequency, Hz			Damping, ζ		Description
	Experimental		Numerical model			
	Mean	Std		Mean (%)	Std (%)	
1	6.672	0.027	9.99	0.743	0.063	Lateral sway 1
2	7.834	0.034	22.09	0.740	0.077	Lateral sway 2
3	12.325	0.137	33.18	1.548	0.455	Lateral bending 1

The mode shapes were obtained by using the quadrature picking method outlined in [7]. Each FRF within the FRF matrix was individually examined in the vicinity of resonance for modes 1, 2, and 3 and the peak value of the imaginary component of the FRF was selected as the mode shape value. With a focus on the 18 partially completed rows of the FRF matrix, mode shape values were interpolated for blank columns using a simple geometric method. The completed rows were then normalized and examined for consistency. Rows that were located near node points of a particular mode shape tended to show a lower consistency and were removed from the analysis. The remaining rows of normalized mode shape values were averaged to produce a smoother mode shape. Figures 18.7, 18.8 and 18.9 show the normalized and averaged mode shapes for modes 1, 2, and 3.

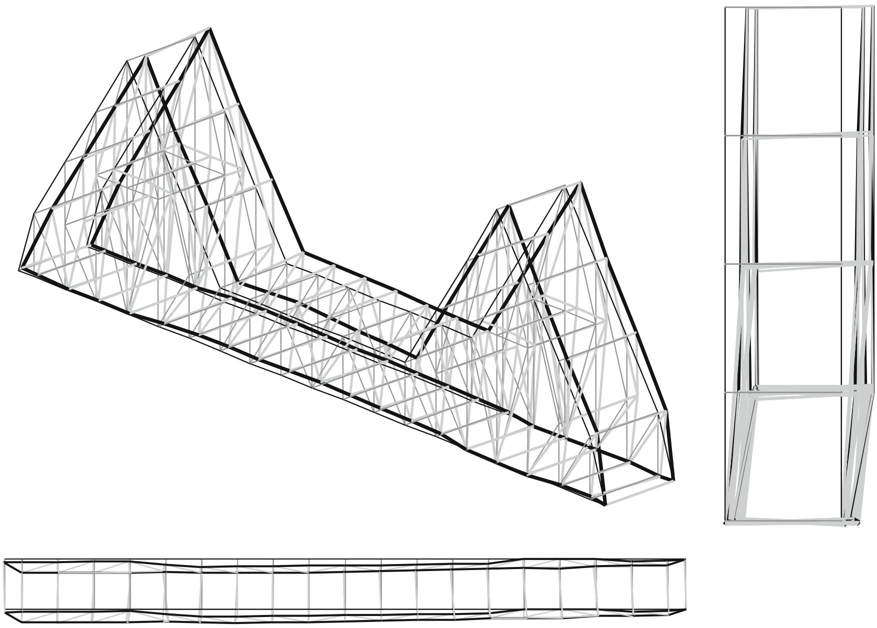


Fig. 18.7 Mode 1—Lateral sway 1, 6.67 Hz, 0.74 % damping

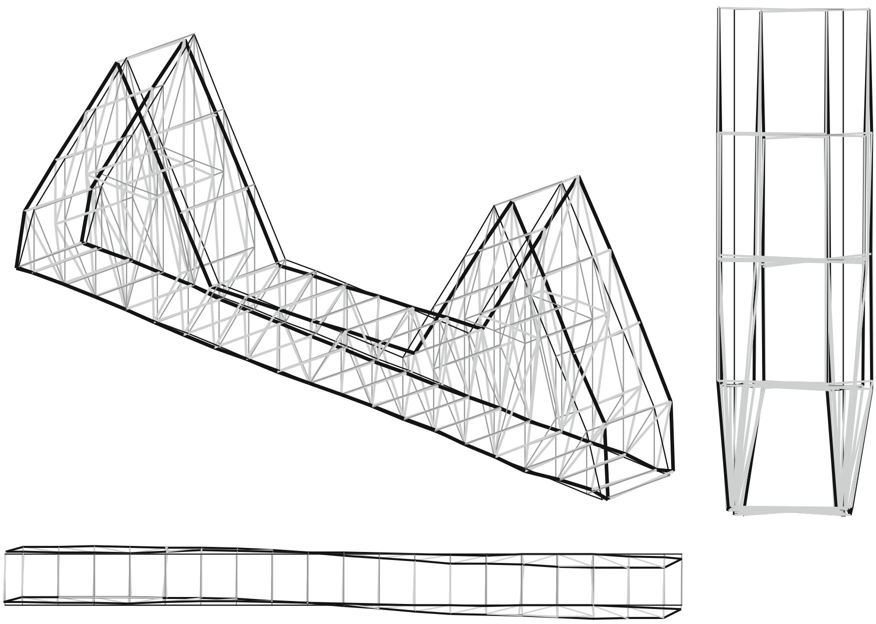


Fig. 18.8 Mode 2—Lateral sway 2, 7.83 Hz, 0.74 % damping

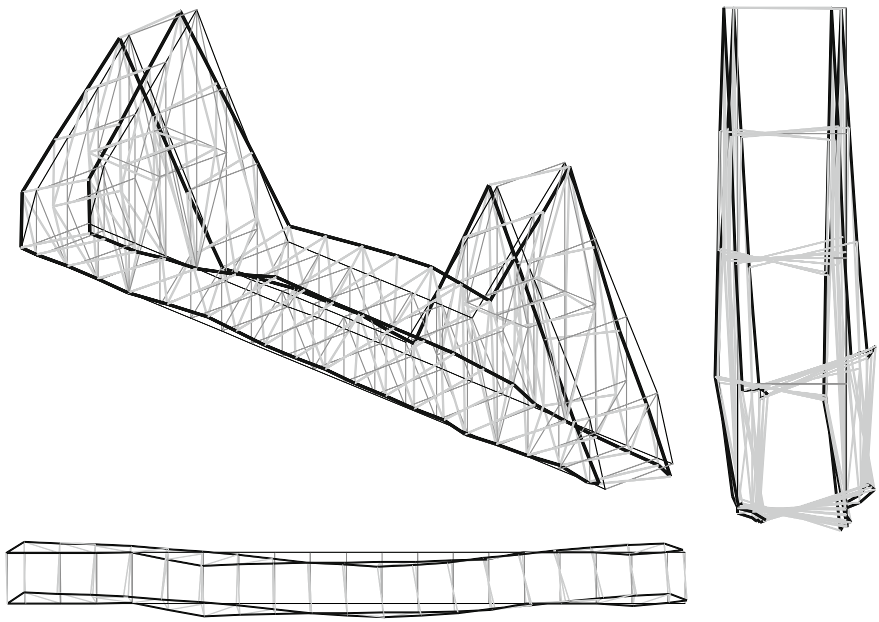


Fig. 18.9 Mode 3—Lateral bending 1, 12.33 Hz, 1.55 % damping

18.3.5 Discussion

The experimentally obtained frequencies did not match the results of the dynamic analysis of the numerical model; however, the order and shapes of the first three modes did match those of the numerical model. It should be noted that the boundary conditions of the numerical model were simplified as pins and rollers for the purpose of determining the design actions of the structure. These boundary conditions do not adequately represent the actual boundary conditions and consequently make the numerical model a lot stiffer than the actual structure. The numerical model will be updated as future work.

18.4 Conclusion

This paper has presented the preliminary modal testing and analysis of the bridge model which was described in the companion paper [1]. The aim of the experiment was presented. The equipment was listed, and the test method was detailed. The data were analyzed, and results were presented and discussed.

This experiment aimed to identify modal information of the first few modes of vibration, using impact hammer tests, and compare these results with the numerical model of the structure. Modal information was obtained for the first three modes of vibration.

The mode shapes correlated well with those of the numerical model. The order of the modes obtained experimentally and numerically were the same; however, there was a significant difference between the frequencies of the experimentally obtained modes and those from the numerical model, which indicates that the numerical model needs refining. Experimentally obtained damping results were somewhat variable.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centres Programme. The primary author is a postgraduate student, studying at Queensland University of Technology, Brisbane. The primary author wishes to acknowledge the Australian Research Council for providing a living allowance scholarship, and the Cooperative Research Centre for Integrated Engineering Asset Management, Queensland Department of Transport and Main Roads and Brisbane City Council for providing top-up scholarships. Taringa Steel Pty. Ltd. are also acknowledged for their high quality craftsmanship in the fabrication of the bridge model.

References

1. Cowled CJL, Thambiratnam DP, Chan THT, Tan ACC (2012) Structural complexity in structural health monitoring: design of laboratory model and test plan. Paper presented at the 7th world congress on engineering asset management, Daejeon, Korea, 8–10 Oct 2012
2. Engineering Systems Pty Ltd (2010) Microstran v.8.11.100930 (software). Engineering Systems, Turramurra
3. Wang L (2012) Innovative damage assessment of steel truss bridges using modal strain energy correlation. Dissertation, Queensland University of Technology
4. National Instruments Corporation (2011) LabVIEW Signal Express 2011 v.5.0.0 (software). National Instruments Corporation, Austin, Texas
5. The Mathworks Inc (2012) MatLab R2012a (7.14.0.739) (software). The Mathworks Inc, Natick, Massachusetts
6. He J, Fu Z-F (2001) Modal analysis. Butterworth–Heinemann, Oxford
7. Agilent Technologies Inc (1997) The fundamentals of modal testing (Application Note 243-3). Agilent Technologies, Santa Clara, California

Chapter 19
Stakeholder Engagement and Asset Management: A Case Study of the Gold Coast Airport, Queensland

Subas P. Dhakal, Muhammad N. Mahmood, Anna Wiewora, Kerry Brown and Robyn Keast

Abstract One of the fastest growing industries—aviation—faces serious and compounding challenges in maintaining healthy relationships with community stakeholders. One area in aviation-creating community conflict is noise pollution. However, current understandings of the factors that affect noise annoyance of the community are poorly conceptualized. More importantly, the way community needs and expectations could be incorporated in airport governance has been inadequately framed to address the issue of aircraft noise. This paper proposes the utility of adopting an integrated strategic asset management (ISAM) framework [1] to explore the dynamic nature of relationships between and airport and its surrounding area. The case of the Gold Coast Airport (OOL) operator and community stakeholders is used. This paper begins with an overview of the ISAM framework in the context of airport governance and sustainable development—as a way to find a balance between economic opportunities and societal concerns through stakeholder engagement. Next, an exploratory case study is adopted as a method to explore the noise-related complaints, complainants, and possible causes. Following this, the paper reviews three approaches to community stakeholder engagement in Australia,

S. P. Dhakal (✉) · K. Brown
Curtin Business School, Kent Street, Bentley, WA 6102, Australia
e-mail: subas.dhakal@curtin.edu.au

R. Keast
Southern Cross Drive, Bilinga, QLD 4225, Australia
e-mail: robyn.keast@scu.edu.au

K. Brown
e-mail: kerry.brown@curtin.edu.au

M. N. Mahmood
University of Southern Queensland, West Street, Toowoomba, QLD 4350, Australia
e-mail: nateque.mahmood@usq.edu.au

A. Wiewora
QUT Business School, 2 George St, Brisbane, QLD 4000, Australia
e-mail: a.wiewiora@qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_19,

Japan, and UK and discusses their implications in the context of OOL. The paper concludes with a contention that airport governance is likely to be much more effective with the adoption of ISAM framework than without it.

19.1 Introduction

Aviation is one the fastest growing industries in the world. The aviation industry can be broadly categorized into two sectors: airports and airlines [2]. This paper focusses on the airport as an infrastructure asset, particularly in relation to its operation. Aviation is an industry of national strategic importance to Australia [3]. The significance of airports as essential infrastructures for overcoming the tyranny of distance and fostering sustainable development is nowhere more evident than in the vast Australian continent. It is often argued that the changes in the governance structures—from state-owned assets to fully privatized entity—of Australian airports since 1998 have encouraged the operators to be fixated on maximization of the profits [4]. This lopsided emphasis on economic growth is argued to be eventually unsustainable because of the actual and potential adverse environmental, economic, and social impacts, such as noise pollution and loss of biodiversity, loss of property or land value, and disruption to lifestyle and community activities and functioning. If airports are to be considered vehicles of sustainable development, operators must find a way to maintain healthy relationships with community stakeholders and address societal concerns such as those relating to noise pollution associated with aircraft movements. However, current understandings of the factors that affect noise annoyance of the community and, more importantly, the way community needs and expectations should be incorporated in airport governance are inadequate. It is in this context, this paper examines the utility of an integrated strategic asset management (ISAM) framework [1], developed in conjunction with asset management industry associations, to examine the dynamic relationships between Gold Coast Airport (OOL) operator and community stakeholders.

This paper begins with an introductory overview of ISAM framework and relates this framework to airport governance and sustainable development. Next, an exploratory case study is adopted as a method to explore the noise-related complaints at OOL, complainants, and possible causes. This paper then reviews three facets of community stakeholder engagement in Australia, Japan, and UK and discusses their implications for OOL. The paper concludes with a contention that the adoption of ISAM framework for OOL operation can improve airport governance.

19.2 Integrated Strategic Asset Management Framework and Airport Governance

Assets can be either tangible, e.g., airport infrastructure or intangible, e.g., network knowledge that has a certain value or utility over the period of its life cycle. Optimum management of assets is a desired objective of airport operation. According to the Australian Asset Management Collaborative Group (AAMCoG), asset management is the process of organizing, planning, designing and controlling the acquisition, care, refurbishment, and disposal of infrastructure to support the delivery of services [1]. Recent approaches to asset management advocates the 'life cycle' view of an asset as a systematic and structured process that allows greater improvements in long-term performance, safety, and productivity. The entire life cycle of an asset can be multifaceted and involve several stages, e.g., acquisition, operation, maintenance, and disposal. Asset management is therefore a complex task mainly because the asset being managed may have a series of owners during various stages of its life cycle with different objectives, planning horizons, problems, stakeholders, and values [5]. Consequently, unilateral focus on technological aspects of asset and its management has gradually transformed to recognize the significance of human and social factors in the governance of airports [6]. The ISAM framework [1] is based on the following five principles:

1. Assets exist to support service delivery. Therefore, non-asset solutions should be considered
2. Agencies should manage assets consistent with whole-of-government policy frameworks and take into account whole-of-life costing, future service demands, and balance between capital expenditure and maintenance requirements
3. Asset management should be integrated with agency strategic and corporate planning
4. Asset management decisions should holistically consider sustainability outcomes: environmental, social, economic, and governance
5. Governance arrangements should clearly establish responsibility for functional performance of, and accountability for, the asset and service delivery (p. 5).

These principles are particularly useful for shaping airport governance mechanisms in order to internalize the needs and expectations of community stakeholders regarding noise annoyance.

The term governance captures a shift from the traditional hierarchical structure toward a horizontal decision-making process in which formal and informal relationships among the private sector, government representatives, and community stakeholders are valued [7]. The premise behind airport governance is that external actors, e.g., community stakeholders exhibit a range of interests and influence that needs to be addressed during airport operation. Although there is no unanimous definition of what constitutes a genuine community stakeholder, an individual or an organization with a stake or an interest in various stages of asset life cycle can be considered one. For the purpose of this paper, community stakeholders represent organizations with a stake—direct or indirect and beneficial or otherwise—in the way airport is governed. The theory of stakeholder engagement embraces the idea of corporate social responsibility [8] and assumes that airport operators have obligations to a broader society than just their shareholders. In other words, airport governance is said to be better, when operators invest in strategic relationships with community stakeholders rather than acting unilaterally. Several case studies in Australia and elsewhere have highlighted the fact that engaging with communities is vital for public image of the airports that have increasingly positioned themselves as the drivers of sustainable development [9–11]. Community stakeholder engagement is therefore central to the idea of airport governance for sustainable development—the notion which advocates community involvement as necessary to ensure not only economic prosperity but also environmental and social well-being [12]. It is in this context, this paper explores the dynamics of stakeholder engagement and asset management around the issue of noise pollution at the OOL.

19.3 Methodology

Gold Coast Airport (OOL) was chosen as a subject of case study for the paper because of two of the following reasons. Firstly, OOL (a) is one of the fastest growing Australian airports in terms of average annual growth of passenger movements [13] and (b) is expected to be one of the primary hubs for the visitors of 2018 Commonwealth Games to be held at Gold Coast [14]. Secondly, gauging by recent coverage in the local media, community stakeholders, (a) seem rather unimpressed by the future expansion of the airport and (b) have serious reservations about the ways airport operators are interested in addressing the issue of noise annoyance. An exploratory case study approach was adopted in order to investigate the nature of community stakeholders–OOL relationships using multiple sources of information. Case studies are particularly useful in exploring and comprehending diverse perspectives within the community because the method is open to the use of theory or conceptual categories that guide the research and analysis of data [15]. In order to triangulate the findings of the case study, this paper makes use of:

- Informal conversational interviews in which the researcher relies on the interaction with the interviewees to guide the structure [16], e.g., with key community representatives during the Airport Noise Abatement Consultative Committee (ANACC) meeting.
- Content analysis as an intellectual process of categorizing textual data into clusters of conceptual categories in order to identify consistent patterns between themes [17], e.g., local media coverage and publicly available minutes of ANACC meetings.

- Document analysis as a way to focus on conduits of meaningful communication of messages between the writer and reader [18], e.g., systematic analysis of current OOL master plan.

19.4 Findings

The Gold Coast is the sixth largest Australian city with a population of about half a million people. The city attracts more than 10 million tourists who collectively spend nearly $5 billion dollars annually [19], making it one of the most popular tourism destinations in Australia. In this regard, the OOL—located in southeast Queensland (QLD) with some portion of the runway within northern New South Wales (NSW)—is an economically significant infrastructure for the region. The airport was built in 1930s as an emergency landing ground for aircrafts flying between Sydney and Brisbane on the airmail services. The existing terminal building was completed in the 1980s. As a result of the privatization policy in the late 1990s, Queensland Airport Limited (QAL) purchased the 'Coolangatta Airport' in 1998 and renamed it OOL as known today under the management of OOL Private Limited [20]. OOL is Australia's fifth busiest international airport and the fastest growing one in terms of annual growth of passenger movements. The total number of passenger movements has nearly tripled from 1.9 million in 1998/1999 to 5.5 million in 2010/2011 since the changes in governance structure [21]. Because of growing interests of several airlines based in Asia and the Pacific, e.g., China Southern Airlines, Scoot (a subsidiary of Singapore Airlines) to establish direct connection between various Asian cities and the city of Gold Coast, the recently approved master plan predicts that OOL will service more than 16 million passengers by the year 2031/2032. In order to cope with this predicted increase in passenger and associated aircraft movements, an ambitious new construction plans to extend the runways and to improve the terminal facilities have been proposed in the 2011 master plan [19]. This scenario of extensive growth has alarmed community stakeholders in the region already frustrated with existing level and frequency of noise pollution associated with the aircraft movements (Fig. 19.1).

19.4.1 Community Stakeholder Engagement

OOL has embraced community stakeholder engagement as a part of the legislative requirement since the change of ownership in the late 1990s. There are two different forums, the *Airport Noise Abatement Consultative Committee* (ANACC) was established in 1999 and the *Community Aviation Consultation Group* (CACG) was established in 2011 [20]. The content analyses of 36 publicly available minutes of ANACC meetings between 1999 and 2012 suggest that there are a total of

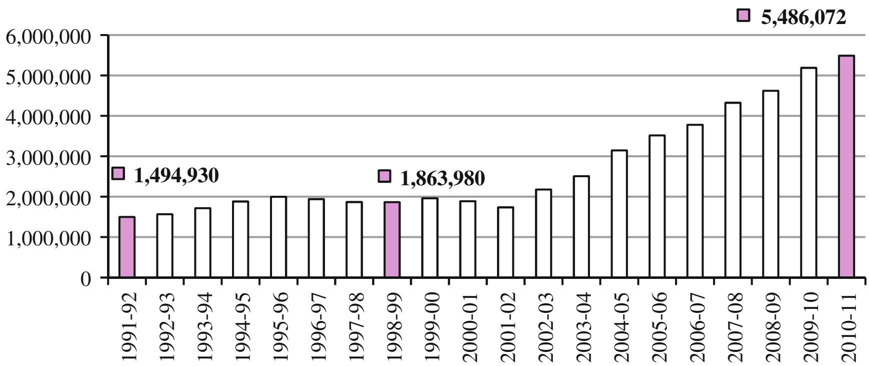


Fig. 19.1 Twenty years trend of passenger movements in Gold Coast Airport (*Source* [21])

18 active community stakeholders—10 in north of OOL (QLD) and 8 in south of OOL (NSW). On the one hand, ANACC has evolved to become a forum for particularly shaping noise abatement procedures over the years. For instance, minutes of the December 8, 2011 ANACC meeting indicates that stakeholders are generally appreciative of OOL and Airservices Australia—government-owned corporation responsible for ensuring services of aviation industry are safe and secure—efforts to work with the Department of Infrastructure and Transport in order to engage with the community [22]. On the other hand, recently formulated CAGG aims to be more holistic in its scope and proposes itself as a medium for broader issues related to airport development including that of the noise annoyance. For example, during the CACG meeting held on April 3, 2012 (attended by the lead author), the Airservices Australia representative willingly followed up on technical information related to the noise level and airport expansion requested by the community during February 12 meeting [22]. In this regard, the purpose CACG at the moment appears to be ensuring community views that are effectively heard by the airport as well as to inform community about broader activities of the airport operation. The CACG membership is open to residents affected by airport operations, local authorities, airport users, and other interested parties, and the CACG meetings are used to exchange information on issues relating to airport operations and their impacts. Apart from direct community stakeholder engagement through ANACC and CACG, OOL also sponsors various community programs ranging from extending financial support to the local Wildlife Sanctuary Animal Hospital to tourism-related campaigns totaling to $380,000 per annum [20].

19.4.2 Complaints, Complainants, and Possible Causes

A total of 38,813 complaints were received by the Airservices Australia from 350 community stakeholders in 2011 [23]. Analyses of the OOL documents and local media coverage indicated that complaints related to noise annoyance at OOL

Table 19.1 Noise complaints, complainants, and aircraft movements in various airports

Airport	State	Complaints (2011 in '000)	Number of complainants	Aircraft movements (2011 in '000)
Brisbane	QLD	6.59	322	168.34
Cairns	QLD	0.11	29	42.61
Gold Coast	QLD	38.83	350	37.37
Sydney	NSW	28.778	1,236	290.501

Source [23, 24]

were the highest not only in QLD but also in Australia. For instance, of the three international airports that operate in QLD, OOL received the highest number of complaints, even higher than the busiest airport—Sydney (Table 19.1). The number of complaints received by OOL in 2011 surpassed the actual number of aircraft movements at the airport. A local newspaper recently reported [24] that although the number of flights over the northern and southern areas of the airport was more or less the same, there was a concerted campaign from community stakeholders in NSW (south of the airport) to *make the noise about noise* by lodging thousands of complaints.

Extremely high number of noise-annoyance-related complaints received by OOL can be attributed to two possible causes: (a) Not in My Back Yard (NIMBY) syndrome and (b) North versus South Divide (NSD) reality. Firstly, NIMBY syndrome generally refers to localized resistance to often external development initiatives such as that of airport based on environmental grounds, e.g., noise pollution [25]. While this particular syndrome has been reported by media as being problematic in the context of opposition to airport-related development in Brisbane, and the proposed second airport in Sydney, it has also led to cooperation in case of Canberra Airport and its community stakeholders [8]. The syndrome and its potential association with the unusually high number of complaints are certainly a subject worthy of further research. Secondly, NSD reality is about differences between communities in northern and southern suburbs of OOL. Informal conversations during a recent CACG meeting, the president representing one of the northern community stakeholders indicated that the people living north of the airport understand that the airport is nearby and the associated noise is part of it. The president further asserted that the airport has been in the same location for nearly 80 years, long before people in the south even built houses. On the other hand, southern community representatives were adamant that they are carrying more than their fair share of noise during takeoffs (higher level of noise exposure) on top of southern suburbs because aircrafts mostly land (lower level of noise exposure) through the northern suburbs. An in-depth investigation of the north–south divide and its association with socioeconomic variables is equally worthy of further investigation.

Table 19.2 Comparison of approaches to stakeholder engagement in various airports

Airport	Authors: issue	Main findings	Implications for OOL
Birmingham, UK	Whitfield (2003): noise annoyance	Airport operators need to realize that unlike high level of noise exposure, low exposure affects different communities differently	Address socioeconomic differences between communities in shaping annoyance mitigation initiatives
Canberra, Australia	May and Hill (2006): noise ramifications of airport expansion	Airport operators need to be aware of stakeholder polarization—an alliance between local developers and community groups versus powerful vested interests seeking to manipulate community perception	Adopt a decision-making process to take the relationships within and between various community stakeholders into account
Narita, Japan	Yamada (2004): opposition to airport construction	Airport operators need to utilize deliberative based forum, e.g., regional symposium on airport issues, round table conference in order to gradually reduce community antagonism	Consider organizing a flagship event in which government, community, academics, and OOL can participate and exchange ideas or express concerns

19.5 Discussion

Noise annoyance has been a significant issue for the governance of airports around the world, and it is clear from the findings above that OOL is no different. In accordance with the Air Navigation (Coolangatta Airport Curfew) Regulations of 1999, OOL has adhered to curfew for aircraft movements between 11 pm and 6 am since December 22, 1999, in order to minimize the noise annoyance [20]. However, curfew hours have only partially addressed this thorny issue at the most. An attempt is made here to review and summarize three significant approaches to community stakeholder engagement in Australia, Japan, and UK [8, 26, 27].

Table 19.2 depicts and examines community concerns in these three airports and points out a possible way forward in the context of OOL. In order to improve airport governance through meaningful engagement with community stakeholders, OOL needs to consider: (a) socioeconomic differences within and between stakeholders, (b) significance of relationships among various stakeholders of OOL or social capital—the idea that social connections or relationships matter [12], and (c) a flagship event that can potentially bring variety of stakeholders in one forum.

19.6 Conclusion

This paper began with an introductory overview of ISAM framework, placing it in the context of airport governance and sustainable development—as a way to balance economic prosperity and societal concerns through stakeholder engagement. A case study method was adopted in order to explore noise-related complaints at OOL, complainants, and possible causes of complaint. The findings indicated that complaints related to noise annoyance at OOL were the highest not only in Queensland but also in Australia. The possible associations between extremely high numbers of complaints were made with: (a) Not in My Back Yard (NIMBY) syndrome and (b) North versus South Divide (NSD) reality. The syndrome and reality were also identified as two important areas for future investigations. Then, the paper concisely reviewed three approaches of community stakeholder engagement at airports in Australia, Japan, and UK and discussed their implications in the context of OOL governance. For airports, stakeholder engagement needs to evolve from: (a) a compliance enforcement and (b) 'already have the license to operate' approach toward a model in which the role of community stakeholders is embedded in the decision-making process. As airports play a crucial role in the sustainable development of the regions that they are located in [20], the utility of ISAM framework to manage community needs and expectations as a way to enhance airport governance for sustainable development is significant [28, 29]. It is in this context, the paper contends that airport governance is likely to be much more effective with the adoption of ISAM framework as it offers the start of a guideline to bring together the different and perhaps competing arenas in airport infrastructure management.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management (CIEAM), established and supported under the Australian Government's Cooperative Research Centres Programme, in collaboration with the Collaborative Research Network (CRN) programme within Southern Cross University. The authors gratefully acknowledge the financial support provided by the CRC and CRN.

References

1. AAMCoG (2012) Guide to integrated strategic asset management. Australian Asset Management Collaborative Group (AAMCoG) and Cooperative Research Centre for Integrated Engineering Asset Management (CIEAM), Brisbane, QLD
2. Kaszewski AL, Sheate WR (2004) Enhancing the sustainability of airport developments. Sustain Dev 12:183–199
3. DITRDLG (2009) National aviation white paper. Department of Infrastructure, Transport, Regional Development, and Local Government (DITRDLG), Canberra, ACT
4. Forsyth P (2002) Privatization and regulation of Australian and New Zealand airports. J Air Transp Manage 8:19–28
5. IAM (2011) Asset management: an anatomy. The Institute of Asset Management (IAM), Bristol, UK

6. Keast RL, Baker DC, Brown KA (2008) Balancing infrastructure for the airport metropolis. In: Proceedings of international conference on infrastructure systems: building networks for a brighter future. Institute of Electrical and Electronics Engineers (IEEE), Rotterdam, The Netherlands, pp 246–250
7. Dhakal SP (2011) Can environmental governance benefit from ICT-social capital nexus in civil society? Triple C: Cogn Comm Cooperation 9(2):551–565
8. Greenwood M (2007) Stakeholder engagement: beyond the myth of corporate responsibility. J Bus Ethics 74:315–327
9. May M, Hill SB (2006) Questioning airport expansion: a case study of Canberra International Airport. J Transp Geogr 14:437–450
10. Thomas C, Lever M (2003) Aircraft noise, community relations and stakeholder involvement. In: Upham P, Maughan J, Raper D, Thomas C (eds) Towards sustainable aviation. Earthscan, London, UK, pp 97–112
11. Freestone R, Baker D (2010) Challenges in land use planning around Australian airports. J Air Transp Manage 16:264–271
12. Dhakal SP (2012) Regional sustainable development and the viability of environmental community organisations: why inter-organisational social capital matters? Third Sect Rev [Special Issue in Environment] 18(1):7–27
13. Airservices Australia (2011) Gold Coast noise information pack, Dec 2011. Airservices Australia, Canberra, ACT
14. Willoughby S (2012) Gold Coast Airport plan takes off. Gold Coast Bull, 16 May 2012. http://www.goldcoast.com.au/article/2012/05/16/416101_tweed-byron-news.html
15. Blatter JK (2008) Case study. In: Given LM (ed) The Sage encyclopedia of qualitative research methods, vol 2. Sage Publications Inc, Thousand Oaks, CA, pp 68–71
16. Turner DW (2010) Qualitative interview design: a practical guide for novice investigators. Qual Rep 15(3):754–760
17. Julien H (2008) Content analysis. In: Given LM (ed) The Sage encyclopedia of qualitative research methods, vol 2. Sage Publications Inc, Thousand Oaks, CA, pp 120–122
18. Prior LF (2008) Document analysis. In: Given LM (ed) The Sage encyclopedia of qualitative research methods, vol 2. Sage Publications Inc, Thousand Oaks, pp 230–232
19. Gold Coast Tourism (2010) Gold Coast hits 10 million tourists mark for first time since 2008, Dec 2010 Media Release. http://www.visitgoldcoast.com/media/media-releases/
20. Gold Coast Airport Private Limited (2011) Master plan 2011. Gold Coast Airport Pty Ltd (GCAPL), Bilinga, QLD
21. BITRE (2012) Airport traffic data. Bureau of Infrastructure, Transport and Regional Economics [BITRE]. Australian Government, Canberra, ACT. http://www.bitre.gov.au/publications/ongoing/airport_traffic_data.aspx
22. Gold Coast Airport Private Limited (2011) Aircraft Noise Abatement Consultative Committee (ANACC) MINUTES (8 Dec 2011). Gold Coast Airport Pty Ltd (GCAPL), Bilinga, QLD
23. Menon U, Simpkins M, Keenan E (2012) Gold Coast CACG, 22 Feb 2012. http://goldcoastairport.com.au/community/community-aviation-consultation-committee
24. Ricks B (2012) Coast airport noise worse than Sydney. Gold Coast Bull, 23 Feb 2012. http://www.goldcoast.com.au/article/2012/02/23/393485_tweed-byron-news.html
25. Suau-Sanchez P, Pallares-Barbera M, Paul V (2011) Incorporating annoyance in airport environmental policy: noise, societal response and community participation. J Transp Geogr 19:275–284
26. Yamada I (2004) Partnership for noise management between the airport and communities in Japan. In: Proceedings international commission for acoustics (ICA). ICA, Kyoto, Japan, pp 2047–2050
27. Whitfield A (2003) Assessment of noise annoyance in three district communities living in close proximity to a UK regional airport. Int J Environ Health Res 13(4):361–372

28. Boons FA, van Buuren A, Teisman G (2010) Governance of sustainability at airports: moving beyond the debate between growth and noise. Nat Res Forum 34:303–313
29. Koppenjan J, Mandell M, Keast R, Brown R (2009) Contexts, hybrids and network governance a comparison of three case-studies in infrastructure governance. In: Brandsen T, Holzer M (eds) The future of governance. Fifth Transatlantic Dialogue, Washington, DC, pp 301–325

Chapter 20
An Assessment of Benchmarking Practice Relating to Complex Engineering Assets: A Case Study from an Oil and Gas Production Company

Riad El-Wardani and Jayantha Prasanna Liyanage

Abstract Companies are continuously striving for world-class performance by use of various performance management tools and techniques. One such popular technique, proving effective in achieving best-in-class results, is benchmarking. A case study of an oil and gas production company, aiming to drive business performance based on best practice-based learning rather than compliance, shows how benchmarking can be effectively implemented to maintain continuous improvement. Benchmarking practice is assessed across a variety of different departments affecting operations, giving an insight into the diversity and complexities of implementing the practice yet paying attention on learning opportunities, while maintaining focus on the ultimate goal of achieving world-class performance.

20.1 Introduction

Many methods have been adopted and developed to boost business performance within oil and gas organizations. Some are widely accepted and used, while others are less popular. One of the most popular techniques is benchmarking, so much so that it features as a common job-description competency for production managers in Japan. Having spawned and improved over the past two centuries, it is now one of the most effective tools for continuous performance improvement and innovation. Even though it is believed to have been in practice since the 1700s, when shoemakers used it for relative pricing, it was not recognized as a methodical performance management tool until the 1970s when Xerox first documented and implemented the practice systematically, while competing with its Japanese

R. El-Wardani (✉) · J. P. Liyanage
Centre for Industrial Asset Management (CIAM), Rennebergstien 30, 4021 Stavanger, Norway
e-mail: riad.wardani@gmail.com

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_20,

counterparts. Robert Camp, considered to be the forefather of benchmarking, was part of Xerox's widely acknowledged benchmarking campaign in the late 1970s. It is evident from the literature that most modern benchmarking models are based on Camp's documented method. In some cases, it was considered the key factor in transforming fading businesses into global conglomerates taking over majority market share, as was the case for Xerox in the late 1970s. Nevertheless, the literature around the technique is mainly geared toward case studies and practices, leaving a gap in the comprehensive theoretical description of the practice, its characteristics, strengths, and limitations. The working definition of benchmarking used throughout this paper is based on Camp's [2] concept, stating that "benchmarking is the continuous process of measuring a firm's own performance against industry best practice, leading to superior performance through structured changes to operations."

As alluded to by the definition adapted, to truly capitalize on the gains attainable by benchmarking, a thorough understanding of one's existing business practices, processes, and performance levels is crucial, as well as a clear understanding and continued focus on the strategic objectives of benchmarking. Coupled with a strong and concrete implementation plan, this will give organizations a solid chance in achieving world-class performance. The central question of interest for an oil and gas production context is how should the benchmarking process be carried out to ensure true business implication in terms of learning, continuous improvements, and innovation? This requires a systematic approach to clarify several different aspects such as: What is the main objective of benchmarking for the organization—why benchmark? What are the different benchmarking models within different departments—how to benchmark, adapt common practices, or encourage diversity and tailored solutions? Ideally, such questions shall be raised at a higher level of the organization and a strategy set to drive benchmarking teams forward including a strong commitment to continuously improve and update the company's benchmarking practices.

Subsequently, a detailed study was conducted within a large oil and gas operator organization (denoted as *Company-A* hereafter) on their application of benchmarking. *Company-A* has adopted a major responsibility of "driving simplification and improvement initiatives" by relying on tools such as benchmarking. The overarching aim is to drive business performance according to best practice rather than on compliance basis. To date, the full potential of benchmarking has not been realized since the concept is not easy to define, let alone follow-up. A great deal of knowledge and practice remain hidden within the organization's systems that should be brought to light to drive performance, based on effective benchmarking. The presence of many benchmarking opinions and methodologies, both within and outside the organization, has cluttered the process and made it difficult to compare and rank facilities or projects, using a systematic benchmarking methodology.

As such, a thorough review of the chosen company's benchmarking processes within different departments was conducted to assess internal business practices and identify objectives of each department in relation to benchmarking.

Interestingly, varied performance management objectives and strategies were identified within different departments in the organization. These diverse practices are studied in detail, and conclusions are drawn to aid one of the Norway's largest oil and gas production companies in achieving their strategic benchmarking objectives. More alignment and common practices are recommended as well as improved communication and better understanding of benchmarking principles by staff at all levels of the company.

20.2 Scope and Objectives

Initially, the study examines the company's high-level performance management strategy and implementation process focusing on capturing elements such as the company's strategic means of driving performance, objectives of benchmarking, effective communication of strategic objectives, and finally, at a more detailed level, how do the different departments within the company actually benchmark. Here is where different benchmarking models within the company are analyzed and compared to enhance cross-departmental learning and highlight internal benchmarking best practice. In terms of the bigger project scope, this was the first piece of the puzzle where later external best practice was analyzed and an integrated benchmarking model presented which combines internal and external best practice.

20.3 Methodology

To fulfill the scope and objectives of this project, an in-depth analysis of the subject area was first conducted. Often, the challenge is to find the right questions to ask to get the desired information. Therefore, a thorough background study of the subject matter was necessary to identify the best way of both mining for information within the information management system and when approaching staff. A detailed academic literature review was kicked off relying on the resources of *The University of Stavanger* library services, online journals, and subject matter papers. Very few of the publications reviewed provided a comprehensive review of benchmarking practice; each was centralized around a specific outcome the respective author wanted to achieve. Additionally, many of the publications refer to the same base literature, which shows a consistency of the science over the years in spite of having witnessed continued development. Furthermore, in addition to literature, interviews and workshops with subject matter experts and colleagues enriched the background study, providing a valuable hands-on perspective.

In terms of the empirical component, it was imperative to review the company requirements in contrast to the empirical evidence obtained from those actually involved in performance management within the company. To get this breadth in opinions and variety of context, the first step was to research governing documents, functional requirements, and work processes related to performance measurement and management. This was followed by interviews, workshops with 2–3 people, e-mail correspondence, and phone calls which turned out to be an especially valuable source for understanding the practical implementation of strategies and documented practices.

20.4 State of the Art

As mentioned, one of the most challenging aspects of effectively applying benchmarking is to first gain a common definition and understanding of what the practice entails and what its purpose is. As early as 1992, Spendolini [5] published "*The Benchmarking Book,*" where he already by that time uncovers 49 different definitions for the term "benchmarking."

The author does not wish to attempt and create yet another definition of the term "benchmarking," rather to consolidate the existing definitions and practices as far as practicable and capture the core elements that are important in understanding and implementing benchmarking. These core elements are identified as follows:

- Initially, define the company's strategic direction, objective, and focus.
- Gain a thorough understanding of one's own processes and practices, strengths as well as weaknesses and define areas for improvement.
- Use scaling factors, commonly called "complexity factors," to compare one's own performance internally and/or with other companies identified to be "best-in-class" in that specific practice.
- Learn about the underlying processes and practices leading to best-in-class performance.
- Adapt such processes and practices to one's own operating culture, strategy, and goals to achieve superior performance, i.e., attempt to be best-in-class.
- Continuously evaluate and repeat the process to maintain world-class performance.

Once the working definition and core elements outlining the process are understood, it is useful to go a level deeper into the types of benchmarking that are conducted. This will allow the analysis to be more specific, targeting those types of benchmarking that could help an oil and gas production company drive its business performance objectives. However, it was observed that there was no-to-little consensus on a scheme for classifying benchmarking practices. Fong et al. have documented this same fact in 1998 and so have Anand and Kodali in 2008 [1]. Lack of convergence is still evident in today's literature and practice. In 1998, Fong et al. [4] tabulated the different types of benchmarking in an attempt to

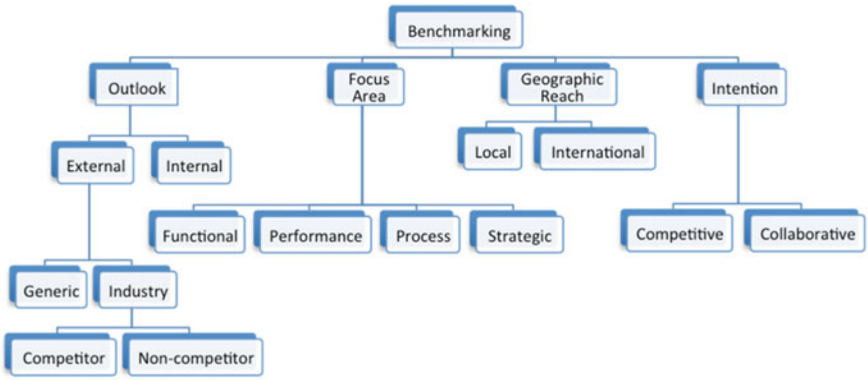


Fig. 20.1 Graphical representation of different types of benchmarking

Fig. 20.2 Decision flow diagram for selecting most appropriate benchmarking type

finally create some sort of a "standardized classification system." This has not succeeded, and although few authors have accepted this classification scheme and its associated definitions, others continue to develop their own. Graphically remodeled, the essence of Fong's classification is demonstrated in Fig. 20.1 and the author encourages adoption of such a system for consolidation and standardization purposes .

An associated decision flow diagram was also developed to assist organizations still new to the benchmarking process to be able to select and implement the most suitable type of benchmarking to cover their individual needs (Fig. 20.2).

Now that a thorough understanding is achieved, it is of value to commence the empirical analysis by interacting with those involved with benchmarking practice within the company as well as the target setters.

20.5 Empirical Study, Results, and Analysis

Noting that *Company-A* has had extensive experience and a long track record of adapting benchmarking as a performance-driving tool, it is only obvious that there is a lot of knowledge stored within the company's systems and held by its personnel. The key was to surface this knowledge, assess it, and try to establish what internal benchmarking best practice looks like. The commonly agreed upon definition within the company is as follows:

> Benchmarking is **measuring** and **comparing** products, services, processes and functions with the best, to **identify**, **understand** and **implement** better ways of conducting business as part of the company's effort for continuous improvement.

Company-A's objective, in terms of performance, is to be recognized as a top-quartile performer within their industry. The company's highest level publication and a fundamental element of the management system, describing the company's most important policies and requirements, considers benchmarking part and parcel of the operating model with respect to "strategy development and target-setting." Furthermore, the first principle of the company's integrated performance management process strongly complements benchmarking practice as shown in the text box below.

> *1st key principle of Performance Management Process: Performance is about performing better than those we compare ourselves with*

Having demonstrated *Company-A*'s strong commitment to continuous performance improvement and their drive for simplification and standardization through benchmarking, it is appropriate to review, in more detail, how different departments within the organization convert strategy into practice. Initially, the envisaged results were that alignment would be demonstrated between corporate and the underlying value chain and support processes, as well as across the different processes.[1] However, after further investigation, it became clear that each of the reviewed departments had a slightly different perspective when it came to benchmarking, significant enough to make it individual. Through interviews, workshops, and the internal management system, a review of five different processes was conducted, the essence of each of their benchmarking practices captured and compared. Having a mixture of value chain and support processes offered a broader perspective at two different levels of the organization. The reviewed processes were the following:

- Drilling (value chain process)
- Project development (value chain process)
- Supply chain management (support process)
- Heath, safety, and environment (support process)
- Operation and maintenance (value chain process).

[1] Value chain processes are the processes related to the company's core business whereas support processes are support functions for the value chain processes.

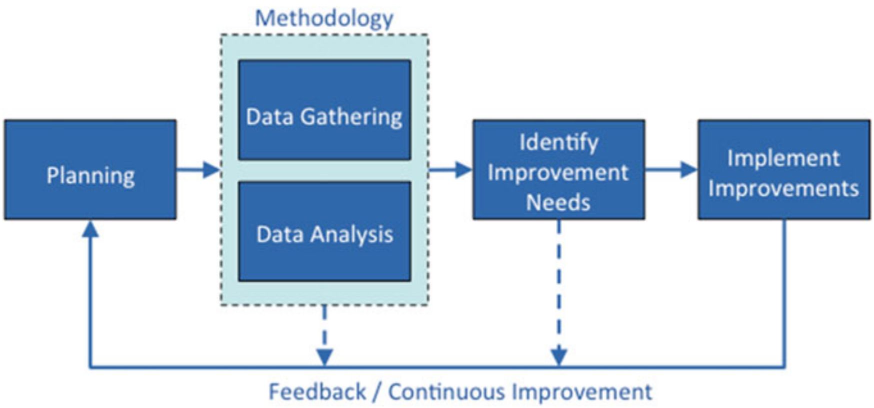


Fig. 20.3 Building blocks of departmental review

Figure 20.3 shows the building blocks of the departmental reviews. In the *Planning* block, objectives, goals, structure, and preparation of systematic benchmarking practice are included. *Methodology* includes *Data Gathering* and *Data Analysis*. Although these are both very important and separate phases, it was clear from the reviews that it is acceptable to combine them. Successful identification and balancing of *Improvement Needs* comes next, encompassing the approach of converting performance gaps into specific, concrete actions that will result in bridging the gap. Without *Implementation of Improvements,* the benchmarking effort would be a waste of resources; therefore, assessing how departments convert gaps, plans, and ambitions into positive outcomes is important to review. Finally, one of the generally neglected aspects of benchmarking, *Feedback and Continuous Improvement*, is not concerned with how each department conducts its business, but rather how they evaluate their benchmarking process and continuously update it to ensure that it adheres to best practice.

20.5.1 Departmental Models Summary

Table 20.1 summarizes the models adapted by each of the five business units based on three main phases: data gathering, data analysis (including reporting), and finally the implementation phase where improvement initiatives are applied as concrete actions.

Table 20.1 Departmental benchmarking methodology comparison summary

		Drilling	Project development	Supply chain management	Health, safety, and environment	Operations and maintenance
Objective	Statement	–	*Building a pathway to asset success*	*Become benchmark within procurement in 2012**	*Industry lead in HSE*	*Operations and Maintenance toward World Class*
	Empirical	Well planning purposes\achieve consistently; comfortable performance level	Comparison/ planning	Competitive focus on bridging the gap	Varied due to lack of clarity on objective from requestors; identify gap and improve performance	Aggressive focus toward best-in-class
Methodology	Data gathering	Internal benchmarking dashboard and external	Part of functional requirements but no separate process and external	External	Dedicated process	External, no separate process found
	Presentation	*proNova* and *Performance Management Process*	Project Performance Improvement Program (PPIP)	External reports and Performance Management Process	Synergy and *Performance Management Process*	Operational performance dashboard (OPD)
	Analysis	No guidelines found	No guidelines found	No guidelines found	No guidelines found	No guidelines found
Implementation		Management strategies and targets translated into *Performance Management Process*	–			

(continued)

Table 20.1 (continued)

	Drilling	Project development	Supply chain management	Health, safety, and environment	Operations and maintenance
	Internal benchmarking dashboard—localized, daily improvements based on short-term data from previous day	–	Dedicated campaign developing specific actions and improvement initiatives	Task teams made up of subject matter experts—leading advisors and chief engineer	–
Challenges	• Conservative culture	• Resource shortage	• Benchmarking culture: getting line organization to accept and take ownership of benchmarking results	• Too much focus on the numbers and not the process/enablers • Clearly defining objectives of benchmarking effort • Some data not reported (cultural barrier)	• Overall understanding of benchmarking process for all involved • Resource shortage • Cross-asset learning
		• Clarity on source of data		• Balancing improvement benefits	• Balancing improvement benefits
	• Results linked to incentives	• Gap between management and line organization • Results linked to incentives	• Clarity on source of data	• Resource shortage • Willingness to learn from others	• Gap between management and line organization • Systematic and structured implementation process

*was the objective between 2008 and 2012, now with strategy update could be different

20.6 Conclusions

In conclusion, it has been found that there is a gap in the literature in terms of accurate and comprehensive explication of benchmarking both as a concept and as a process. There is a need for literature that characterizes benchmarking versus other types of analysis. Further, a more thorough explanation of the process and the expectations behind the main headings is required in terms of assessing one's own performance and finding the enabling processes and practices that lead to best practices. Delving deeper beyond the numbers and understanding such enabling processes is the most critical part of benchmarking, and in the past, it has received very little attention.

In terms of the empirical study, numerous effectual benchmarking processes have been identified within the organization along with a myriad of expertise related to the process, its limitations, and challenges. Expertise, however, need to be more spread throughout the organization and not only be centralized within the corporate and managerial levels to ensure a completely successful implementation of benchmarking and its improvement potential. Training, awareness, and open communication should be enhanced to improve on this aspect, in addition to more accessible systems that communicate all decisions and intentions related to benchmarking.

By combining benchmarking best practice and company know-how, *Company-A* can readily transfer their business performance into world-class performance, achieving their goals and setting the targets for the industry. Refer to El-Wardani and Liyanage [3] for more details on the proposed development of existing benchmarking practices within the oil and gas production company to support them in achieving this world-class performance.

References

1. Anand G, Kodali R (2008) Benchmarking the benchmarking models. Benchmarking: Int J 15(3):257–291 (online Emerald Insight)
2. Camp RC (1989) Benchmarking: the search for industry best practices that lead to superior performance. Productivity Inc, Japan
3. El-Wardani R, Liyanage JP (2012) Creating a learning and improvement scenario based on distinctive internal benchmarking practices and cultures: an adaptive strategy based on experience from an oil and gas production company. The University of Stavanger, Stavanger, Norway
4. Fong SW, Cheng EWL, Ho DCK (1998) Benchmarking: a general reading for management practitioners. Manag Decis 36(6):407–418
5. Spendolini M (1992) The benchmarking book. American Management Association Communications (AMACOM), New York

Chapter 21
Creating a Learning and Improvement Scenario Based on Distinctive Internal Benchmarking Practices and Cultures: An Adaptive Strategy Based on Experience from an Oil and Gas Production Company

R. El-Wardani and J. P. Liyanage

Abstract Contrary to traditional management style, where new managers would introduce immediate drastic changes in an attempt to create short-term impact, more contemporary performance management philosophies rely on embracing and further developing existing internal practices and operating cultures toward world class. Building on the vast knowledge and experience available within the performance management systems of one of the world's largest oil and gas production companies, an Integrated Benchmarking model is developed combining internal and external best practices and continuous improvement tools.

21.1 Introduction

Companies are continuously striving to gain an edge on their competitors to increase market share and ultimately profits. Careful monitoring of key performance indicators, both leading (process) and lagging (outcome), in addition to a thorough understanding of one's own processes, procedures, and working culture, allows organizations to make minor adjustments on key parameters which significantly improve performance in the form of cost, schedule, and quality.

One of the most renowned and widely adapted continuous improvement tools has and continues to be benchmarking. The technique has been around since the eighteenth century and has enjoyed a range of publications ranging from theoretical fundamental type literature, to case studies and innovative adaptations. By

R. El-Wardani (✉) · J. P. Liyanage
Centre for Industrial Asset Management (CIAM), Rennebergstien 30,
4021 Stavanger, Norway
e-mail: riad.wardani@gmail.com

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_21,

adapting this continuous improvement technique, two overarching factors are key: (1) The practice has to be in fact continuous in nature and (2) benchmarking results should not be used to overhaul an organization's operating procedures and should not clash with existing company culture and existing practices. Both of these challenges along with several limitations and reluctances have traditionally plagued the practice; these will be investigated and addressed in the context of one of the world's largest oil and gas production companies, Company A.

Before delving into and resolving some of the benchmarking challenges and limitations faced by Company A, it is important to define the working definition of benchmarking used herein. It is based on Camp's [2] concept, stating that "benchmarking is the continuous process of measuring a firm's own performance against industry best practice, leading to superior performance through structured changes to operations."

Following, the results of a review of the company's benchmarking processes and procedures are used as input into a de-bottlenecking process where a holistic solution is sought to improve the effectiveness of benchmarking within the organization and set clear guidelines on how it should be carried out. The concept of "Integrated Benchmarking" is introduced as a solution and is considered to be a fairly new and underdeveloped hypothesis only documented in few previous publications, none of which empirical in nature. Therefore, this effort is thought to be a genuinely original contribution to the field of benchmarking within performance management.

21.2 Scope and Objectives

Using the results of a parallel study by El-Wardani and Liyanage [3] on the benchmarking processes, culture, and procedures of an oil and gas production company, main areas of improvement are identified in the attempt to take the organization's benchmarking efforts from good to world class. Once the main areas of improvement are identified along with the challenges and limitations inherent in benchmarking practice, a framework based on the existing procedures and operating processes is developed around the unique and effective approach of Integrated Benchmarking. Originally, the concept is developed based on performance management best practices and integrating different tools and techniques accounting for each others' shortfalls within a complex system. Finally, the framework is plugged into the company's benchmarking activities to review how it would be implemented and how it would enhance the outcome of benchmarking practice.

The objective is to evolve the existing benchmarking process, feeding it with internal experience and external best practice to enhance the effectiveness of the practice within the company.

21.3 Methodology

A thorough background study of the subject matter was necessary to identify the best approach of both digging up information within the company's Information Management System and when approaching staff. Furthermore, since the scope is to amalgamate company practices and best practices, a broad performance management literature review was required while maintaining a detailed focus to assess the applicability and value of integrating different techniques into the company's benchmarking profile. The base material for the report, where the company's benchmarking practices are reviewed in much closer detail, is captured in El-Wardani and Liyanage's [3] publication where interviews and workshops with subject matter experts and colleagues provided the necessary hands-on perspective of how the company actually carries out benchmarking.

In terms of the empirical component of this paper, it was imperative to review company requirements in contrast to the empirical evidence obtained from those actually involved in performance management within the company. To get this breadth in opinions and variety of context, the first step was to research governing documents, functional requirements, and work processes related to performance measurement and management. This was followed by interviews, workshops with 2–3 people, e-mail correspondence, and phone calls which turned out to be an especially valuable source for understanding the practical implementation of strategies and documented practices.

21.4 State of the Art

After investigating the several foundational elements of benchmarking such as the origin, definition, evolution, literature diversity, and types and different processes adapted over the years, it was necessary to take a slightly more generic viewpoint. This involved reviewing the ethical responsibilities and challenges/limitations associated with the practice that in turn leads to challenges and barriers while conducting the benchmarking activities. Following is a brief discussion of the main findings in each of the reviewed sections including the extracted best practice lessons learnt.

Firstly, touching on the ethical obligations that need to be considered when conducting benchmarking, there are some fine lines differentiating benchmarking from re-engineering or copycatting. While there is an abundance of literature about benchmarking code of conduct from third-party benchmarking facilitating organizations, there seems to be a lack of mention about ethics in case study/specific application-related literature. Evidently, the most popular "Code of Conduct" publication quoted is that produced by American Productivity and Quality Centre (APQC) upon which the International Benchmarking Clearinghouse and European Benchmarking codes of conduct are based.

The APQC Code of Conduct consists of eight principles; each has its own sub-bullets describing what the heading entails. For a more detailed description, refer to the full report [4]; however, the eight principles are as follows:

- Principle of Legality—no illegal dealings for mutual or individual gain and do not acquire trade secrets
- Principle of Exchange—be willing to provide the same type and level of data as that requested
- Principle of Confidentiality—respect confidentiality and do not share with external parties
- Principle of Use—use information gained only for the stated purposes
- Principle of Contact—respect corporate culture, agreed procedures, and follow correspondence protocol
- Principle of Preparation—be prepared prior to making contact
- Principle of Completion—follow through with benchmarking effort to satisfaction of all, as agreed
- Principle of Understanding and Action—understand expectations and preference of partners

The value of benchmarking should be clear by this stage: (1) It helps achieve world-class performance by setting realistic, achievable goals that others have in fact achieved and (2) it enables a company's resources to maintain focus and alignment with corporate strategy and vision. The technique, as popular as it seems to be, still faces some criticism. Detractors argue that benchmarking is merely a way of copying what others are excelling at and exposes a company's weaknesses and shortcomings by "hanging out their laundry." Opsahl's [5] publication titled "Does this benchmarking make me look inefficient?" discusses how the public sector in the USA seems to want to avoid external benchmarking in an effort to conceal inferior performance. This notion is evident even among the most professional and prestigious private firms. Williams et al. [6] have comprehensively documented different reasons why employees and companies are reluctant to benchmark. They found that these reluctances fit roughly into four categories, as follows:

- Relating to the soundness of benchmarking (i.e., comparing "apples to oranges")
- Concern about lack of resources (financial, time, expertise, data, etc.)
- Inertia: resistance to change (comfort, fear of the unknown, regulations, etc.)
- Concern regarding impact of implementing new practices (loss of influence, job changes, etc.)

Based on Williams et al. [6] study, best practices have been identified which help overcome benchmarking reluctance and encourage organizations to take it on as an overall continuous performance improvement tool. These best practices converge into four sequentially linked categories listed below and presented graphically in Fig. 21.1 as a puzzle, which if completed will result in more effective and efficient campaigns.

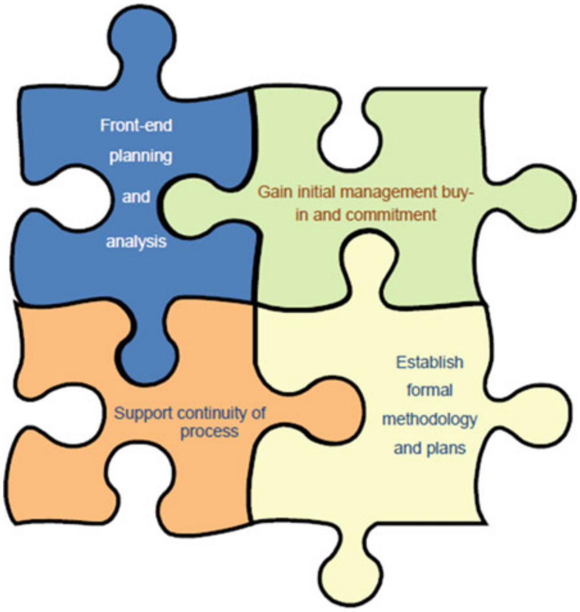


Fig. 21.1 Best practice to overcome benchmarking reluctance

1. Front-end planning and analysis
2. Gain initial management buy-in and commitment
3. Establish formal methodology and plans
4. Support continuity of benchmarking process

It is crucial to start the planning phase for the benchmarking activities early and identify the main players, objectives, gaps, and processes that will apply. Part of this phase includes thoroughly understanding one's own processes and way of operating and taking in the company culture to ensure that no clashes will arise at later stages. Once a clear picture of the company's internal processes and performance gaps is identified, it is imperative to gain management buy-in and support early on to help drive the benchmarking campaign and provide it with the attention and prioritization it requires. Formal methodologies and plans will clearly and concisely communicate the main intention behind the effort, how it will be carried out and what the data will be used for. Having a clear data gathering and processing procedure, for example, will ensure that this phase is streamlined and runs efficiently while gathering the same level and quality of data throughout the organization. Communicating the purpose of gathering the data and how it will be used for continuous improvement will detract staff involved to try to paint a better picture or avoid reporting "bad news." Maintaining an open and clear communication culture will reduce conflicts and staff contesting benchmarking results and redirect the focus on continuous improvement. Last but not least, and this is possibly one of the most overlooked aspects of benchmarking, is to ensure a continuous process that is repeatable and evolving including lessons learned from previous benchmarking campaigns. This helps the organization truly capitalize on their benchmarking efforts and achieve significant gains.

Finally, before presenting the Integrated Benchmarking model that was developed as part of this project, challenges and barriers are yet to be discussed and best practices for overcoming them. These are quite different to the critiques of benchmarking discussed above since they emerge once the benchmarking effort is underway rather than prior to starting. For a comprehensive listing of benchmarking barriers and challenges, refer to Williams et al. [6].

In no particular order, the first challenge relates to overcoming resistance by employees and/or management to benchmarking. Skeptics are commonly nearby when new tools and initiatives are introduced. Some of the reasons employees are skeptical relate to company culture and employee attitude. For those believing that benchmarking is "spying" on the competitors, it must be understood that by benchmarking, management cannot lower their guard in terms of creativity and innovation [7]. According to Walleck [8], "The purpose of benchmarking is to expose managers to new ways of doing things in order to spark creativity, not to create efficient copy cats!" Benchmarking provides a means to identify the gap between one's own performance and that of other companies and in some instances learn and adapt some of the best practices they use. It is still essential to be creative both when adopting practices and when improving on practices to gain superiority.

Other employees are reluctant to get involved and simply refuse to comply with new policies, whether it is due to stress factors, related to working outside of their comfort zones, or learning new skills [9, 10]. Others, who are more innovative and competitive, seem to lose interest and get discouraged especially if they see that the aim of benchmarking is mainly targeted toward matching competitor's performance and simply adopting common competitor practice [11]. In contrast, while some would rather adopt competitor practices, others are reluctant to accept such additions to their organization. The "not invented here" syndrome described by Amaral and Sousa [9] displays a special kind of reluctance where employees are not willing to incorporate any products or practices developed elsewhere. This is also not encouraged since in some instances, it is wise to adapt practices from other organizations by improving and developing them further. Arun Maua, Vice President at Arthur D. Little states that "you can't just impose a best practice. It has to be adapted to your own company's style."

In some cases, it is the organizational culture that induces this inherent resistance. Organizations considered to be "not a learning organization" seem to possess a culture where knowledge transfer, learning from past experience, problem solving, and experimentation are not common. This reflects upon the employees, who are not used to seek or share knowledge [9].

One of the most challenging phases of benchmarking is the implementation phase, which logically poses some of the most delicate challenges to deal with. In some instances, lack of proper implementation is a result of inadequate or insufficient employee skills when it comes to benchmarking implementation. A lack of understanding of the organization's products and services leads to frustration and stagnation in the process, which could be due to inadequate training or poor leadership [9]. Poor planning during the initial phase and failing to clearly identify the goals of the process, tasks, resources, and deadlines appear later down the track and adversely affect implementation. Another common challenge presented by Elmuti and Kathawala [10] as one of the most common implementation challenges is lack of involvement from employees within the line organizations during the benchmarking process. Commonly, the benchmarking process is carried out by a corporate performance management team or a dedicated group that does not include those who are actually responsible for delivering the results. During implementation, challenges arise due to the disconnect between organizational levels, leading to poor implementation and frustration that affect not only the conducted benchmarking effort but also any future benchmarking campaigns due to its bad reputation. Linked to this issue is the inadequate definition of the benchmarking scope. Since, in some instances, the line organization employees are not sufficiently engaged in the process, the topic becomes too broad or poorly articulated using management jargon that is too abstract to measure or implement in practice [9]. So, it is important to actively engage the employees during the process, since they will be the ones actually using the information and improving the process [12].

Regrettably, benchmarking is commonly carried out as a once-off exercise. This is in fact a poorer solution than not carrying out benchmarking at all. The concept of “ignorance is a bliss” may work well as long as the company is profitable. In cases where benchmarking is used as a last resort when companies are failing, it is then too late to benchmark. Conducting benchmarking as a one-time project results in a very limited comparison scope, poses the risk of misinterpreting or misusing the information, and does not allow the organization sufficient time to get acquainted with best practices in order to adopt and adapt them. Carrying out benchmarking as an ongoing process allows the company to develop its knowledge and expand its scope of companies studied: looking at competitors, non-competitors within the same industry, and other companies in different industries [10]. This is the exposure a company needs to advance to the top by applying benchmarking effectively.

21.5 From Good to World Class

Generally, when it comes to suggesting an improved model for performance measurement and management, it is important not only to consider what has worked well for other organization or think about ideal theoretical scenarios, but to incorporate company culture, operating philosophy, and existing practices into that model. The underlying proposition of using “Integrated Benchmarking” aims to set a standardized, systematic method of conducting tasks related to benchmarking and continuous improvement that is clear to everyone, from senior staff within corporate business units to the maintenance supervisor guiding the crew on the tools. Many diverse initiatives are already in place, which seem to work to varying

Table 21.1 Reference model element summary

Element	Purpose/Function	Organizational level
Benchmarking	Identify gap/compare practices	Corporate/Managerial
Balanced scorecard	Balancing improvements	Corporate + line Org
Force field analysis	Implementation—drivers and restraints	Corporate + line Org

degrees of success. Understanding and amalgamating those practices and coupling them with industry best practices would align the company's operating model with industry best practices and reduce the level of change management required. It would take the organization's benchmarking efforts "from good to world class."

So, how does this translate into tangible impacts to the benchmarking process? The elements of the model translate into concrete gains as presented by Table 21.1.

After careful review of the benchmarking model adapted by the different departments within Company A, the following are the main identified areas of improvement:

- Data gathering phase
- Defining improvement initiatives
- Implementing improvement initiatives
- Feedback (closing the benchmarking cycle)

One way to tackle these challenges is to introduce the unique and effective concept of "Integrated Benchmarking," where benchmarking is combined with other techniques to reap the full benefits of the practice and compensate for where it lacks. The graphical representation of the model is presented in Fig. 21.2. The circular theme in the representation is designed to give the impression of a never-ending, continuous process of improvement. Gears illustrate the intertwining and meshing of the different techniques, which must be managed carefully or else it will diminish the effectiveness and efficiency of the improvement process. Notice the opposite direction of the arrow on the Balanced Scorecard (BSC) gear. This demonstrates that using BSC gives a different twist to the performance improvement cycle by taking a step back and reviewing what the value is with respect to strategic priorities. Finally, the pathway from the Current Performance level point along the different techniques up to the Target Performance level is "dashed" to highlight that this is a progressive process that takes time and needs to be carefully planned, monitored, and implemented rather than being a quick fix.

Table 21.1 shows the different elements of the model, their purposes, and at what organizational level they should be applied. It is important to address the issues at the right level to be able to get the relevant effect.

According to the model, the process of striving from *Current* to *Target Performance* level is as follows:

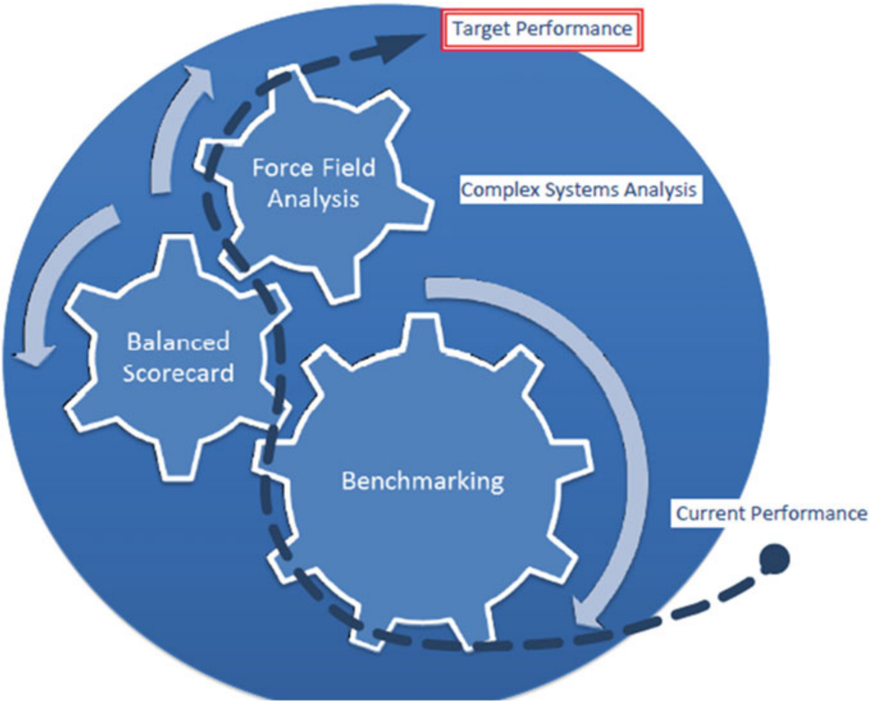


Fig. 21.2 Integrated Benchmarking model

1. Understanding one's own processes, evaluating performance, and identifying the gap to be bridged to reach world class, all as part of an internal or external benchmarking campaign
2. Putting a different spin on the analysis; utilize *BSC* to effectively trade-off between improvement initiatives to achieve the best results with the available resources
3. Communicating through dynamic *Force Field Analysis* the benefits, drivers, and restraints from senior management to the line organization and vice versa

All of these techniques interrelate within a complex system, where each relies on elements of the others in a dynamic and continuous fashion. Next is a detailed description of how using such an integrated approach may contribute to the areas of improvement identified previously.

Initially, the data gathering and analysis phase is identified as one of the main areas of improvement in terms of staff awareness and training. Staff on facilities, responsible for gathering data, is not kept in the loop as to the purpose of collecting the required data and how it will be used in the bigger picture. Consequently, a challenge emerges when data do not correlate at the analysis stage because every department or group participating in a project is reporting based on their individual

Fig. 21.3 Quality versus cost quadrant

interpretations and in the manner that best suits them. A high-level overview and awareness will enable them to provide data that are better aligned and easier to use by analysts and business improvement groups. By the same token, the concern that incentives are tied to benchmarking rankings may also result in inaccurate data reporting. Therefore, moving away from this reward method and communicating factual benefits of benchmarking toward continuous improvement is advised.

When it comes to defining improvement initiatives, currently through the external benchmarking activities, the company at a very high managerial level receives a summary presentation that includes the areas of potential significant improvement. Next, area managers gather their teams of operations managers to discuss improvement initiatives and implementation plans. High-level initiatives and targets are agreed; the operating team is to come up with detailed plans and initiatives to implement. An aspect of defining improvement initiatives that needs to be enhanced is balancing different elements to ensure that the best overall results are attained. A recurrent theme was identified through the review of benchmarking models showing that quality is generally very good within the company, while costs are unacceptably high. World-class performers are characterized by conducting their business efficiently (low cost) and with high quality (effectiveness) as highlighted in the top right quadrant of Fig. 21.3. To get to this optimum balance between several different measures, *Balanced Scorecard* (BSC) can be used. This is actually not a new development within the company, and the *Information Management System* is in fact based on this for the myriad of key performance indicators (KPIs) used within the company.

Developing the BSC concept, already used within the company, beyond balancing KPI performance to using it for prioritization and definition of the most beneficial improvement proposals to be carried out will ensure that resources are allocated to those proposals yielding the highest gains. Forward feeding the results of the gap analysis, in terms of areas of improvement and associated actions, into the BSC assessment, will result in a list of prioritized tasks to be carried out according to balanced gains. Once the list of improvement needs is defined and prioritized at a higher level in alignment with strategic objectives, it is wise to

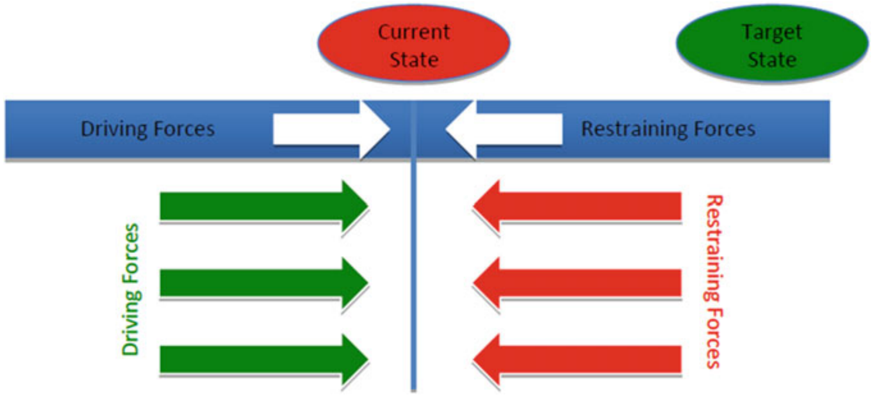


Fig. 21.4 Force field analysis [1]

include the line organization to flesh out the details. At this stage, it is very important to appreciate the limitations that line organizations are expected to maneuver around. Therefore, trying to decide on how to implement improvement plans without including them would be largely a theoretical exercise. This is one of the issues the organization is currently facing: disconnect between management and the line organization. Therefore, Force Field Analysis (FFA) has been recommended as a downstream process after BSC.

In the Integrated Benchmarking model, incorporating the line organizations into the goal setting and implementation phases will ensure more realistic and practical solution as well as provide ownership for them to improve their own performance. One of the most effective tools in converting plans, strategies, and targets into concrete, specific, and achievable actions, while at the same time communicating the challenges and drivers, is Force Field Analysis (FFA), as shown in Fig. 21.4. FFA offers a simple but clear way of communicating current performance state, target as well as barriers, challenges, and drivers. Driving forces in this case are motivations for why it would be good to achieve the target state (improve HSE performance, reduce cost, become best-in-class, etc.). More importantly though, are the required actions taken to reach the target or improved state. On the other hand, restraining forces are those elements or critiques describing what would hold the team back from achieving their goals.

Finally, the feedback loop currently only exists in the form of future benchmarking data collection, which gives a very late indication of whether the process was or was not successful. In some instances, individuals suggest modifications to the process. This is not a systematic or regular occurrence. By adapting the Integrated Benchmarking model, several feedback loops are indirectly included throughout the process. These are discussed in more detail in Sect. 21.6 about the road map.

21.6 Conclusion

In conclusion, this work based on the findings of El-Wardani and Liyanage [4] has demonstrated the importance of keeping ethical responsibility at the foreground of benchmarking efforts and incorporating it early into the planning phase. It also brings together years of benchmarking experience and presents the main lessons learnt and best practices that will drive benchmarking to be used as a continuous improvement tool while developing and evolving the practice according to company culture and needs.

Once these concepts are understood and aligned into the benchmarking procedures, limitations of the technique must be identified in accordance with the objectives, which will then introduce other techniques to fill the gap and ensure organizations get the most of their benchmarking campaigns.

Using an Integrated Benchmarking framework and incorporating BSC, FFA and spider web diagrams for presentation purposes will ensure achievable, tangible gains are realized assisting organizations to take their performance from good to world class.

References

1. Burnes B (2004) Managing change. Prentice Hall, Englewood Cliffs
2. Camp RC (1989) Benchmarking: the search for industry best practices that lead to superior performance. Productivity Inc, Japan
3. El-Wardani R, Liyanage JP (2012) An assessment of benchmarking practice relating to complex engineering assets: a case study from an oil and gas production company. The University of Stavanger, Stavanger
4. El-Wardani R, Liyanage JP (2012) Developing a reference-model for benchmarking: performance improvement in operations and maintenance. The University of Stavanger, Stavanger
5. Opsahl A (2006) Does this benchmarking make me look inefficient? http://www.govtech.com/e-government/Does-This-Benchmark-Make-Me-Look.html (April)
6. Williams J, Brown C, Springer A (2012) Overcoming benchmarking reluctance: a literature review. Benchmarking Int J 19(2):255–276
7. Boxwell RJ (1994) Benchmarking for competitive advantage. McGraw-Hill, New York
8. Walleck AS (1991) A backstage view of world-class performers. The McKinsey quarterly reprint series, pp 2–5
9. Amaral P, Sousa R (2009) Barriers to internal benchmarking initiatives: an empirical investigation. Benchmarking Int J 16(4):523–542 (online Emerald Insight)
10. Elmuti D, Kathawala Y (1997) An overview of benchmarking process: a tool for continuous improvement and competitive advantage. Benchmarking Qual Manage Technol 4(4):229–243
11. Shetty YK (1993) Aiming high: competitive benchmarking for superior performance. Long Range Plan 26(1):39–44
12. Omachonu VK, Ross JR (1994) Principles of total quality. St. Lucie Press, Delray Beach, pp 137–154

Chapter 22
Data Mining the Relationship Between Road Crash and Skid Resistance

Daniel Emerson, Richi Nayak and Justin Z. Weligamage

Abstract Road asset managers are seeking analysis of the whole road network to supplement statistical analyses of small subsets of homogeneous roadway. This study outlines the use of data mining capable of analyzing the wide range of situations found on the network, with a focus on the role of skid resistance in the cause of crashes. Results from the analyses show that on non-crash-prone roads with low crash rates, skid resistance contributes only in a minor way, whereas on high-crash roadways, skid resistance often contributes significantly in the calculation of the crash rate. The results provide evidence supporting a causal relationship between skid resistance and crashes and highlight the importance of the role of skid resistance in decision making in road asset management.

22.1 Introduction

Road safety is a major concern worldwide with road crashes costing countries between one and three percent of annual gross domestic product. The World Health Organization predicts road traffic crashes emerging as the 3rd leading cause of disease or injury burden [1]. Research is continually contributing to development of roadway design principles [2], and this study contributes to enhancing the knowledge of the role of road surface friction (skid resistance) in crashes.

The effect of skid resistance in braking is well understood [3], and skid resistance is routinely investigated in case studies of crash hot spots [4]. Historically, statistical, and data mining studies have been used for analysis of crashes on small-scale, homogeneous road subsets. However, there is an emerging demand for whole of road network analysis. The Australian road research organization, the ARRB Group, contends that use of risk assessment at the network level (rather

D. Emerson (✉) · R. Nayak · J. Z. Weligamage
Queensland University of Technology, 2 George Street,
Brisbane, QLD 4001, Australia
e-mail: d.emerson@connect.qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_22,

than concentrating on single sites) can help apply consistency of approach to road safety across the entire network [5].

As late as 2008, Cairney writing in an Austroads Skid Resistance Review [6] described skid resistance as the best associated link with crash risk, citing before and after road resurfacing cases. A subsequent study by Piyatrapoomi et al. [7] in 2009 supports this observation. Analysis of the effect of skid resistance on crash rate was conducted on sections of national highway. The study established a skid resistance threshold for classes of roads where the probability of crash decreased at a given skid resistance value.

Our paper provides a new method of analysis for the whole road network. The method quantifies the magnitude of contribution of skid resistance and other roadway features in crash risk. The influence of the features is shown in selected regression formulae produced by the regression tree algorithm drawing on all classes of roads from the network. This method contrasts with the statistical studies that often attempt to control the effects of roadway features. In this study, the rule sets derived from the regression tree analysis generalize the range of relationships between crashes and roadway features for all classes of roadway.

This study relied on the outcomes of a series of studies by the authors, where data mining methods were applied to the skid resistance/crash rate problem. These studies contributed the following key understandings: (1) decision tree method utilizing roadway features can successfully predict the probability of a roadway having crashes [8]; (2) the existence of road segment crash proneness [9]; (3) cluster patterns based on roadway features delineate strong roadway relationships including the number of crashes experienced, seriousness of crashes, and wet crash proneness of roads [10]; (4) demonstration that the crash risk measure of absolute crash count (4-year crash count per km of roadway) was a superior method for data mining in comparison to the traffic-moderated crash rate measure (crashes per million vehicle km travelled) commonly used in statistical studies [11]; and (5) demonstration that models developed with the regression tree algorithm can predict the roadway crash count measure using roadway features [11]. These outcomes provided the foundation of this study.

This paper examined our prior roadway crash models and sought evidence for the proposition that the inverse covariance shown between skid resistance and crash rate is a cause/effect relationship.

The paper has the following sections. Section 22.2 describes the data understanding and preparation stages. Section 22.3 discusses the methods used to develop and evaluate the models. Section 22.4 presents and analyzes the results. The conclusion provides a summary of the outcomes and proposes future directions.

22.2 Data Understanding and Preparation

Data provided for data mining are generally collected for some other purpose and require analysis to assess its quality and develop understanding of links between the data and the data mining goal. Subsequently, transformational operations are required to produce the formatting required for the selected modeling processes [12]. The sequence of tasks in this study was as follows: reevaluation of the crash risk measure, subsequent investigation of the relationships between the roadway features and crash risk, and relational transformation to meet modeling requirements. Because of space limitations, only key points are discussed.

The sources of data were as follows: the annual road snapshot reports providing the characteristics of 1-km road segments, crash data, skid resistance surveys, texture depth surveys and road construction, and traffic source tables for the whole road network. The data mining analysis relied on a level of homogeneity of features along the length of each roadway segment, and results indicated that the homogeneity was adequate. The expected problem of missing data arose once the scope of analysis encompassed the whole network. Problems existed primarily with missing values and inconsistencies in roadway data and the skid resistance coverage of roadways. Skid resistance surveys were available for only 25 % of the roadway segments and limited the proportion of instances available for model training. In comparison, crash data were complete and consistent and had the advantage of providing the road segment details specific to the crash site.

22.2.1 Meeting the Business Goal

The project objective was the development of a method to identify the optimal skid resistance value for a given road segment using the crash count data mining model. This paper examined the rules produced from the models and drew conclusions about the relationship between skid resistance and road segment crash rate.

The simple and understandable modeling objective of '*predicting road segment crash rates from the roadway variables*' based on road attributes alone had to be abandoned because of the acute problems with data consistency and missing values described above. Thus, a new modeling objective was required. A *crash centric* method was devised to allow inclusion of the roadway data associated with each crash, and while more complex, the new method provided the advantage of a series of evaluations for each roadway along its length, one at each crash site, as well as allowing removal of the problem roadway attributes. Thus, the modeling objective changed to: '*predicting the crash count of a 1-km road segment at multiple crash sites using a combination of generalized road data and localized road data at each site*'. Hence, the new objective provided the opportunity to assess and compare the roadway's crash risk at a number of sites, where each had slightly different data.

22.2.2 Representing Crash Risk

The target crash count variable, produced for both 1-year and 4-year intervals, was calculated for each 1-km road segment by simply summing crashes that occurred on the roadway over the time period. Annual crash counts ranged from 1 to 32 crashes per annum, while 4-year crash counts ranged from 0 to 100 crashes, with averages calculated from the 4-year totals ranging from 1 to 25 crashes. The number of road instances dropped exponentially as crash counts increased, with the majority of crashes occurring on roads with few crashes. At the high end of the crash rate scale, the maximum value of 100 crashes per 4 years was limited to a single road segment.

The distribution of crashes and crash counts is shown on a chart of skid resistance versus daily traffic averages (AADT) in Fig. 22.3. The distribution demonstrates that most crashes occur on roads with 4 or fewer crashes per 4 years in the zone of low daily traffic rate. While this zone is of concern, these low-crash roads fail to respond to prediction by modeling with road attributes as input, indicating that their crash cause may lie elsewhere other than roadway features. The crash-prone roads of interest are those whose crash count is predictable using roadway features and generally have crash rates consistently above 4–8 crashes per 4 years (1–2 crashes per year) [9].

The concept of crash proneness is supported by the observation of the limited variation of annual crash counts over time. Trends showed that roads maintained their crash rate from year to year within a significantly narrow range. From this observation, deduction was made that the consistency was due to the action of a set of factors, and our objective was to link the roadway features, in particular skid resistance, with these factors.

When the roadway crash counts were aggregated into the average crash count groups, the distribution of the annual average crash values versus observed crash count approximated a Poisson distribution. Across the averages, the count of instances dropped exponentially with the increase in crash count, but the weight of observations in the distributions of the average had sufficient data elements to maintain their ‘normal-like’ distribution in lower to middle crash range. In higher ranges, the distributions became random yet remained discriminatory. To quantify the significance of this discrimination between the averages, a test of equality of averages was conducted using the ANOVA method. Resulting p values between 0 and 0.07 provided evidence supporting the significant or near-significant difference in distribution between adjacent annual crash count values. These results provided the comfort required to use average crash count as the data mining target.

22.2.3 Skid Resistance Versus Crash Count

Evidence was sought to verify the existence of a relationship between skid resistance and crash count. To provide comparative skid resistance groups, instances were categorized into the classes with the set of increasingly higher ranges of skid resistance shown in Fig. 22.1.

An ANOVA was applied to examine the status of the equality of the crash averages across the set of decreasing skid resistance categories, and the results

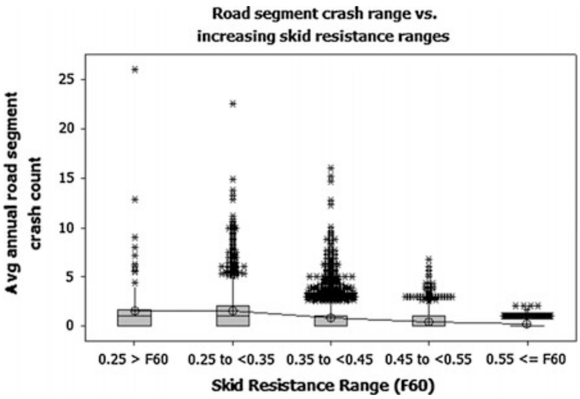


Fig. 22.1 Skid resistance versus road segment crash counts

showed that the average crash rates decrease with increase in skid resistance. The *p* value of 0 showed that differences were statistically significant below the skid resistance value of 0.35.

This relationship, illustrated in the skid resistance/crash count plot in Fig. 22.1, confirms the presence of the inverse skid resistance/crash count correlation throughout the whole range of values, with substantial reduction in crash rates in both the averages and outliers below the skid resistance value of 0.35.

22.2.4 Roadway Features

The following roadway features were assembled into a flat data file for modeling: (1) **roadway design** including roadway types (highway, urban arterial, etc.), speed limit, divided road (dual or single carriageway), lane count, and lane types (median, passing, etc.); (2) **roadway surface** including road surface friction (skid resistance), texture depth, seal age, and seal type; (3) **road surface wear and damage** with rutting and roughness; (4) **roadway features** such as roundabouts, bridges, intersections, and rail crossings; (5) **geometry** with horizontal and vertical alignment and terrain; (6) **demography and settlement** with urban or rural; (7) **traffic control** with traffic lights, signs, barriers, etc.; and (8) **traffic rates** including average annualized daily traffic (AADT) and percent heavy vehicle (% heavy). As modeling progressed, the list was refined to a set of 29 attributes from the classes listed.

22.3 Data Modeling and Results

Data mining algorithms were selected from predictive data mining family, with the regression tree type chosen for the following reasons: its ability to predict numerical results, the accuracy of models and ease of understanding of the tree diagram. The tree diagrams showed the contribution of input variable in classifying roadways into classes and optionally provided reference to regression formula for each class, thus providing a very rich basis for analysis.

Models were configured to predict crash rate, with the roadway variables as input and the road segment crash count as target. Once the configured algorithm was executed, the following outputs became available: a data mining model suitable for deployment with new data, rules suitable for analysis and the predicted crash count for each instance. Comparison of the actual and predicted crash counts allowed for calculation of the model correlation coefficient (r) and provided one of the key statistics used for assessment of models.

Models were trained on the set of instances with road segment having skid resistance values and subsequently deployed and tested on the whole data set. Regression tree configuration parameters were used to grow successive trees, with each iteration reducing the size and increasing the generality of the resulting model, producing each new model with fewer classes. Assessment was made by comparison of the resulting deployment correlation coefficients to find an optimal tree. The objective of selection was to choose the most generalized tree with the smallest number of classes that was capable of producing a high-deployment correlation coefficient in deployment.

22.3.1 Predictive Ability of Models

Most modeling was performed with *REPtree* because of its extreme speed, accuracy, and ease of interpreting the tree. Generation of the sequence of trees required for evaluation was managed by incrementally modifying the parameter *minimum total weight of instances per leaf (minNum)*. *MinNum* restricts tree growth by demanding the presence of certain number of instances meeting a condition before a split, and new branch in the tree is allowed. The set of *minNum* parameter values ranged from 2 to 128.

Trees were assessed by applying the criterion of selection of picking the smallest tree with near-optimal deployment correlation. The model with the *minNim* value of 8 was chosen. The results are shown in Table 22.1 and Fig. 22.2 with the optimal tree highlighted in the table and labeled A in the chart.

The chosen model had 488 nodes and approximately 224 leaf nodes/classes from almost 17 k of training instances. Following candidate models with 295 and 181 nodes were not chosen because of their loss of ability to describe the full range of crash values with the consequent drop in their number of classes.

22.3.2 The Regression Tree Mechanism

To demonstrate the mode of operation of the regression tree, a model with only two variables was built. The variables selected were average daily traffic rate

Table 22.1 Predictive accuracy of the baseline table: training and testing versus deployment

Tree decision splitting count (*minNum*)	Tree size	Training correlation coefficient	Deployment correlation coefficient
2	1696	0.9237	0.7787
4	751	0.9301	0.7801
8	**488**	**0.9217**	**0.7773**
16	295	0.9024	0.7741
32	181	0.8583	0.7499
64	89	0.8189	0.7097

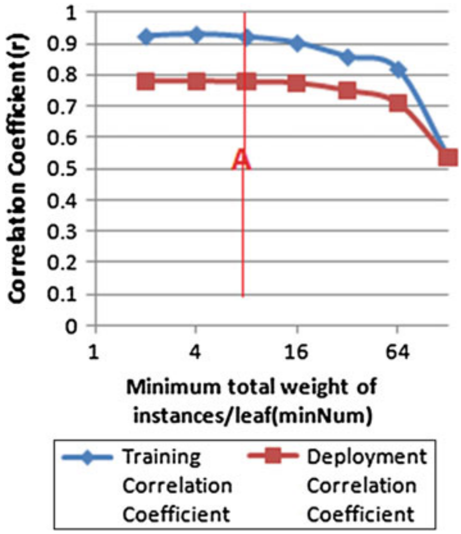


Fig. 22.2 Predictive accuracy of the baseline table: training/testing versus deployment

(AADT) and skid resistance (AVG_FRICTION_AT_60_100m). The fragment of the tree listed in Table 22.2 shows the mechanism of selection for the highest crash value of 48.42 available from this tree.

Visualization of the process outlined above is shown in Fig. 22.3. The chart shows the crash instances selected by the AADT range between 17,478 and 25,925 vehicles per day and subsequently by skid resistance value to below 0.23. The chart shows the intersections of the rules working to define the specified area.

The area of interest is the high-crash area with an average crash rate of 67 crashes, where the maximum rate of 100 crashes per 4 years is located.

The chart shows that the relationship between crashes and AADT is not linear, and thus, multiple rules are required to define the area. The target value predicted by the rule is calculated by the average crash counts of instances included in the scope of the rule. The process in this two variable example can be shown on a three-dimensional chart; however, in multivariate models, the selection is performed in multidimensional space.

Table 22.2 Rule selection with average daily traffic (AADT) and skid resistance (AVG_Friction_AT_60_100m)

Tree segment
| | **AADT >= 17478**
| | | **AADT < 25929**
| | | | **AVG_FRICTION_AT_60_100m < 0.28 : 48.42 (272.93/983.5) [137.03/1005.74]**

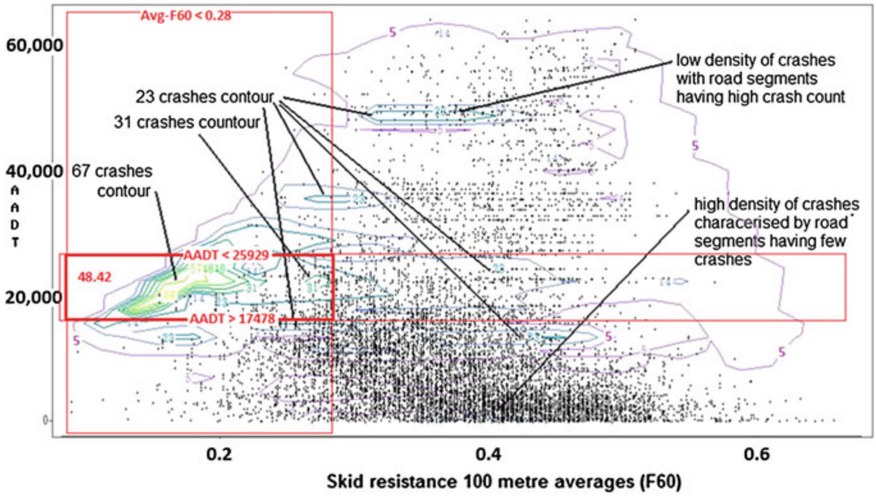


Fig. 22.3 Rule selection process using AADT and skid resistance to calculate crash count

22.3.3 Role of Skid Resistance in the Tree Structure Representing High Crash Rate

A more complex but still simplified multivariate model with a small tree and consequently with low correlation coefficient was selected for exploration of the most prominent variables. While it does not resolve the tree branches to the extent of the optimal tree, it shows the branching below the root node similar to that in the optimal tree.

The first split was performed based on daily traffic averages (AADT) of 9,872 vehicles per day. The section of the tree below that level was omitted because it generally describes roads with low crash rates, with many falling below 8 crashes per 4 years into the class of non-crash-prone roads [9] where resistance skid is of lower relevance. In the high-crash section of the tree (Table 22.3) for roads with a traffic

Table 22.3 Rule selection with average daily traffic (AADT) and skid resistance

<table>
<tr><th>Tree segment</th></tr>
<tr><td>AADT >= 9872</td></tr>
<tr><td>| AVG_FRICTION_AT_60_1km < 0.23---A</td></tr>
<tr><td>| | ROAD_TYPE = Main : **23.97** (29/101.74) [8/61.38]-------------------------------------B</td></tr>
<tr><td>| | ROAD_TYPE = Highway : **100** (69/0) [31/0]--C</td></tr>
<tr><td>| AVG_FRICTION_AT_60_1km >= 0.23</td></tr>
<tr><td>| | CRASH_SPEED_LIMIT < 65--D</td></tr>
<tr><td>| | | LANE_COUNT < 3.46 : **21.89** (2220.52/148.3) [1095.25/146.5]-------------------E</td></tr>
<tr><td>| | | LANE_COUNT >= 3.46 :**43.09** (358.05/469.08) [198.29/497.61]----------------- F</td></tr>
<tr><td>| | CRASH_SPEED_LIMIT >= 65</td></tr>
<tr><td>| | | AVG_FRICTION_AT_60_1km < 0.37 :**17.37** (1268.06/139.92) [655.06/133.25]--G</td></tr>
<tr><td>| | | AVG_FRICTION_AT_60_1km >= 0.37 :**10.96** (1649.13/37.09) [766.08/34.66]--H</td></tr>
</table>

Note Skid resistance is the variable AVG_Friction_AT_60_100m

threshold above 9,872 vehicles per day, skid resistance (AVG_FRICTION_AT_60_1km) is the second most significant variable in classifying the roads.

In this high-crash section of the tree, road segments have an average of over 20 crashes per 4 years with the highest crash count road having 100 crashes/4 year over that period. Examination of the table shows that the key roadway variables that combine with *skid resistance* are *road type, speed limit* and *lane count.* Skid resistance partitions the tree into classes above and below the 0.23 threshold. The zone below 0.23 has the highest crash count values, which is subsequently divided between *main roads* and *highways*. *Main roads* (label B) have an average of 23.97 crashes per 4 years, while highways have an averages crash rate of 100 (Label C). However, in this generalized model, subsequent splits are not present and consequently the model lacks detail. This highway passes through a regional city as a four lane, undivided highway with intersections at each bock, making a high maneuver environment requiring a high-demand skid resistance classification. A more detailed tree and access to satellite mapping facilities provide this crucial context.

Above the skid resistance value of 0.23, speed limit is significant and divides the roads into groups above and below 65 km/h, into two zones: the zone below 65 km/h (Label D) and the 80, 100 or 110 speed limit roads above. The 60-km speed limit zone is divided into two categories at or below 4 lanes by a threshold of 3.64.

The average crash count for 1 and 2 lanes is around 21 crashes per 4 years (Label E), whereas roads having 4 lanes in the 60-km speed limit (Label D) experience a doubling of crash rate with an average 4-year crash count of 43 crashes/4 year (Label F). When the speed limit is above 65 km/h, i.e., 80, 100, and 110, the effect of skid resistance on crash rate is shown in subsequent rules. The average crash rate is 17.37 for the group of low skid resistance roads below 0.37

Table 22.4 The effect of skid resistance in low and high crash count rules

Rule 1: Road segments with low crash rate	Rule 100: Road segment with high crash rate
IF AADT <= 6105 CRASH_SPEED_LIMIT > 85 AADT <= 1028.5 THEN CrashCount_4Yr_Total = + 0.0002 * SEAL_AGE + 0.0012 * Text_Depth_SPTD_OWP **- 0.165 * AVG_FRICTION_AT_60_1km** + 0 * AADT - 0.0007 * TRAFFIC_PERCENT_HEAVY + 1.4762 [668/4.613%]	IF LANE_COUNT <= 2.075 Text_Depth_SPTD_OWP <= 1.075 AADT > 15080 SEAL_AGE <= 18.9 AVG_FRICTION_AT_60_1km <= 0.347 THEN CrashCount_4Yr_Total = - 31.7714 * LANE_COUNT + 0.5181 * Text_Depth_SPTD_OWP **- 49.0451 * AVG_FRICTION_AT_60_1km** + 0.0001 * AADT + 96.7313 [83/40.725%]

Note Skid resistance is the variable AVG_Friction_AT_60_1 km

(Label G), but with the higher skid resistance, the group crash rate lowers to an average of 10.96 (Label H).

These rules show a strong relationship among skid resistance, other roadway variables and crash rate.

22.3.4 High Crash and Low Crash Rules

Regression trees algorithms may be configured to show a regression formula for each leaf node.

The formula displays the contribution of features participating in a rule to calculate crash count. In this example, two sample rules are used to contrast the effect of skid resistance in crash rate calculation in low-crash and high-crash environments. The model was developed using an alternative regression tree method designed to generate rules, aptly named M5 Rules [13]. The tree used the 29 roadway attributes described above, with the *mimNum* parameter set to 32, resulting in an optimal tree with 109 rules. Only key attributes are shown in the samples to avoid unnecessary complexity. While these rules are taken from the full set for examination, the set needs to be executed in sequence and instances removed once identified: thus given rules cannot operate in isolation. The selected rules are shown in Table 22.4.

The rules show that in low-crash situations skid resistance in Rule 1 has only a miniscule effect with a coefficient of 0.165, whereas in high-crash situation, along with lane count and AADT, skid resistance is a dominant variable with a coefficient of 49.05. In rules with high crash count, skid resistance frequently is a significant variable.

22.4 Summary and Conclusion

The study shows that data mining methods are capable of modeling road segment crash rates based on roadway features. In relative importance, the road traffic variable AADT plays the most significant part by partitioning roads into low- and high-crash groups above and below 9,872 vehicles per day, respectively. In the high-crash group, the primary partitioning attributes for crash rate are *skid resistance, road type, speed limit* and *lane count*. S*kid resistance* is a significant variable in many road segments. In the high-crash roads, the regression formulae provide evidence supporting causative relationship between skid resistance and crash count. In the lower crash zone below 9,872 vehicles per day, skid resistance is of low importance, where roads have more in common with non-crash roads than roads with crashes.

Further work is required for detailed examination of the rules and trees to work toward defining road classes, relating their characteristics to crash types and understanding the dynamics so that action can be taken to produce safer roads.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management (CIEAM), established, and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by CIEAM. This CIEAM study is a component of a project involving QUT and QDTMR in the development of a skid resistance decision support system for road asset management. Queensland Department of Transport and Main Roads provided the road and crash data sets. Data mining operations and presentations were performed in the WEKA and SAS platforms, and substantial computer processing was performed on the QUT High-Performance Computing Facility (HPC). The views presented in this paper are of the authors and not necessarily the views of the organizations.

References

1. World Health Organization (2003) Global health: today's challenges. World Health Organization. http://www.who.int/whr/2003/en/Chapter1.pdf. Accessed 12 Nov 2011
2. QDTMR (2005) Road planning and design manual, design philosophy, Chap. 2. Queensland Department of Transport and MainRoads. http://www.tmr.qld.gov.au/Business-and-industry/Technical-standards-and-publications/Road-planning-and-design-manual.aspx. Accessed 11 Feb 2011
3. Norsemeter (2009) Primer on modern runway surface friction measurement. Oscarsgate 27, Oslo, Norway: Norsmeter Friction. http://www.norsemeter.no/Downloads/Files/Runway%20Primer.pdf. Accessed 15 June 2009
4. Venz P (2005) Skid resistance survey Bruce highway (Middle Creek to Christies Rd), Queensland, Ed. Department of Transport and Main Roads Queensland, Brisbane
5. ARRB Group (2011) Tools for improving road safety. ARRB Group. Accessed from http://www.arrb.com.au/Safe-Systems/Tools-for-improving-road-safety.aspx, July 2011
6. Cairney P (2008) Road surface characteristics and crash occurrence: Austroads skid resistance literature review. Austroads Aust. Accessed from http://nla.gov.au/nla.cat-vn4406507. Accessed 12 June 2010

7. Piyatrapoomi N, Weligamage J (2008) Probability-based method for analysing the relationship between skid resistance and road. In: Proceedings of the laboratory of transportation engineering and infrastructure
8. Nayak R, Emerson D (2010) Using data mining on road asset management data in analysing road crashes. In: Proceeding of the 16th annual TMR engineering & technology forum, Brisbane, Australia, 20 July 2010
9. Nayak R, Emerson D, Weligamage J, Piyatrapoomi N (2011) Road crash proneness prediction using data mining. In: Proceedings of the 14th international conference on extending database technology, Uppsala, Sweden, 22–24 March 2011
10. Emerson D, Nayak R, Weligamage J (2011) Identifying differences in safe roads and crash prone roads using clustering data mining. In: Proceedings of the World Congress on Engineering Asset Management, Cincinnati, Ohio, USA, 3–5 Oct 2011
11. Emerson D, Nayak R, Weligamage J (2011) Using data mining to predict road crash count with a focus on skid resistance values. In: Proceeding of the 3rd international road surface friction conference, Gold Coast, Queensland, Australia, 15–18 May 2011
12. Witten I, Frank E (2005) Data mining: practical machine learning tools and techniques, 2nd edn. Morgan Kaufmann, San Francisco
13. Quinlan JR (1992) Learning with continuous classes. AI 92

Chapter 23
Value Creation Through Collaborative Operating Environments: Challenges in the Offshore Petroleum Sector

Ole-Erik V. Endrerud and Jayantha P. Liyanage

Abstract Since late 80s and early 90s, the offshore oil and gas industry has seen a decline in the rate of large discoveries on the Norwegian Continental Shelf; moreover, a majority of fields on the NCS are mature. To cope with increasing risk and lower margins, integrated operations have become a modern solution that introduces novel business models innovations. This greatly opens up many opportunities for value creation, inclusive strong business-to-business relations, new organizational forms, and technological solutions that can improve decision making, operational effectiveness, safety, and production leading to increased recovery, lower operational costs, and extended field life. Collaborative operating environments are an important element of integrated operations, functioning as nodal points for data, people, decision, and work processes, and play an important role in bringing operators, service companies, and experts closer together in collaborative relationships. In this new operating environment, business-to-business relationships have a defining role in increasing value creating potential. This paper elaborates on the current performance gaps based on a project that aimed at identifying performance enhancement potential in B2B context, between operator and service companies in the oil and gas industry operating in collaborative operating environments.

23.1 Introduction

Energy is the catalyst for societal and economical development. Until today, petroleum resources have primarily met energy demands, especially oil but more recently also gas. The Norwegian continental shelf (NCS) has been an important

O.-E. V. Endrerud (✉) · J. P. Liyanage
University of Stavanger, Kristin Bonnevies vei 22, 4021 Stavanger, Norway
e-mail: ole-erik.v.endrerud@uis.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_23,

provider of petroleum resources to international markets, but has recently seen a decline in production due to challenges in oil recovery and reserves replacements [1, 2]. In addition, increasing scope of maintenance on aging platform and field assets together with modifications and development to capitalize on value potentials contributes to drastically rising operating costs [1, 2]. Generally, operational and business risks of operators of complex industrial assets on NCS are on an increasing trend [3, 4].

To cope with increasing business risk and lower margins, integrated operations have been introduced as a modern solution that introduces novel innovations in organizational and business models. This opens up many opportunities for value creation, inclusive strong business-to-business (B2B) relations, new organizational forms, and application solutions that can improve decision making, operational effectiveness, safety, and production leading to increased recovery, lower operational costs, and extended field life. A manifestation of the integrated operations strategy are collaborative operating environments (COEs), which are centers where service companies, experts, and operators work together in collaboration to solve complex challenges supported by advanced information and communication technology (ICT). In this new operating environment, collaborative B2B relationships have a defining role in increasing value creating potential. Service companies in the oil and gas industry strongly affects an operator's performance effectiveness, as operators rely on increasingly more outsourcing to perform complex operations. Several areas of improvement in the B2B partnership between service companies and operator have been identified, addressing the opportunities for additional value creation by implementing integrated operations and COEs.

Integrated operations have a high value potential. In 2006 and 2007, OLF—Norwegian Oil Industry Association—released two new reports on the value potential of changing operating model on NCS, which indicated that implementing integrated operations had a value potential of 300 billion Norwegian kroner (NOK), assuming an oil price of USD 55 per barrel of oil [5, 6].

This paper elaborates on the current performance gaps based on a project that aimed at identifying performance enhancement potential in B2B context between operator and service companies in the oil and gas industry operating in collaborative operating environments.

23.2 Methodology

The project was conducted through observations in the facilities of the operator and service company A and through group interviews with 2–4 managers or professionals from each company separately. In total, 12 persons were interviewed. Interviews were purely qualitative and were used to establish three criteria: (1) level of integration between the operator and service companies, (2) decision environment and work practices, and (3) relationship health. These three categories were divided into sub-categories.

Service company A was located in their own physical facilities collaborating in a virtual facility with other service companies and the operator, both offshore and onshore. Service companies B and C were located at the operator's own onshore facilities and had the possibility to collaborate virtually and physically with other COE partners. Service companies A and B operated key business processes for the operator, while service company C delivered support functions to the key business process. In the next chapter, the most important findings are presented.

After interviews and observations had been conducted, discussions and interviews with academic experts at University of Stavanger and professionals at the operating company were used to analyze and identify the key challenges indicated in the interviews and observations.

23.3 Status Review: The Growth of Complexity

In the modern business environment characterized by increasing complexity, rising costs, globalization, and rapid growth of collaborative application technologies, companies see potential in outsourcing non-core functions and as a consequence joining supply chain networks [7]. Operators have to put more focus on being excellent at core competencies, thus requiring services from third party contractors to perform business functions [8], often specialized and complex functions [7]. For example, in drilling, specialized personnel from a service company are contracted to perform highly specialized drilling services, e.g., geosteering; meanwhile, the operator focuses on reservoir engineering and well planning.

As a response to rapid growth of complexity on the NCS (driven by a need to increase recovery and slow down cost acceleration), integrated operations are proposed as a model of operations for complex offshore assets [9]. Integrated operations are defined as "*use of information technology to change work processes to achieve better decisions, remote control equipment and processes, and to move functions and personnel onshore*" by the Norwegian Ministry of Oil and Energy [10]. Moreover, integrated operations transform traditional offshore assets into smart plants (smart sensory equipment, real-time monitoring, efficient real-time communication, 3D visualization, etc.) supported by highly advanced onshore operational centers [4, 11], which is named COEs. Hence, COEs function as nodal points for data, people, decision, and work processes, and play an important role in bringing service companies, operators, and experts closer together in collaborative B2B relationships.

Integrated operations are one of the areas of priority for the petroleum industry today to reduce risk and sustain commercial success [2, 9]. In order to achieve these ambitious goals, traditional organizational barriers and functional silos have to be torn down, and a whole new way of collaboration and B2B partnerships must emerge.

23.4 What Is a COE?

Collaborative operating environments (COEs) in principle apply information and communication technology (ICT) and advanced application solutions to integrate decisions and work management processes in operational centers. This involves bringing together and integrating functions and disciplines across organizational boundaries and is a complex coordination and managerial effort only attainable by the use of ICT and other technological solutions. COEs have six building blocks: (1) humans, (2) organizations, (3) technology, (4) facilities, (5) work practices, and (6) business networks, as illustrated in Fig. 23.1.

Quality, standardization, storage, transfer, and visualization of data are challenges in COEs. These challenges are met with attempted novel solutions such as Web 2.0 and other groupware solutions, as well as acceptance of industry data standards by all major operators and service companies, such as WITSML (Wellsite Information Transfer Standard Markup Language) for drilling data [8].

Reengineering the operational organization from a traditional silo structure into a multi-functional and multi-disciplinary structure involves restructuring layouts of facilities, organizations, and work practices. Facilities in COEs are both virtual and physical. Virtual in the sense that partners of COEs—operator and service companies—collaborate in virtual spaces through technology by sharing video, audio, data, information, etc., through computer screens, phone and radio, software applications, etc. Physical facilities transform from traditional office landscapes organized according to organizational functions and disciplines into multi-disciplinary and multi-functional operational centers with visualization tools, communication tools, analyzing tools, etc. COEs can represent a business process, such as drilling or production optimization. These two types of facilities are combined in COEs to access distributed knowledge, information, and data across dispersed locations and organizations in real time and to collaborate on achieving the best decisions and solutions.

The road map of development illustrated in Fig. 23.2 was created by OLF as a vision of how integrated operations could mature over time and realize enhanced value creation [12]. Two important steps of progress should be noticed—Generation 1 and Generation 2—that represents changes on managerial, organizational, and technological side of asset management.

Traditional operational models had a clear division between offshore and onshore organization in operating companies; functions and disciplines were organized in silos; and experts were field or platform specific. When integrated operations and COEs started emerging on the NCS, the main focus was to integrate onshore and offshore organizations within operators and provides onshore support to offshore operations, referred to as Generation 1. This primarily lead to enhanced communication and coordination between personnel offshore and onshore, which means less traveling for experts, effective sharing of information, and increased productivity as operations and teams became more integrated [11].

Fig. 23.1 Six important elements of modern collaborative operating environments

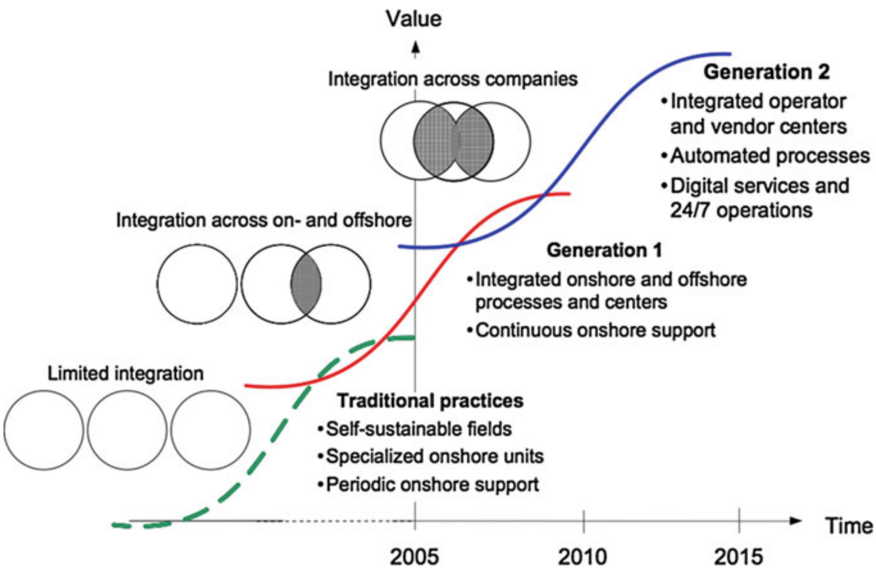


Fig. 23.2 Development stages of integrated operations and collaborative operating environments

The current development of integrated operations and COEs within several operators and service companies on the NCS is moving toward Generation 2, a more comprehensive and holistic model of operations. In Generation 2, a restructuring of physical and virtual facilities has been accomplished, and focus is shifted to broad integration of vendors and service companies in collaborative B2B relationships. In the modern business environment characterized by increasing complexity, rising costs, and rapid growth of collaborative application technologies, operators see potential in outsourcing non-core functions and as a consequence establishing collaborative B2B relationships with service companies and vendors in integrated operations and COEs.

23.5 Performance Gaps and Challenges Identified in COEs

A number of interesting issues emerged during the study varying from more technological aspects to the ones that can be discussed as soft issues. For the simplicity of explanation, they were classified into five main segments, namely: (1) *decision support tools*, (2) *knowledge integration*, (3) *B2B ties between service companies*, (4) *relationships within onshore/offshore virtual teams*, and (5) *partner perception*. Underlying performance gaps of each segment are discussed in the following sections.

23.5.1 Decision Support Tools

Application solutions are one of the benefits of advanced ICT developments, and collaborative decision support tools are one category of such application solutions that can take COEs one step further to enhance decision quality and coordination, as well as enhance knowledge integration. Hence, all companies working in COEs should implement a decision support tool in their work processes. The mission of collaborative decision support tools is to handle data integration and visualization, improve awareness and collaboration, and support faster and better decisions, work process compliance, and a decision model (e.g., based on Bayesian networks) to act as an expert system for decision support [13]. Collaborative decision support tools address coordination complexity in virtual teams by effectively filtering and transferring information and create a mutual understanding in problem solving. A decision model can give advice on what decision gives the best outcome.

However, the scope of involvement and type of services a company takes on should play a role to whether develop a company-specific tool or not. For example, a drilling contractor is responsible for delivering a borehole for an operator, thus having the responsibility for maintenance and operations of the drilling rig. A service of this scope could justify that a drilling contractor develops its own decision support tool, to incorporate all of its own lessons learned and knowledge created during drilling for different operators, and its own work management system. On the other hand, a service company with a supporting role could not justify developing a company-specific decision support tool. For example, a company that monitors and simulates fluid properties in the borehole is not dependent on having a sophisticated decision support tool. However, it is important to be a part of an operator's decision support tool in order to comment on challenges and notify about developing challenges.

It was evident from the case study that even though technological advanced, the COE was lacking an advanced decision support application. However, development and implementation of a collaborative decision support tool was under ways in the studied COE, especially by the operator, as they started to see decision complexity as a challenge when a vast amount of locations and companies were involved. Most service companies had not developed any specific collaborative decision support tools, but plans were there to do so in the future.

23.5.2 Knowledge Integration

Knowledge integration is the synthesis of individuals' specialized knowledge into situation-specific systemic knowledge [14]. To achieve value creation through COEs, it is important to achieve knowledge integration in the virtual teams (teams where participants originate from multiple companies and locations). Everybody cannot know everything, but should know where to find the needed knowledge, and everybody should have a basic understanding of different disciplines, in other words, creating a transactive memory. Problem solving is a social arena where knowledge integration can take place and be applied, and problem solving involving many a group of multi-disciplinary experts is an important part of COEs. In particular, because several companies are involved in COEs, knowledge integration between companies, functions, and disciplines is crucial to access and capitalize on distributed knowledge available in a large network of experts within all companies involved. The project identified a performance gap for knowledge integration in COEs, particularly in making relationships and locating experts efficiently. There does not seem to exist some kind of "yellow pages" or similar applications to locate experts with a specific knowledge or specialty in COEs, as a consequence location of an expert relies on personal relations or that an expert is located within the team in the studied COE. Such "yellow pages" solutions have a potential to enhance knowledge integration and make it easier to locate the right experts even though one has never met the relevant expert.

Another part of knowledge integration is to create a large and robust knowledge repository to access previous lessons learned. For example, if a situation is starting to develop, it is valuable to be able to easily access what decisions were taken in similar situations before and what the outcome was, for making a good decision now. A service company was working on such a knowledge repository to collect reports on incidents, lessons learned, project reports, and other valuable documentations. Such a solution can be integrated into collaborative decision support systems as a knowledge base. The operator was also developing a comprehensive collaborative decision support tool also supporting knowledge integration, and as part of that, a knowledge repository was to be created. However, all companies should have a user-friendly and searchable knowledge repository as a part of their knowledge integration strategy.

23.5.3 Few Strong B2B Ties Between Service Companies

Companies with strong ties interact frequently and directly, compared to companies with weak ties that are characterized by infrequent interactions and distant relationships [15]. Strong ties promote trust and reciprocity. In the case of COEs where companies operate under long-term and stable conditions, strong ties are favorable because of the close collaboration between participants is needed for performance efficiency in COEs. Traditionally, operators and service companies are related with strong ties. Despite so, service companies do not traditionally share strong ties, probably due to protectionism of intellectual property and fear of loosing competitive advantage. However, if service companies collaborate and share strong ties, COEs become more autonomous, and it gives access to a larger base of distributed knowledge for service companies involved. Studies from other industries have shown there is a large potential in sharing knowledge and expertise among service companies in business networks [16]. On the other hand, challenges are introduced when strong ties are created between service companies. Lowered betweenness centrality for the operator can induce a power asymmetry in business networks. Also protecting intellectual property can also become an issue that needs to be resolved to make collaboration possible.

The project identified that service companies were collaborating without involvement from the operator, in spite of being competitors, in COEs. However, this collaboration was only in a small scale. It was highlighted by several companies that such strong ties and collaboration between service companies were valuable when it occurs.

23.5.4 Deficiency in Relationships Within Onshore/Offshore Virtual Teams

Teams are distributed between offshore and onshore locations, and communication between offshore and onshore occurs through video conferencing quite often, or through phone calls. Two persons interviewed during the project experienced performing a function from an offshore location first, but was later moved onshore. They highlighted that communication between offshore and onshore personnel was at times a challenge. There can be several reasons for this: First, a lack of mutual understanding of the benefits from onshore support. Offshore personnel might not see the benefit because they do not believe that someone sitting hundreds of kilometers away onshore can understand the issue. Second, a difficulty in building relationships and trust. When these two individuals moved onshore and into the

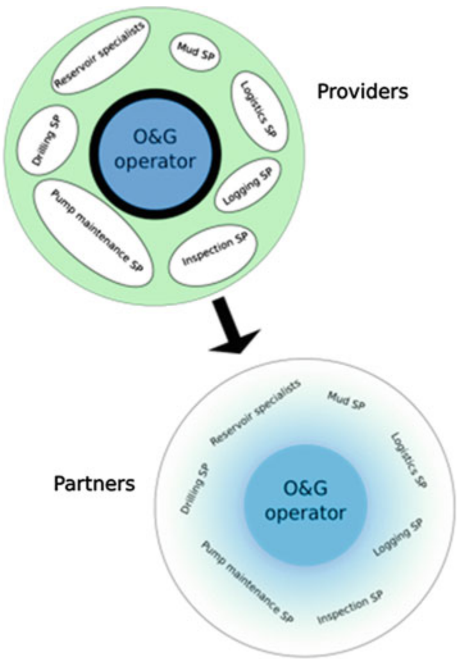


Fig. 23.3 Partner versus provider perception

operator's location, it was much easier to build strong relationships between professionals at the onshore location and get that mutual understanding and high level of trust needed to perform well in a team. This proves there exists a performance gap in relationships between participants in virtual teams, especially between the onshore and offshore participants.

23.5.5 Perception of Service Companies

The difference between partner perception and provider perception has implications for communication, decision making, knowledge and information sharing, partaking in risks and benefits, and incentive systems. Figure 23.3 illustrates the differences between the two perceptions, where organizational borders are cut and clear in the provider circle, and each company has an information and knowledge boundary. However, as relationships move toward partner perception, boundaries are blurrier and there exist no information and knowledge boundary.

If service companies are considered a partner, communication is not command based but rather information based, and decisions can be made by service companies based on information. This enhances the ability to contend with a dynamic operating environment and deal with operational risks and opportunities swiftly as they develop. Because service companies are considered partners, not providers, decision making authority must be transferred from operator to service companies, which entails a great deal of trust between partners. However, decentralization of decision authority is essential for COE to be effective and handle dynamic complexities.

The project identified that service company perception is changing in COEs. Service company employees feel as a part of the operator's own staff, communication is information based and informal, and everyone has access to data, information, and facilities. However, perception change has to take place in all parts of a COE; hence, a high level of trust must be established. None of the service companies asked during the project indicated a perception of being a full partner. In addition, there is still a relatively centralized decision authority.

A number of solutions can be implemented in order to overcome some of the principal issues highlighted in Sect. 23.5. These solutions are discussed in Endrerud and Liyanage [17].

23.6 Summary and Conclusion

Many studies have focused on technological developments, but not a lot has been done in the past looking at the role of collaborative B2B relationships in COE. Drivers of performance are not exclusively a question of technology any more. Five segments were identified as challenging: (1) *decision support tools*, (2) *knowledge integration*, (3) *B2B ties between service companies*, (4) *relationships within onshore/offshore virtual teams*, and (5) *partner perception*. In particular, decision support tools and knowledge integration play an important role in realizing some of the value potential forecasted by OLF. Integrated operations and COEs are here to stay as a modern and novel solution for the petroleum industry, and capitalizing on improvement potentials in the relationship between operator and service provider is one important part of improving value creation and performance.

Future research should extent the scope of the study to include offshore and thus avoid bias. In addition, it was only a qualitative study without a possibility of comparison with other COEs. Another important part of B2B relationships is the incentive system, which has not been covered extensively in this study. A recommendation for future work is to look at how traditional incentive systems have to change to create a win-win situation for operators and service companies in collaborative B2B relationships as a part of COE.

References

1. The Norwegian Oil and Energy Ministry (2010-2011) White Paper No. 28 On petroleum activity, The Norwegian Oil and Energy Ministry, Oslo
2. KonKraft (2009) KonKraft report 2: production development on the Norwegian continental shelf. KonKraft, p 97

3. Liyanage JP, Bjerkebaek E (2007) Use of advanced technologies and information solutions for North Sea offshore assets: ambitious changes and socio-technical dimensions. Int J Int Technol Inf Manage 15(4):1–10
4. Liyanage JP, Herbert M, Harestad J (2007) Smart integrated eOperations for high-risk and technologically complex assets: operational networks and collaborative partnerships in the digital environment. In: Wang WYC, Heng MSH, Chau PYK (eds) Supply chain management: issues in the new era of collaboration and competition. Idea Group Inc, pp 387–414
5. Oljeindustriens Landsforening (OLF) (2006) Verdipotensialet for Integrerte Operasjoner på Norsk Sokkel, 2006, Oljeindustriens Landsforening (OLF): Stavanger
6. Oljeindustriens Landsforening (OLF) (2007) Oppdatert verdipotensiale for integrerte operasjoner på norsk sokkel, 2007, Oljeindustriens Landsforening (OLF): Stavanger
7. Gereffi G, Humphrey J, Sturgeon T (2005) The governance of global value chains. Rev Int Polit Econ 12(1):27
8. Mundim APR, Rossi A, Stocchetti A (2000) SMEs in global market: challenges, opportunities and threats. Braz Electron J Econ 3(1)
9. Liyanage JP (2008) Rapid virtual enterprising to manage complex and high-risk assets. In: Putnik GD, Cruz-Cunha MM (eds) Encyclopedia of networked and virtual organizations. Business Science Review, Hayden, pp 697–704
10. The Norwegian Oil and Energy Ministry, Olje-og-Energidepartement DK (ed) (2004) White Paper No. 38 On petroleum activity (2003–2004), Oslo
11. Vindasius J (2008) The integrated collaboration environment as a platform for new ways of working: lesson learned from recent projects. In: SPE intelligent energy conference and exhibition 2008, Society of Petroleum Engineers (SPE), Amsterdam, The Netherlands, p 10
12. Oljeindustriens Landsforening (OLF) (2005) Integrated work processes: future work processes on the Norwegian Continental Shelf, S.G.I. Operations. Oljeindustriens Landsforening (OLF), Stavanger, p 29
13. Fjellheim RA, Bratvold RB, Herbert M (2008) CODIO—Collaborative decision making in integrated operations. In: SPE intelligent energy conference and exhibition 2008, Society of Petroleum Engineers (SPE), Amsterdam, The Netherlands
14. Alavi M, Tiwana A (2002) Knowledge integration in virtual teams: the potential role of KMS. J Am Soc Inform Sci Technol 53(12):1029–1037
15. Hansen M (1999) The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits. Adm Sci Q 44:83–111
16. Dyer JH, Nobeoka K (2000) Creating and managing a high-performance knowledge-sharing network: the Toyota case. Strateg Manag J 21:345–367
17. Endrerud OE, Liyanage JP (2012) Collaborative operating environments in the offshore petroleum sector—Performance solutions for added value in B2B partnerships. In: World congress on engineering asset management 2012, Daejong, South Korea

Chapter 24
Collaborative Operating Environments in the Offshore Petroleum Sector: Solutions for Added Value in B2B Relationships

Ole-Erik V. Endrerud and Jayantha P. Liyanage

Abstract Collaborative operating environments (COE) are a new concept for the offshore petroleum sector and have been introduced as a part of integrated operations as a solution to mitigate risks and to capitalize on value creating potentials. To some extent, military, aviation, and aerospace industries have used COE with encouraging success, yet still left with some notable challenges concerning modern developments in social, economical, technological, and political terms. To reflect on the developments in the Norwegian petroleum sector, the Norwegian Oil Industry Association (OLF) published a report in 2007 on the value potential for integrated operations, which concluded that a value of as much as 350 billion NOK could be realized by the offshore petroleum sector in Norway. However, realizing this value potential does not solely rely on pure technological innovations, but various other parameters that contribute to boost performance. This particularly involves relations and interactions between stakeholders of producing assets in COE—operational centers used in integrated operations. In a project conducted recently within the offshore petroleum sector, it was revealed that, regarding today's implementation status of COE, there exist a number of performance gaps and challenges, particularly in the business relationships between operators and service companies. The findings from this project have been elaborated in another paper. As a continuing action, this paper tries to propose improvement solutions to capitalize on these performance gaps, create the right win–win scenario for all stakeholders, and present factors affecting performance in business-to-business (B2B) relationships.

O.-E. V. Endrerud (✉) · J. P. Liyanage
University of Stavanger, Kristin Bonnevies vei 22, 4021 Stavanger, Norway
e-mail: ole-erik.v.endrerud@uis.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_24,

24.1 Introduction

The world depends on energy to develop. Petroleum resources have primarily met energy demands as a cheap energy resource, especially oil but in recent years also gas. The Norwegian continental shelf (NCS) has been an important provider of oil and gas for decades, but production has seen a noticeable decline since year 2000 [1, 2]. However, both the petroleum industry and the national government believe that even though NCS is mature, the sector still holds a high value potential if increased recovery of fields already in production can be achieved, and cost levels can be controlled. On another hand, squeezing that last drop of oil out of a mature reservoir and controlling costs of asset management is a complex and demanding task, in need of a holistic and systematic approach. Changing asset management model by implementing integrated operations is a hot topic in the offshore petroleum industry in Norway.

Collaborative operating environments (COE) are an important part of an integrated operations strategy, functioning as a hub to capitalize on dispersed knowledge, collaboration, and advanced information and communication technology (ICT). Using COEs to manage complex assets have been investigated by several researchers. Liyanage [3] investigates how the combination of advanced ICT, social networking, and business networks creates a state of "hybrid intelligence" to manage complex assets. Vindasius [4] elaborates on experiences from implementation of COEs and defines three different development paths implementation can take, and what value one can gain from the different implementation strategies. Liyanage et al. [5] look into the current development and change process ongoing in the Norwegian petroleum industry, and point out that this change creates a new organizational setting and important changes in decision loops and work processes. Cohen [6] presents an early example of benefits from virtual teamwork in early COEs, which show remarkable improvements in cost reduction and problem solving effectiveness.

Having said that, the solution to improve recovery and lower costs is not a purely technical question and a traditional engineering challenge. A previous paper by the authors Endrerud and Liyanage [7] elaborated on the identified performance gaps and challenges in collaborative business-to-business (B2B) relationships in COEs in a project from the offshore petroleum industry. The findings from this project indicated that technology is not the major constraint any more. Moreover, soft characteristics, such as trust and strong relationships between service companies and between service companies and operator, have a potential to increase asset performance if proper solutions can be found. The scope of this paper is to present important success factors for performance effectiveness in B2B relationships between operators and service companies and propose solutions to meet the challenges presented by Endrerud and Liyanage [7].

24.2 Methodology

In order to get a deeper understanding of success factors for B2B relationships between operators and service companies, up to date, literature on relevant theories was evaluated and is given in Sect. 24.4. In addition, a project was conducted to identify challenges related to B2B relationships between operator and service companies working together in COEs in the Norwegian offshore petroleum industry. The finding from this project is briefly given in Sect. 24.5 and was presented in another paper by Endrerud and Liyanage [7]. The solutions presented in this paper to meet these challenges were found through discussions with industry professionals and academic experts, as well as from evaluations of industrial experiences and developments through literature and discussions.

24.3 What Is a COE?

A COE is a particular setting, physical or virtual, where partners actively work together to control the functioning of a machine, process, or system with the support of advanced application technologies. This definition includes people working together in physical facilities, virtual facilities, or both, supported by advanced application technologies, which is the key enabler for COE. Examples of COE are onshore operating centers with collaboration rooms, videoconferencing capabilities, simulation tools, visualization tools, remote operation capabilities, and online-shared workspaces, which are used to collaborate between dispersed locations and across organizational borders in real time. For more information on COEs, see Endrerud and Liyanage [7].

24.4 Escalating Complexity and Factors Affecting Performance in B2B Relationships

As said before, squeezing that last drop of oil out of a mature reservoir and controlling costs of asset management is a complex and demanding task. In response to this, the petroleum industry has introduced integrated operations as a new operating model. Integrated operations are defined as "*use of information technology to change work processes to achieve better decisions, remote control equipment and processes, and to move functions and personnel onshore*" [2]. Gereffi et al. [8] evaluate global value chain networks and have identified a current trend of increased outsourcing and focus on core competencies, which is at the core of integrated operations and work in a COE. The mission is to integrate multiple functions and disciplines from various locations and companies to achieve better decisions through collaboration and real-time asset data. Benefits from integrated operations are many, for example, increased effectiveness, improved HSE performance, synergy effects, enhanced communication, and risk and gain sharing.

Gulati et al. [9] define three dimensions that affect a focal company's performance in a business network: centrality within the network, configuration of ties, and partner profiles.

- ***Centrality*** refers to how exposed a company is to information and what control a company has over information. Three measures of centrality exist. Degree centrality refers to how many ties a company has in a business network. Betweenness centrality refers to what degree information has to pass through a company to reach other network participants. Closeness centrality refers to how fast information moves in a B2B tie [9]. COEs will lead to a higher degree centrality, and a high degree of closeness centrality must be maintained in order to have a fast and reliable information flow. Also, a company with high degree of centrality can achieve a higher degree of continuous improvement due to the larger access to knowledge, innovations, and experiences.
- ***Configuration of collaborative B2B ties*** refers to how interconnected companies in a B2B network are, and how frequent they interact [9]. Ties can be strong or weak. Strong ties between two companies mean frequent and direct interactions, and weak ties the opposite. Hence, strong ties are favorable between operator and service companies and between service companies to achieve productive and effective collaboration.
- ***Partner profile*** is a binary dimension between two companies and is divided into three categories: status, technology distance, and partner centrality. Status refers to benefits small firms can realize by partnering with a high status firm. Technology distance is defined by Gulati et al. [9] as similarity between two companies' innovative activity, but can also refer to similarity in ICT and application technologies between two companies in a COE context. Partner centrality refers to whether the partner has many bridging ties, hence access to knowledge and innovations outside the business network. For companies in a COE, to interface well, technology distance must be small. Status is relevant for network governance, as high status companies can create relationships with a more equal power symmetry. For operators, it can be beneficial to include service companies with high partner centrality to access a larger distributed knowledge base for problem solving and innovations.

Furthermore, knowledge management lies at the core of COEs. Liyanage [3] presents a model of "Hybrid Intelligence" saying that modern complex business environments require companies to acquire core distributed knowledge in order to achieve sustainability and performance, and these processes are extensively supported by "virtuality" orchestrated by modern ICT and socialization processes within business networks. In that sense, it is appropriate to address knowledge integration. ***Knowledge integration*** is a process where individual specialists' tacit knowledge is combined to solve a challenge at hand [10]. Knowledge integration is effectively achieved through virtual teams, e.g., teams in a COE, but requires team members to develop a transactive memory and mutual understanding.

It was also discovered through the project that one of the drivers of value creation in COEs was the ***level of trust*** between service companies, and between service companies and the operator. There existed different levels of trust, and one of the service companies (with longest partnership history) even got the task to perform quality assurance of work done by the operator and other service companies. Another interesting observation was that even though most service companies were competitors, they collaborated in problem solving as partners. It was said by one of the interview subjects from one of the service companies:

> "We are very glad when other service companies get their jobs done successfully, and the operator don't loose money. Therefore we help them out when they need."

It shows that integration has come a long way in this COE, and the level of implementation is moving toward Generation 2 (OLFs model of development, see Endrerud and Liyanage [7] for more information) where a strong collaborative network emerges involving the operator and service companies.

Another factor affecting performance is ***team structures***. In the even more complex and disperse operating environments, virtual teams emerge as an effective form of team structure. According to Thompson [11] and Ale Ebrahim et al. [12], the operational and business environment is so different in the modern world, much caused by novel and innovative communication and information sharing technologies. Only during the last decade, there has been an exploding upward trend for social media and communication such as Facebook, Twitter, LinkedIn that connect people across time zones and geographical location. Thompson [11] and Ale Ebrahim et al. [12] both agree it is rare to find organizational teams where everybody knows each other personally. Ahuja and Galvin [13] indicate that in 2001, in USA alone, some 8.4 million employees were members of one or more virtual teams or groups.

24.5 The Principal Improvement Potentials Identified in a Operator–Service Company Collaborative Setting

Findings of the study have been presented in another paper by Endrerud and Liyanage [7] and are shortly summarized here for practical reasons:

- Provider perception of service companies. A fundamental change in how operators perceive service companies, especially changing from provider to partner perception. This can change communication, and risk and gain sharing, and is very dependent on trust.
- Deficiency in relationships within onshore/offshore virtual teams. It has been identified that it is challenging to establish trustful relationships and mutual

understanding between onshore and offshore personnel through virtual communication in some cases.

- Lack of collaborative decision support tools to address coordination complexity in virtual teams and enhance decision quality and coordination.
- Few strong B2B ties between service companies, which can make COEs more autonomous and effective, in addition to enhancing continuous improvement and service quality.
- Inadequate knowledge integration in virtual teams has a performance gap, especially in locating a specific expert outside personal relationships quickly.

24.6 Some Measures as Improvement Solutions

With the identification of the performance gaps, the next obvious challenge was to derive suitable performance solution to overcome some difficult situations and to capitalize on the additional value adding potential. Some of the solutions are already in place within the development phases of several companies in the offshore petroleum sector in Norway. Over a long period, technology development has enhanced performance and value creation in operations of offshore assets. However, it became evident that additional measures can boost the quality and effectiveness of B2B relationships between operators and service companies working with COEs in an organizational context. Some of the main solutions resulted from the project are elaborated in the sections below.

24.6.1 Promote a Change of Perception of Business Partners

One of the proposals for changing the perception within a COE is to increase ***socialization efforts among service partners and operator***. To do this, common goals and ownership have to be created toward the COE and business processes, for example, through awareness campaigns, and joint training events. In particular, socialization between the onshore and offshore workers should have high priority, in addition to realizing a high level of ***technology adaption***.

Such social and training events could be done through virtual solutions for videoconferencing to include both offshore and onshore. For example, ***small learning events should be broadcasted on the platforms as well as in the premises onshore***. Another proposal is to foster the traditional “coffee corner chat”, also including the offshore “coffee corner”. A video screen and a camera could be put up in a coffee corner at the onshore facility and the same in the offshore coffee corner that broadcasts all the time. This might invoke social interactions between the offshore and onshore facilities and personnel in an informal setting and bring the two organizations together not only in problem solving and during operations but also in a pure social context.

Attitude campaigns should also be organized to help change operator's perception of service companies, as well as service companies' perception of themselves. In a provider perception situation, there exist knowledge and information barriers where flow of information and knowledge is restricted. Furthermore, service companies considered providers will not likely partake in risk sharing simply because they do not see the benefit. In addition to enhancing socialization among partners, ***incentive systems have to be changed*** to create a win–win situation and equal sharing of risks and benefits. For example, service companies can provide more functional services where the operator buys performance instead of a conventional product or support.

24.6.2 Create a Framework for Generating Personal Collaborative Relationships

In the previous paper, it was identified that it was difficult to establish valuable relationships between onshore and offshore personnel in virtual settings, for example, in a videoconference meeting. In order to accomplish valuable collaboration between functions, disciplines, and locations, trust and good relationships need to be developed. ***Coaching and virtual meeting training*** should be carried out across all collaborative partners from time to time to create a solid foundation for collaboration and an understanding of the benefits of virtual collaboration. Creating a coaching team as suggested by Mueller et al. [14] that can organize and follow up virtual teams can be one good solution to close the gap between traditional face-to-face and virtual interactions.

Another solution can be to have a ***rotation of personnel***, so that onshore support personnel can get the possibility to create a mutual understanding and contextual knowledge, as well as closer relationships with offshore personnel. Lastly, Toyota has managed to create an exceptional supplier network, where ***small and large events are held to foster experience transfer between suppliers and a chance to socialize and create personal relationships across disciplines, functions, and organizations*** [15]. Organizing such events for all collaborative partners, both onshore and offshore, should be initiated. Today, there exist "lunch and learn" events in the onshore offices in the studied COE, which can be a good starting point for expanding to include offshore and other distributed locations as well. Onshore personnel indicated that these "lunch and learn" events were valuable.

24.6.3 Develop and Integrate Collaborative Decision Support Tools

The previous paper emphasized the importance of developing collaborative decision support tools and integrated work processes in order to realize the potential in COE and enhance value creation. A flow chart of a decision process

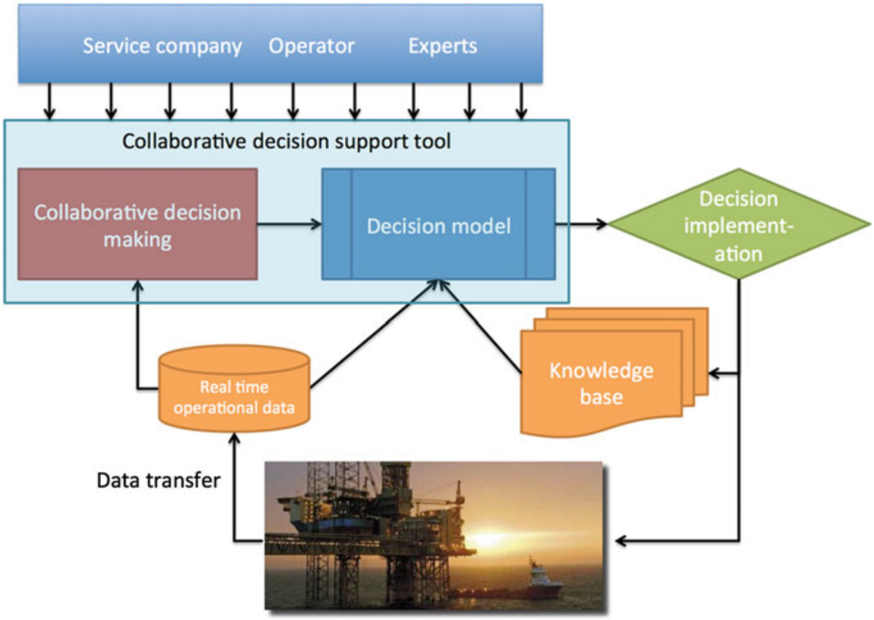


Fig. 24.1 Architecture of a collaborative decision support tool

with a collaborative decision support tool and real-time data from the offshore asset is shown in Fig. 24.1.

Collaborative applications can be a solution to several of the identified improvement potentials. *First*, a collaborative application can ***provide work process compliance*** by integrating roles, responsibilities, and processes in an ICT system that ***supports coordination and communication*** between different participants in a certain process [16]. This system has to be accessible by all partners of the COE to provide transparency and make it easy to follow the workflow at all times. Hence, it makes it easy for partners to know what to do next and who is responsible for doing it. When COE get large and complex, like the studied COE, integration of disciplines and functional areas is key for success, and an application solution that provides coordination of large work processes has a large value creating potential. For example, when planning a well, several functions and disciplines are involved. Well planners are responsible for making the well trajectory and plan, but need support from reservoir engineers responsible for production to optimize well location. Directional drillers might have some comments to optimize the well trajectory, and the drilling contractor knows about the limitations of their equipment. In other words, an integrated work process is extremely complex. Because experts and professionals are dispersed, an effective social network communication solution should be in place to make it easy to include the right people at the right time without delays.

Second, a collaborative application should include an ***expert system*** that can give advice on determining the best solution, in a combination between existing explicit knowledge from a knowledge base, and tacit knowledge from experts. A collaborative application like this will need to have a smooth interface between the different modules (decision support, communication, work process compliance, etc.)

For example, if a drilling team comes to a situation where fluid properties are developing in the wrong direction, they should be able to access previous events with similar characteristics and see what solution was selected then and what the outcome of that decision was. The team should be able to do problem solving in the application by communicating and sharing pictures, videos, audio, and other contextual knowledge. During the problem solving, team members should be able to plot different solutions into a decision model to get a recommended solution to implement. *Third*, it should be able to ***simulate and visualize the different solutions in real time*** as the problem solving session goes on.

24.6.4 Develop Stronger Ties Between Service Companies

Collaboration between operator and service companies in the studied COE is a frequent activity. However, collaboration between service companies happens much more infrequently. This represents a value creating potential. If service providers would collaborate thus accessing a wider network of distributed knowledge, value creating would positively benefit from it. Service companies are often strong competitors and are reluctant to help or collaborative with other service providers. *First*, to motivate service companies to share tacit knowledge and collaborate more in problem solving on their own initiative, one needs to ***create a common goal and a common purpose behind collaboration*** [15]. Maybe the operator has to force or strongly encourage service companies to collaborate in the start to make them become aware there are more benefits to be realized from sharing tacit knowledge and collaborate in problem solving, than safeguard their knowledge and keep it only to themselves.

Second, to ***make sure no "blind passengers"*** are getting the benefits in being a part of such a collaborative partnership with other service companies without contributing anything, a legally binding contract could be signed giving the operator the right to punish service companies that only exploits the collaborative partnership. Service companies would probably not share their product specifications, but can share their experiences and tacit knowledge.

24.6.5 *Enhance Knowledge Integration Between Collaborative Partners*

Knowledge integration has been identified as a key element of knowledge application in COEs. Knowledge integration is the synthesis of distributed experts' tacit knowledge into a situation-specific knowledge for the challenge at hand [10]. An important element of knowledge integration is to know who knows what—transactive memory—which relies on personal relationships and socialization. Another element was mutual understanding, or sharing the same language, terms, and having an idea about several disciplines without being a specialist. ***Training and learning events*** between organizations should be encouraged to create a common socializing arena where professionals can meet from all partner organizations to share experiences, challenges, plans, and knowledge. Another suggestion to enhance the transactive memory is to create a ***social network to establish weak inter-organizational ties***, hence make it easy to search for a specific specialist or expert and get a list of names within the collaborative partners that can be included in a virtual team to support problem solving. Applications like this already exist in many companies without being used actively, for example, in Microsoft Share Point Server.

24.7 Summary and Conclusion

There is a value creating potential in B2B collaborative relationships between operators and service companies on the oil and gas industry. However, realizing this value potential does not solely rely on pure technological innovations, but various other parameters that contribute to boost the performance from a cross-sectorial perspective. This particularly involves relations and interactions between stakeholders of producing assets. Several recommendations for future development were proposed, and even though addressing multiple issues, the most important recommendation is to create arenas for socialization among partners of the COE. Another recommendation was to develop and implement collaborative applications for decision support and work process compliance, which will be particularly important when complexity of coordinating the COE increases beyond human capabilities.

Furthermore, by capitalizing on the improvement potentials, one can achieve a higher utilization of capabilities of the COE and increase productivity and effectiveness, which can result in increased value creation in the end. COE is not a new concept but has been and is used by several industries, e.g., aviation, aeronautical, and military, and many of the solutions proposed in this paper have proven to enhance performance. Hence, transfer of experiences from other industries can provide solutions to challenges related to COE in the petroleum sector.

Incentive systems are an important element of COE, but have not been subject to thorough investigation. Hence, a recommendation for further work is to develop an incentive system for COE that creates a win–win situation for operator and service company.

References

1. KonKraft (2009) KonKraft report 2: production development on the Norwegian continental shelf, KonKraft, p 97
2. The Norwegian Oil and Energy Ministry (2011) White paper no. 28 on petroleum activity (2010–2011), The Norwegian Oil and Energy Ministry, Oslo
3. Liyanage JP (2012) Hybrid intelligence through business socialization and networking: managing complexities in the digital era. In: Carter E (ed) Handbook of research on business social networking: organizational, managerial, and technological dimensions. Business Science Reference, Hershey, PA, pp 567–582
4. Vindasius J (2008) The integrated collaboration environment as a platform for new ways of working: lesson learned from recent projects. In: SPE intelligent energy conference and exhibition 2008, Society of Petroleum Engineers (SPE), Amsterdam, The Netherlands, p 10
5. Liyanage JP, Herbert M, Harestad J (2007) Smart integrated eOperations for high-risk and technologically complex assets: operational networks and collaborative partnerships in the digital environment. In: Wang WYC, Heng MSH, Chau PYK (eds) Supply chain management: issues in the new era of collaboration and competition. Idea Group Inc, Hershey, PA, pp 387–414
6. Cohen D (1998) Knowing the drill: virtual teamwork at BP. In: Perspectives in business innovation, pp 14–19
7. Endrerud OE, Liyanage JP (2012) Value creation through collaboration in complex operating environments: challenges in the offshore petroleum sector. In: World congress on engineering asset management, Daejong, South Korea
8. Gereffi G, Humphrey J, Sturgeon T (2005) The governance of global value chains. Rev Int Polit Econ 12(1):27
9. Gulati R, Dialdin DA, Wang L (2002) Organizational networks. In: Baum JAC (ed) The Blackwell companion to organizations. Blackwell Business, Malden, Massachusetts, p XXXI 957 s
10. Alavi M, Tiwana A (2002) Knowledge integration in virtual teams: the potential role of KMS. J Am Soc Inform Sci Technol 53(12):1029–1037
11. Thompson K (2008) Bioteams: how to create high performance teams and virtual groups based on nature's most successful designs, 1st edn. Meghan-Kiffer Press, Tampa, FL
12. Ale Ebrahim N, Ahmed S, Taha Z (2009) Virtual teams: a literature review. Aust J Basic Appl Sci 3(3):16
13. Ahuja MK, Galvin JE (2001) Socialization in virtual groups. J Manage 29(2):1–25
14. Mueller K, McClelland C, Anvar A (2010) The role of human factors integration and change management coaching in integrating multidisciplinary teams working in collaborative work environments SPE 128616. In: SPE intelligent energy conference and exhibition. Society of Petroleum Engineers, Utrecht, The Netherlands
15. Dyer JH, Nobeoka K (2000) Creating and managing a high-performance knowledge-sharing network: the Toyota case. Strateg Manag J 21:345–367
16. Fjellheim RA, Bratvold RB, Herbert M (2008) CODIO: collaborative decision making in integrated operations. In: SPE intelligent energy conference and exhibition. Society of Petroleum Engineers (SPE), Amsterdam, The Netherlands

Chapter 25
Unlock the Value of Unstructured Data in EAM

Jing Gao and Andy Koronios

Abstract Gartner suggested that 80 % of organizational data is unstructured and has not been made available to the users. As a result, most organizations have far more data than they possibly use, yet at the same time, they do not have the quality data they really need. With the participation of an Australian Power Utility Company, this research demonstrated the value of unstructured data found in the plant incident reports (which are Word documents stored in the public directory of the exchange server) by conducting text analysis with special in-house-developed software.

25.1 Introduction

Data quality is a critical issue for effective information system management. Modern organizations, both public and private, are continually generating large volumes of data. According to Steenstrup from Gartner Research [15], each person on the planet generates an average of 250 MB of data per annum, and this volume is doubling each year. At an organizational level, there are incredibly large amounts of data, including structured and unstructured, enduring and temporal, content data and an increasing amount of structural and discovery metadata.

Most organizations have far more data than they possibly use, yet at the same time, they do not have the quality data they really need [8]. In many cases, data quality problems can result in severe negative consequences for an organization. Many information systems studies indicate that the majority of these contain substantial errors. For example, Klein et al. [6] indicate that mission-critical databases generally contain errors ranging from 1 to 10 %.

J. Gao (✉) · A. Koronios
University of South Australia, GPO Box 2471, Adelaide, 5001, Australia
e-mail: jing.gao@unisa.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_25,

Our previous nationwide survey [4] has indicated that over 60 % of Australian public and private organizations have at least three different types of data quality problems (e.g., accuracy, completeness, and consistency). The field of data quality and data management has become an important and vibrant research domain. Numerous researchers have attempted to define data quality and to identify its dimensions [7, 10, 16–19].

Traditionally, data quality has been described from the perspective of accuracy. However, most of the research has indicated that data quality should be defined beyond simple accuracy, and they have identified it as encompassing multiple dimensions. In literature, many authors have tried to explain the meaning of all relevant dimensions from several points of view [1, 3, 9, 18]. In this paper, the most common definition "fitness to use" is adopted.

Among all data quality problems, it is recognized that unstructured data such as e-mails, documents, and engineering drawings is a significant contributor to the data quality problems, and Gartner suggested that 80 % of organizational data is unstructured and has not made available to the users. As a result, most organizations have far more data than they possibly use, yet at the same time, they do not have the quality data they really need. With the participation of an Australian Power Utility Company, this research demonstrated the value of unstructured data found in the plant incident reports (which are Word documents stored in the public directory of the exchange server) by conducting text analysis with special in-house-developed software.

25.2 Research Domain: Engineering Asset Management

Assets are the lifeblood of most organizations. They may include digital assets, human assets, and financial assets. Most companies also have physical assets. These physical engineering assets (e.g., machinery, plant, and equipment) can be used to turn raw material into finished goods, supply electricity and energy, provide transportation services, or control huge utility operations. Many organizations rely heavily on these engineering assets to maintain and monitor daily operations. During the life cycle of these engineering assets, an enormous amount of data is produced. The data is captured, processed, and used in many computer information systems such as supervisory control and data acquisition (SCADA) systems, facility maintenance and management systems (FMMS), and geographic information systems (GIS), to name a few. It is therefore acknowledged by many researchers (e.g., [5, 12, 14]) that the study into asset management is worthy of investigation.

In general, there are two major types of engineering asset data (configuration data and transaction data). The first type of data (configuration data) is that associated with the physical asset attributes. Included in this data would be data such as acquisition date, acquisition cost, and physical location of the asset. The second type of data (transaction data) is data generated as a result of operation (or use) of the assets. This data can be self-generated (i.e., the asset has embedded sensors that can track when maintenance is necessary and has been completed) or manually generated (i.e., a service technician may perform routine maintenance checks, complete required activities, and record this data in a system separate from the asset itself).

In practice, both the configuration and transaction data can be captured automatically and manually and may involve sensors, field devices, human operators, field technicians, and contractors, in a variety of formats, processed in isolation and stored in an array of legacy systems. Data captured and processed by these systems is generally not comprehensive; it is usually process dependent, making it difficult to be reused for other processes or process innovation. For example, the data captured by the sensor may only be readable in the specially designed monitoring systems and cannot be exported and used for any other purposes.

According to Eerens [2], Spires [13] and IPWEA [5], the objective of asset management is to optimize the life cycle value of the physical assets by minimizing the long-term cost of owning, operating, maintaining, and replacing the asset, while ensuring the required level of reliable and uninterrupted delivery of quality service. At its core, asset management seeks to manage the facility's asset from before it is operationally activated until long after it has been deactivated. This is because, in addition to managing the present and active asset, asset management also addresses planning and historical requirements. The process of asset management is thus sophisticated and involves the whole asset life cycle that can span a long period of time [14].

The process of engineering asset management is sophisticated as shown in Fig. 25.1. The process itself is data centric—relying heavily on input data and simultaneously producing data. Following the principle of cause and effect ("garbage in, garbage out") and considering the large amount of data produced during the process, the quality of the data eventually becomes an issue.

Within the scope of AM life cycle management, a special focus must be placed on the field engineers and technicians (also known as the field force). The field force has been enhanced significantly with the emergence of the Internet and other mobile communication systems as well as the commercial utilization of GIS and GPS technologies. These technologies afford the organizations the opportunity to develop global integration systems which have the potential to enhance maintenance management systems. Although monitoring and diagnosis through remote technologies have developed quickly during recent times, it is still necessary to send maintenance staff out to the field to identify and fix faults as well as carry out periodic preventive or corrective maintenance.

An essential feature of effective asset maintenance is the ability of field force to have access to the currently undocumented practices that the asset management enterprise has accumulated over the years; these may include best practices, failures, and subsequent root causes incorrect diagnostic methods that should be avoided and other best practice organizational learning. At the same time, this information provides managers with important indicators on critical asset information such as wear-and-tear profiles, lifetime statistics, and operating efficiency

Fig. 25.1 Asset life cycle stages [11]

analyses, which provide for the identification of optimum maintenance strategy, process improvement structure, and centralized decision making.

Among the asset management life cycle, any incidents must be carefully dealt with in organizations. For example, a standard incident-reporting process is common in many utility companies. Engineers are required to investigate the issues and document all findings, solutions applied, and recommendations for the future. Usually, a semi-structured Word template is provided and the final document will be delivered to all relevant stakeholders (e.g., plant managers). POR reports are stored in the document management systems or in a public folder on the exchange e-mail server for later retrieval. However, over time, the number of incident reports accumulated, without being properly tagged (e.g., asset id), the incident documents became irretrievable to the field force. As a result, the critical historical data and information became unavailable to the field engineers for more effective and efficient problem solving.

25.3 Unstructured Data Problems

Many EAM organizations are using MS Word documents for incident reports. Examples of structures are explained in Sect. 25.4. However, this research has found a number of problems in relation to the Word-based incident reports:

- Unstructured document: Incident reports created in MS Word are often stored without proper tagging. For example, the file name of these incident documents may not contain any forms of the asset id, thus making it difficult to retrieve at later stages.
- Unstructured data/information within the document: Although many incident reports have simple templates as shown in Sect. 25.4, they are largely consisting of free-flow texts embedded in tables, dot points, and paragraphs.

Without performing text analyzing, it is difficult to grasp the context easily and quickly. Additionally, the rich information in the incident report has different types (e.g., a number can be asset serial id or work orders).

It must be pointed out that the incident report problem is just an example of unstructured data problems in EAM organizations. Drawings, manuals, policies, and procedure documents are stored in similar ways. By studying the incident report problems, learning can be useful in improving other areas.

25.4 Research Design

This research was conducted with an Australian Power Utility Company. The research consists of three aspects:

1. Investigate the current unstructured incident reports in Word documents and understand the data and information requirements.
2. Provide a proof-of-concept software tool to automatically parse the documents in order to provide searchable content.
3. Conduct data profiling on extracted data and information for more informed decision making.

It must be noted that one of the most valuable research outcomes is to obtain an insightful understanding on the difficulties of dealing with unstructured data in order to improve data and information quality in organizations. The proof-of-concept software was not intended to be a solution to the addressed problem. Instead, it provides useful learning toward a proper and complete solution.

A number of incident reports were collected from the participating organization for more than 10-year period. It is noted that the Word template for the incident reports has been changed in 2003. Prior to year 2003, it was called incident reports, and after which, it was renamed to plant outage reports. The data and information contained changed slightly due to the new compliance requirements. The old incident report has a simpler structure (Fig. 25.2).

The relatively new plant outrage report has information on the investigation team, safety issues as well as attachment section where photographs and drawings could be included to describe the incident situation (Fig. 25.3).

It was also determined that these incident reports are currently stored as an attachment in a public directory of the corporate exchange server. The incident reports were distributed monthly to all plant managers. Some plant managers may have manually saved the incident reports on the team shared drive, but others may not. Thus, the only reliable and complete source is the exchange server.

Incident Report

text removed

Report No.:		
Incident Type:	Safety:	No
	Environment:	No
	Plant:	Yes
PSAVAIL No.:		
EMIS No.:		
Date of Incident:	24/6/2000	
Plant Area:		
Production Lost:		

Incident

Text removed

Evidence

Text removed

Findings

Text removed

Recommendations

	Action by	WO Number
1. Issue Work Order ...	Shift Manager	12345 (completed)

Report completed by:

Text removed

Fig. 25.2 Old POR report template

25.5 Research Findings

In order to extract useful contents from the semi-structured POR Word documents, the proof-of-concept software has been developed as shown below (Fig. 25.4).

In addition to the XML output as shown above, the proof-of-concept software also generates the HTML and MS Access view (Fig. 25.5).

During the process, a number of issues were identified as summarized below. This has also demonstrated various difficulties in dealing with unstructured data in organizations, which are considered as most useful research outcome from this study.

1. Access Policy: As discussed above, the incident reports are currently stored in the corporate exchange server. A special account is required to access the public directory. Although it is technically straightforward, the IT governance policy currently in place in the organization requires the research team applying for special permission (as the intended account use is not a general-purpose one).

POR No.:		
REPORT TITLE:	"A" EFP Overheated	
UNIT:		
DATE & TIME:		
MW LIMITATION:	0	
MWh LOSS:		
INCIDENT:	Safety:	No
	Environment:	No
	Plant:	Yes
	NEL:	No (Check Status Bar below or press F1)
FAILURE COSTS ($):	Assets:	
	Production:	
ASSIGNED TO:		

Brief Description of Incident
Text Removed

Investigation team:

Text removed

Evidence

Text removed

Findings

Text removed

Recommendations

		Action by	WO Number	Estimated Complete Date
1.	Inspect ...	Plant Owner	123456789	
2.				

Attachments:

Fig. 25.3 New POR report template

2. Access Techniques: The Microsoft Exchange Server has provided sufficient APIs for software tools to retrieve e-mails automatically (including attachments). However, it was not optimized for performance efficiency especially for large batch jobs (thousands of e-mails and attachments). Each attachment may take approximately 3 to 5 s to download. (It appears that the exchange server also deliberately slows the downloading in order not to create deadlocks on the files). Thus, the proof-of-concept software is only able to run after hours in order not to interfere with the normal network traffic.
3. Incremental updating: Since each batch downloading took a long time, an incremental approach was used to extract and update the new incident e-mails and extract the useful data and information to the output database. This appears to be a more acceptable approach as it will not constantly increase the network traffic and system loads.

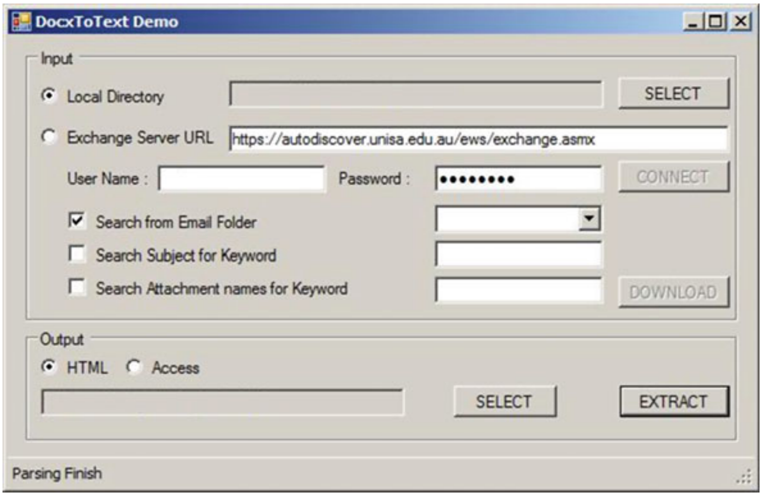


Fig. 25.4 Data extraction software

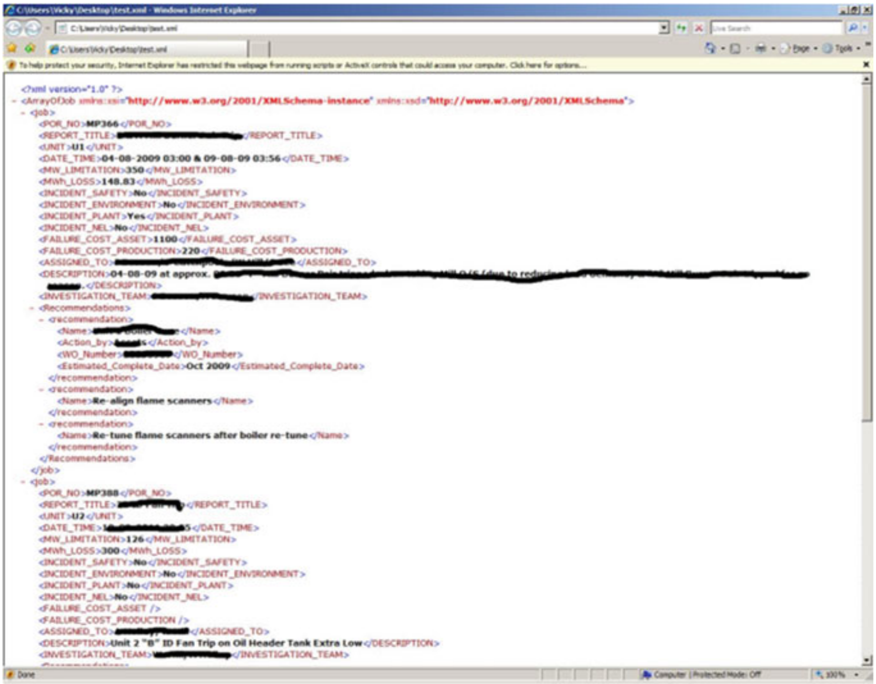


Fig. 25.5 XML output

4. Format conversion: In addition to the different semi-structured templates, over the years, the actual MS Word file format has also been changed. It is found that RTF, DOC, and DOCX are all used in the incident reports. Thus, the proof-of-concept software has to detect the format and use internal conversion tool for text extraction.
5. Duplication and Naming Standard: Although it is not common, it is found that different e-mails contained same incident reports (because that incident reports were manually attached and distributed via a person). To make the situation worse, different incident reports may use the same file name (e.g., the asset name) as there is no naming standard enforced.
6. Varity of data and information types: In addition to the free-flow text, the new incident reports also contained the drawings and photographs of the assets (e.g., before and after the incidents). Currently, it is possible to extract and save these data and information to an external sources with hyperlinks amended to the incident output database. It is, however, difficult to match these photographs and drawings to any other sources (e.g., the corporate document management system) as there are not file name information with these graphic contents unless the OCR is performed. Even then, OCR software is only limited to the text information which is found limited use in handling engineering drawings.
7. Future use and integration: It is valuable to access to the data and information from the incident reports. However, integrating the data to other systems (structured and semi-structured) will contribute significantly to the current maintenance practices. For example, by retrieving the work-order details from the maintenance systems, the engineer will be able to obtain a comprehensive view of all historical jobs performed on this particular asset. Thus, the proposed software produced XML output on default for other integration/cross-checking purpose.
8. Data Quality Issues: The proof-of-concept software is able to extract the text contents into a structured access database with XML and HTML outputs. Unfortunately, a number of data quality problems were discovered. For example.
    (a) Missing Data: Many fields are empty such as MwH loss, missing work-order number in the recommendation section (each recommendation should be associated with a work-order number).
    (b) Wrong Values: For example, the safety textbox should contain value "Yes" or "No". It actually contains "Yes", "Y", "No", "N", "T", "F", and many other varieties.
    (c) Wrong format: The work-order field (in the recommendation section) should only contain numbers. However, some people decided to put (12345Comp) to indicate that the job was completed. It was perhaps very useful to the human readers, but made difficult to link the work order to the actual maintenance system for further information.
    (d) Timeliness: Many incident reports were completed weeks, if not months after the actual event.

25.6 Summary and Conclusion

The proposed prototype system has demonstrated a possible approach to unlock the value from the unstructured data sources in organization. It has also demonstrated the possibility of using these data and information to improve data quality which will result in more effective and efficient asset management in organizations. It is believed that more informed decisions usually came from better information and knowledge management.

As presented, any approach of dealing with unstructured data is more difficult as compared to the approach for structured data sources. During this study, a number of technique and organization issues were identified and summarized. In particular, it is also proven that despite the technical capability, the underlying data quality should always be regarded as a major issue. Without quality information, the extracted content may present limited value to the end-users. Further, with these findings, the organization is recommended to consider a better approach of incident reporting in the future. Perhaps, a Web-based form system may be a better answer.

Overall, by integrating technology, business processes, workflows of asset management, and knowledge management practices, the organization will no doubt improve data and information quality management, resulting in more efficient and effective asset management practices.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Ballou DP, Pazer HL (1995) Designing information systems to optimize the accuracy-timeliness tradeoff. Inf Syst Res 6(1):51–72
2. Eerens E (2003) Business driven asset management for industrial & infrastructure assets Australia. Le Clochard
3. English LP (1999) Improving data warehouse and business information quality: methods for reducing costs and increasing profits. Willey, NY
4. Gao J, Koronios A, Lin S (2006) Data quality in engineering asset management organisations—Current picture in Australia. In: The international conference of information quality, Boston
5. IPWEA (2002) International infrastructure management manual Australia/New Zealand Edition
6. Klein BD, Goodhue DL, Davis GB (1997) Can humans detect errors in data? Impact Base Rates Incentives Goals MIS Q 21(2):169–194
7. Kriebel CH (1979) Evaluating the quality of information systems. In: Szysperski N, Grochla E (eds) Design and implementation of computer based information systems, Sijthroff and Noordhoff, Germantown

8. Levitan AV, Redman TC (1998) Data as a resource: properties implications and prescriptions. Sloan Manag Rev 40(1):89–101
9. Orr K (1998) Data quality and system theory. Commun ACM 41(2):66–71
10. Shanks G, Darke P (1998) Understanding data quality in a data warehouse. Aust Comput J 30(4):122–128
11. Snitkin S (2003) Collaborative asset lifecycle management vision and strategies. Research Report ARC Advisory Group, Dedham
12. Sokianos N, Druke H, Toutatoui C (1998) Lexikon productions management. Landsberg
13. Spires C (1996) Asset and maintenance management—becoming a boardroom issue. Managing Serv Q 6(3):13–15
14. Steed JC (1988) Aspects of how asset management can be influenced by modern condition monitoring and information management systems. IEEE
15. Steenstrup K (2005) Enterprise asset management thrives on data consistency. Gartner Research Article
16. Strong DM (1997) IT process designs for Improving Information Quality and reducing exception handling: a simulation experiment. Inf Manag 31:251–263
17. Strong DM, Lee YW, Wang RY (1997) Data quality in context. Commun ACM 40:103–110
18. Wang RY, Kon HB (1993) Toward total data quality management (TDQM). In: Wang RY (eds) Information technology in action: trends and perspectives. Prentice Hall, Englewood Cliffs
19. Wang RY, Strong DM (1996) Beyond accuracy: what data quality means to data consumers. J Manage Inf Syst 12(4):5–33

Chapter 26
Non-parametric Bootstrap Confidence Limits for Age-Dependent Failure Tendency Using Incomplete Data

Per-Erik Hagmark and Jouko Laitinen

Abstract This paper addresses modeling of age-dependent failure tendency. The field data of an aircraft device record the follow-up start and end of the units and the failures in chronological order. The data are incomplete since they lack the exact age of the unit at some failures in a way that seems to be random. We construct a non-parametric point estimate of the nonlinear cumulative failure rate, and we model the failure point process behind. Firstly, we point out how effectively the quantiles of the distance to next failure capture and interpret the information in real data. Further, since the characteristics of the actual field data are recognizable enough, it is possible to generate artificial (bootstrap) data sets that imitate the field data set. Each simulated data set determines a nonparametric estimate of the cumulative failure rate, and these estimates together provide fully nonparametric bootstrap confidence bounds for the cumulative failure rate. For practical purposes, we carry out some computational demonstrations on how the data set size and the censoring rate affect the confidence bounds.

26.1 Introduction

The object under study can be any clearly defined device or equipment which is intended to execute a certain task. A *failure* is an unwelcome event that interrupts the task execution causing unavailability and extra costs. *Age* means cumulated execution time or some other suitable measure of usage, so age does not include downtime for repair or maintenance or other stops. From the point of view of this study, a failure data set originating from several units of the same kind is qualitatively *complete* if it meets the following conditions:

P.-E. Hagmark · J. Laitinen (✉)
Tampere University of Technology, Korkeakoulunkatu 10, P.O. Box 527,
33101 Tampere, Finland
e-mail: jouko.laitinen@tut.fi

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_26,

1. The data include the age of every unit at the start and the end of its follow-up(s).
2. The follow-up intervals of the units together cover the age interval to be studied.
3. For each unit, the exact age at every failure has been recorded so the data are not censored.

This paper considers failure data stored in a database during the lifespan of 142 units of an aircraft device, an air turbine starter control valve (ATSCV). The age of a unit is measured in cumulative flight hours. Missing, incorrect, or inaccurate failure data are a rather general problem, and especially for small aircraft fleets because the sample sizes tend to be small. The ATSCV field data are slightly incomplete; it fulfills (a) and (b) but not (c). More exactly, we know all failures of each unit in chronological order, but the *exact* age of the unit is missing for some failures in a way that seems to be random. Even though the collecting system does not compel exact recording of ages at failures, it was possible to determine 93.9 % of them in some way or another. All other failures are localized between two known age points. Thus, 6.1 % of the failures are interval censored. Table 26.1 lists some typical follow-ups.

The motivation for this study comes from a demand for better maintenance policy to increase availability and reduce operation costs. To this purpose, the most indispensable modeling object is the cumulative failure rate CFR(x) as a function of the unit's age x. More exactly, CFR(x) equals the expected number of failures in the open–closed age interval $(0, x]$, $x \geq 0$. In simple modeling, the CFR is often assumed linear, and if the exact age at a failure is lost, the service sojourns on both sides of the unknown failure age are normally skipped. This means loss of information and rough simplification. The failure tendency of a device is nearly always *age-dependent*; that is, the distance between failures depends on the current age of the unit, so a nonlinear CFR model is required. In parametric estimation of the CFR, the two-parameter power law model is popular, but also nonparametric estimates have been proposed, especially for complete data; see, e.g., [1, 2].

We showed in [3] how incomplete data of the type described above can be taken into account properly, and we presented a nonparametric estimate and a three-parameter maximum likelihood estimate. In this study, we employ the nonparametric estimate, although the maximum likelihood setting could be used as well. First, we use it to model the distance to next failure as a function of the current age. The domains of the related quantiles are restricted since field data always support a finite interval only. We make clear how probabilistic information about failing can effectively be obtained from quantiles.

Further, the characteristics of our ATSCV field data are quite recognizable, so we can simulate a big collection of analogous bootstrap data sets using the nonparametric estimate. We generate the failure age points in the follow-up intervals of the real data. Then, we censor the right amount (6.1 %) of the exact ages randomly. Each bootstrap data set generated this way yields a distinct nonparametric estimate, and these together can be used to obtain approximate

Table 26.1 The age points in four follow-ups selected from the ATSCV field data (in total 142 realizations)

Unit	Start	1st fail	2nd fail	3rd fail	4th fail	5th fail	6th fail
0009	0	?	363	583	1,230*	–	–
0023	0	113	393	394	515	856*	–
0078	0	?	377	569	?	1,251*	–
0136	0	?	?	367	?	691	1,490*

Notation The symbol '*' denotes the end of follow-up, while '?' indicates a censored failure age

nonparametric confidence bounds for the real CFR. Thereafter, we experiment with data set size and amount of censoring to see the effect on the confidence interval.

We close the paper with a short discussion.

26.2 Concepts and Terminology

26.2.1 A Model for Age-Dependent Failing

Different failure point processes can lead to the same CFR. Therefore, it is often necessary, in stochastic failure modeling, to go behind the CFR. Our specific need is a comfortable model of failure point processes, which produces CFRs of any possible shapes. Let $\Lambda(x)$, $0 \le x < \omega$, be a continuous and strictly increasing function such that $\Lambda(0) = 0, lim_{x\to\omega}\Lambda(\mathrm{x}) = \infty$, where $\omega \le \infty$ is a constant. Let $V(y)$, $0 \le y < \infty$, denote the inverse function of $\Lambda(x)$. We consider a one-parameter cumulative distribution and the related quantile function:

$$G_\alpha(x) = 1 - e^{\Lambda(\alpha)-\Lambda(x)}, \quad 0 \le \alpha \le x < \omega, \tag{26.1}$$

$$G_\alpha^{-1}(u) = V(\Lambda(\alpha) - \ln(1-u)), \quad 0 \le u < 1. \tag{26.2}$$

The interpretation of Eqs. 26.1 and 26.2 will be as follows. If the *current* age of the unit is α (parameter), we assume that the age at the *next* failure obeys the cumulative distribution $G_\alpha(x)$. In simulation terms, the age at the next failure is $X_1 = V(\Lambda(\alpha) - \ln(U_1))$, where U_1 is a random variate from the uniform distribution on the unit interval. Repetition defines a failure point process $X_0, X_1, X_2, \ldots$ as follows: $X_0 = \alpha$, and

$$X_i = V(\Lambda(X_{i-1}) - \ln(U_i)) \text{ or } \Lambda(X_i) = \Lambda(X_{i-1}) - \ln(U_i)\; i = 1, 2\ldots, \tag{26.3}$$

where the uniform variates U_i are independent. The right-hand version of Eq. 26.3 expresses that $Y_i = \Lambda(X_i)$ is a renewal process, where the steps, $-\ln(U_i)$, are independent and exponentially distributed with mean 1. In other words, the point

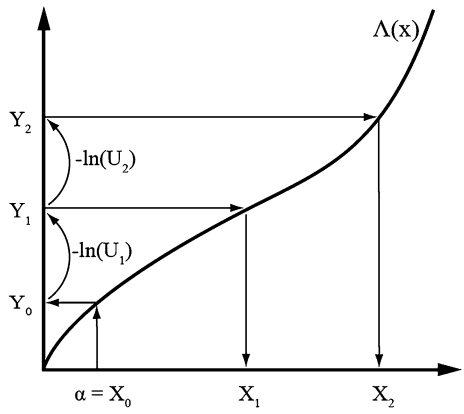


Fig. 26.1 Model of the failure point process

process Y_i is a homogeneous Poisson process (HPP). This means that the failure point process X_i is a non-homogenous Poisson process (NHPP), the CFR of which automatically becomes $\Lambda(x)$, [4, 5]. Figure 26.1 is an illustration.

We summarize. The process X_i generates the correct average number of failures in every age interval. The distance to next failure depends on the current age α, but *not* on what has happened in the open–closed interval $(0, \alpha]$. For example, if a unit has failed at age α, it behaves after 'repair' statistically in the same way as a non-failed unit of the same type and age. This feature is frequently described with slightly inexact phrases like 'as good as old' or 'minimal repair.'

26.2.2 Parameterization of Incomplete Failure Data

We portray field data that fit the model in Sect. 26.2.1. The data of a unit come from one or several follow-up intervals, and follow-up intervals of different units can overlap. A follow-up interval of a unit is of course uniquely represented by the related age interval $\alpha \ldots \beta$. Let $x_1, x_2 \ldots x_n$ denote the *non-censored* (exactly known) failure age points in the open interval (α, β); if there are no non-censored failures in (α, β), put $n = 0$. Denoting $x_0 = \alpha$, $x_{n+1} = \beta$, we have the follow-ups in the form

$$(\alpha, \beta] = \bigcup_{i=0}^{n} (x_i, x_{i+1}]. \quad (26.4)$$

Let $(a, b]$ be any subinterval $(x_i, x_{i+1}]$ of a follow-up of a unit from Eq. 26.4. The open interval (a, b) can also contain *censored* failures (of the same unit), i.e., failures where the age of the unit is unknown; we assume that the number of these failures, $k \geq 0$, is known. The last piece of information in the subinterval $(a, b]$ remains: The *type* $c = 1$ if there was a failure at b, and otherwise, $c = 0$. Let us

Table 26.2 Sample bouquets from the ATSCV field data, which contains 403 bouquets (columns) in total

a (start)	0	860	…	0	377	569	…	0	…	0
k (censored)	0	0	…	1	0	1	…	2	…	2
b (end)	860	1,535	…	377	569	1,251	…	367	…	1,777
c (type)	1	0	…	1	1	0	…	1	…	1

call the four-tuple (a, k, b, c) a *bouquet*. Thus, the bouquets parameterize the information related to the subintervals $(x_i, x_{i+1}]$ of the follow-ups Eq. 26.4. Table 26.2 lists a few bouquets from the ATSCV data in Table 26.1.

26.2.3 Non-parametric Point Estimation of the Cumulative Failure Rate

Let $(0, L]$ be an age interval covered by the union of all follow-ups, let $0 = \xi_0, \xi_1, \xi_2 \ldots \xi_m = L$ be an equidistant division, and let (a_j, k_j, b_j, c_j) be an indexing of the bouquets. The *total number of failures* and the related *total length of follow-ups* in a subinterval $(\xi_{v-1}, \xi_v]$ will be approximated by

$$S_v := \sum_j \left(\mathrm{Lap}\left(a_j, b_j, \xi_{v-1}, \xi_v\right) \frac{k_j}{b_j - a_j} + \left(\xi_{v-1} < b_j \leq \xi_v\right) c_j \right), \tag{26.5}$$

$$H_v := \sum_j \mathrm{Lap}\left(a_j, b_j, \xi_{v-1}, \xi_v\right), v = 1, 2 \ldots m. \tag{26.6}$$

Here, $\mathrm{Lap}(a, b, \xi, \varsigma)$ is the length of the interval $(a, b) \cap (\xi, \varsigma)$, and the (*inequality*) in Eq. 26.5 equals 0 or 1 according to whether the inequality is false or true. The former term in Eq. 26.5 is based on the *assumption* that the k_j censored failure ages in (a_j, b_j) are uniformly distributed, and the latter term takes care of the possibility of failure at b_j (i.e. $c_j = 1$).

The quotient $\lambda_v := S_v/H_v$ approximates the failure density per unit in the age interval $(\xi_{v-1}, \xi_v]$, so the *number of failures per unit in the age interval* $(0, \xi_v]$ sums to

$$y_v := \frac{L}{m} \sum_{r=1}^{v} \lambda_r, \; v = 1, 2 \ldots m. \tag{26.7}$$

We can now define the nonparametric estimate (NPE) of the CFR according to [3]: This is the interpolation function $\Lambda_0(x)$ through the points (ξ_v, y_v); that is, $\Lambda_0(\xi_v) = y_v$, and $\Lambda_0(x)$ is linear between ξ_{v-1} and ξ_v. Similarly, the inverse function $V_0(y)$ is the interpolation function through the points (y_v, ξ_v). To be exact, the array y_v is non-decreasing by definition Eq. 26.7, but for computational fluency, it should be strictly increasing; the modification done is of course slight enough not to deform the results numerically.

26.3 Modeling and Analysis of Failing

26.3.1 Estimation of the CFR of the ATSCV

The follow-ups of the units in the ATSCV field data (Table 26.1) cover the interval (0, 2,174], so we begin with the equidistant partition $0 = \xi_0, \xi_1, \xi_2 \ldots \xi_{100} = 2{,}174$. However, we have a nonparametric setting, and there are no failures after $L = \xi_{82} = 82 \times 2{,}174/100 \approx 1{,}782.7$, so it is not either trustworthy to model beyond that point. Thus, we use the shortened partition $0 = \xi_0, \xi_1, \xi_2 \ldots \xi_m = L$, where $m = 82$, and we form the NPE $\Lambda_0(x)$ and its inverse function $V_0(y)$ in line with Sect. 26.2.3. The result is in Fig. 26.2; for example, the expected number of failures in the age interval (0, 800] is estimated by $\Lambda_0(800) = 1.509$.

26.3.2 Knowledge About the Next Failure

If the current age of an ATSCV unit is α, then based on the process model in Sect. 26.2.1 and the field data, what can be said about the *distance to next failure,* T_α? According to Eqs. 26.1 and 26.2, the random variate T_α and the related quantile function (inverse of the cumulative distribution) are

$$T_\alpha = V_0(\Lambda_0(\alpha) - \ln(U)) - \alpha, \tag{26.8}$$

$$T_\alpha(u) = V_0(\Lambda_0(\alpha) - \ln(1-u)) - \alpha. \tag{26.9}$$

As mentioned above, the field data offer no sensible ground for assessing the nonparametric $\Lambda_0(x)$ after $L \approx 1{,}782.7$. Thus, the distribution of the variate T_α is available in $[0, L-\alpha]$ only, and $0 < \alpha < L$. Exactly expressed, we have the equivalent inequalities

$$\alpha + T_\alpha(u) \le L,\ \alpha \le V_0(\Lambda_0(L) + \ln(1-u)),\ u \le 1 - e^{\Lambda_0(\alpha) - \Lambda_0(L)}. \tag{26.10}$$

The inequalities Eq. 26.10 can be applied and interpreted from different points of view. Figure 26.3 depicts three quantiles as functions of the current age α. The 0.75-quantile $T_\alpha(0.75)$ is available for $\alpha \le V_0(\Lambda_0(L) + \ln(1-0.75)) = 424.3$, the median $T_\alpha(0.5)$ for $\alpha \le 812.0$, and the 0.25-quantile $T_\alpha(0.25)$ for $\alpha \le 1{,}188.5$. Further, let the current age of a unit be $\alpha = 740$, say. Then, no trustworthy information about the 0.75 quantile is available, the median to next failure is

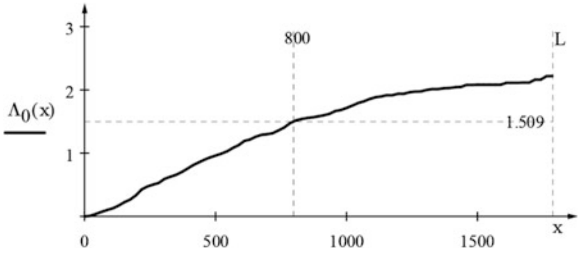


Fig. 26.2 Non-parametric approximation of the CFR of the ATSCV device

$T_{740}(0.5) = 700.7$, and there is a probability of 0.25 that there will be a failure before the age $740 + T_{740}(0.25) = 740 + 207.8 = 947.8$. Clearly, if a more exact probabilistic analysis is needed, a denser swarm of quantiles can be applied, e.g., $u = 0.1, 0.15, 0.2 \ldots 0.95$.

26.3.3 Approximate Bootstrap Confidence Bounds for the CFR

The main target is to construct approximate *confidence bounds* for the CFR of the ATSCV. Let D_0 denote the field data set (Table 26.1). From the bouquet version of D_0 (Table 26.2), we count to $\Sigma c = 263$ non-censored failures and $\Sigma k = 17$ censored failures, so the fraction of censored failure ages in D_0 is $p = \Sigma k/(\Sigma c + \Sigma k) \approx 0.061$. Using p and the NPE $\Lambda_0(x)$, Fig. 26.2, we can simulate a collection of bootstrap data sets D_s ($s = 1 \ldots N$) that imitate D_0 statistically. Each D_s leads to a NPE $\Lambda_s(x)$, and these together provide confidence bounds for the real CFR. The following step-by-step program declares the details:

1. Let the follow-up intervals $\alpha \ldots \beta$ in D_s be exactly those of D_0 (where $\alpha < L$).
2. In each follow-up interval $\alpha \ldots \beta$ in D_s, the stochastic failure process starts from $X_0 = \alpha$ and continues by repetition of $X_i = V_0(\Lambda_0(X_{i-1}) - \ln(U_i))$ until $X_i \leq \min\{L, \beta\} < X_{i+1}$ (Sect. 26.2.1).
3. Each failure age X simulated in Step 2 is censored with probability p, but its existence and chronological position remain known. Construction of D_s is now complete.
4. D_s is transformed to bouquet form (Sect. 26.2.2).
5. NPE $\Lambda_s(x)$ is derived from D_s in the same way as $\Lambda_0(x)$ was obtained from D_0. $\Lambda_s(x)$ has the same partition $0 = \xi_0, \xi_1, \xi_2 \ldots \xi_m = L$ as $\Lambda_0(x)$ (Sect. 26.2.3).
6. When Steps 1–5 have been done for all $s = 1, 2 \ldots N$, go to Step 7.
7. For each $0 < x < L$, let $\Psi_1(x), \Psi_2(x) \ldots \Psi_N(x)$ denote the order statistics of $\Lambda_1(x), \Lambda_2(x) \ldots \Lambda_N(x)$.
8. Choose the confidence level $0 < C < 1$ and define the approximate confidence bounds for the real CFR,

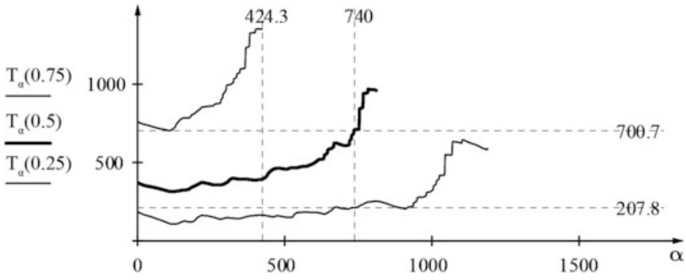


Fig. 26.3 The quantiles $u = 0.25$, 0.5, and 0.75 of the distance to next failure

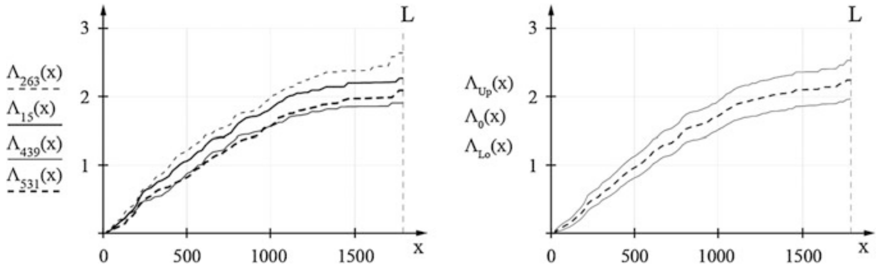


Fig. 26.4 Four bootstrap NPEs $\Lambda_s(x)$ and confidence bounds $\Lambda_{Lo}(x)$, $\Lambda_{Up}(x)$, for $p = 0.061$ and $C = 0.95$

$$\Lambda_{Lo}(x) := \Psi_{\text{Round}(N(1-C)/2)}(x) \;\&\; \Lambda_{Up}(x) := \Psi_{\text{Round}(N(1+C)/2)}(x) \tag{26.11}$$

where *Round* denotes the nearest integer. Figure 26.4 presents results for censoring probability $p = 0.061$ (original rate), confidence level $C = 0.95$, and number of bootstrap sets $N = 600$. (A bigger N would give smoother curves.)

26.3.4 Experimentation

The bootstrap program 1–8 above can be used in experimentation for many purposes. We consider two practically relevant questions concerning the confidence bounds, $\Lambda_{Lo}(x)$ and $\Lambda_{Up}(x)$.

1. How does the data set size affect the confidence interval $\Lambda_{Up}(x) - \Lambda_{Lo}(x)$? First, we ran the program 1–8 with the 142 follow-up intervals of D_0 and with $p = 0$ (non-censored data). As the other alternative, we ran 71 randomly chosen follow-up intervals from D_0 (without returns). As expected, the confidence interval is narrower for bigger data sets, Fig. 26.5. One can of course do the same with any censoring rate $0 < p \leq 1$.

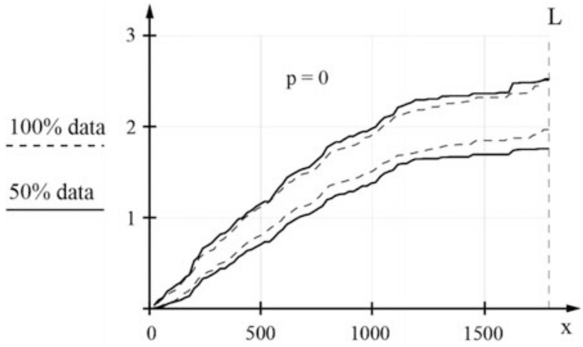


Fig. 26.5 The effect of data set size

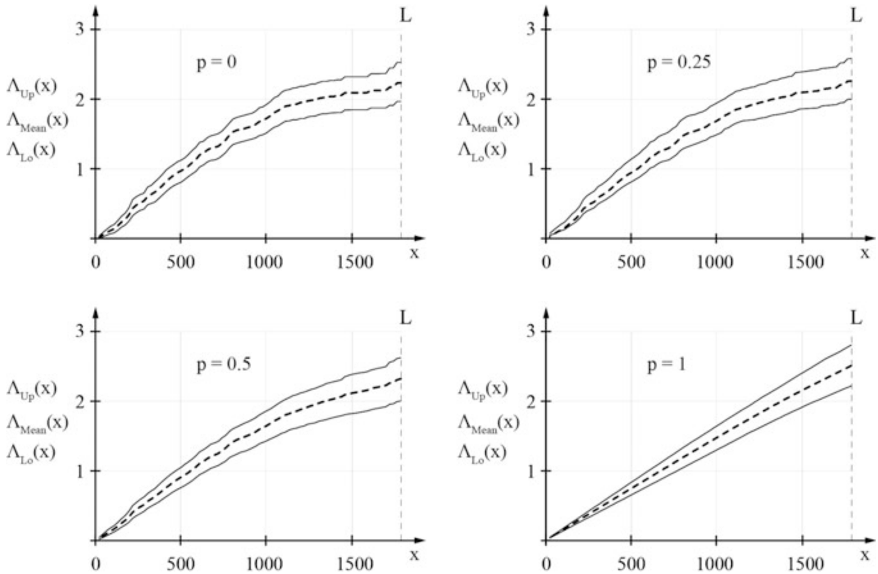


Fig. 26.6 Confidence bounds for different censoring probabilities [Here, $\Lambda_{\text{Mean}}(x)$ is the mean of the generated swarm of NPEs, $\Lambda_s(x)$]

2. How does the censoring rate p affect the confidence bounds? We ran the program 1–8 with four p values, 0, 0.25, 0.5, and 1 (Step 3), and we found that p does *not* affect the confidence interval $\Lambda_{Up}(x) - \Lambda_{Lo}(x)$ essentially, Fig. 26.6! Further, just as with the average NPE point estimate (in [3]), the shapes of the bounds are rather immune against slight and moderate censoring ($p < 0.25$). Bigger p values make the shapes more linear; this is an unmistakable consequence of the definition of NPE, which assumes that censored failure ages are uniformly distributed in their subintervals (Sect. 26.2.3).

26.4 Conclusions

The ATSCV starter valve is an extremely important device. It directly affects the availability of the aircraft, because there is usually no standby system. Malfunction immediately prevents takeoff, and if the plane is airborne, a need to restart the engine can unexpectedly lead to a fatal accident. For devices which have such an impact, the analyst should be clearly aware of the requirements to be put on the data.

When making decisions about the maintenance strategy for devices, a crucial issue is to build point estimates and confidence bounds for the failure tendency as a function of age. Schematically, the confidence gap suffers if insufficient data are available, while the shape (age dependency) suffers if too many failures are interval censored. The methods presented in this paper and in [3] can be used to experimentally determine how much data are needed and how much censorship can be tolerated in order to satisfy stipulated reliability requirements. Prediction of the time to next failure is also briefly discussed, based on the failure rate estimate.

References

1. Anacleto A, Louzada F (2012) Bootstrap confidence intervals for industrial recurrent event data. Pesquisa Operacional 32(1):103–119
2. Arkin BL, Leemis LM (2000) Nonparametric estimation of the cumulative intensity function for a non-homogeneous poisson process from overlapping realizations. Manage Sci 46(7):989–998
3. Hagmark PE, Laitinen J (2011) Modeling of age-dependent failure tendency from incomplete data. Sixth World Congress on Engineering Asset Management (WCEAM), Cincinnati, OH, USA, 2–4 Oct 2011
4. Pasupathy R (2011) Generating non-homogeneous poisson processes. In wiley encyclopedia of operations research and management science. DOI: 10.1002/9780470400531.eorms0356
5. Ross SM (1996) Stochastic processes, 2nd edn. Wiley, USA

Chapter 27
Asset Lifecycle Data Governance Framework

Abrar Haider

Abstract Engineering asset lifecycle management is information intensive. The variety of asset lifecycle processes generate, process, and analyze enormous amount of data on daily basis. Information systems utilized for asset management not only have to provide for the control of lifecycle management tasks, but also have to act as instruments for decision support. Asset lifecycle management can be viewed as a combination of decisions associated with strategic, planning, and operational levels of the organization. Information systems in asset management, in theory, thus facilitate data-enabled view of asset management. However, realization of such a view of asset lifecycle through information systems requires appropriate hardware and software applications; quality, standardized, and interoperable data; appropriate skill set of employees to process data; and the strategic fit between the asset lifecycle management processes data requirements and the information systems. This paper sketches out a framework for asset lifecycle management data governance, which highlights the rights and accountabilities related to asset data lifecycle management. The framework describes how common business data and metrics should be defined, propagated, owned, and enforced throughout the organization, thereby allowing for better quality and faster decision-making, business intelligence reporting, cost reductions, compliance, and better controls of business processes.

27.1 Introduction

Asset management is an information intensive paradigm that generates, processes, and analyses enormous amount of information on daily basis. Information systems utilized for asset management not only have to provide for the control of lifecycle

A. Haider (✉)
Mawson Lakes Campus, University of South Australia, Adelaide, SA 5095, Australia
e-mail: abrar.haider@unisa.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_27,

management tasks, but also have to act as instruments for decision support. However, current information systems in operation within engineering enterprises have outlived their productive lives, as the methodologies employed to design these have evolved over considerable time. As a result, these systems are not forward compatible [1]. For example, the maintenance information system development that has attracted considerable attention in research and practices is far from being optimal. While maintenance activities have been carried out ever since advent of manufacturing; modeling of an all inclusive and efficient maintenance system has yet to come to fruition. This is mainly due to the continuously changing and increasing complexity of asset equipment, and the stochastic nature or the unpredictability of the environment in which assets operate, along with the difficulty to quantify the output of the maintenance process itself. Current information systems employed for condition monitoring identify a failure condition when the asset is near breakdown and therefore serve as tools of failure reporting better than instruments for pre-warning the failure condition in its development [2]. On the other hand, owing to increased competitive pressures, maintenance strategies that once were run-to-failure are fast changing to being condition based, thereby necessitating integration of asset lifecycle data. This requires integration of decision systems and computerized maintenance management systems in order to provide support for tasks such as maintenance scheduling, maintenance workflow management, inventory management, and purchasing. However, in practice, data is captured both electronically and manually, in a variety of formats, shared among an assortment of off the shelf and customized operational and administrative systems, communicated through a range of sources and to an array of business partners and subcontractors; and consequently inconsistencies in completeness, timeliness, and inaccuracy of information leads to the inability of quality decision support for asset lifecycle management. In these circumstances, existing asset management information systems could be best described as pools of isolated data that are not being put to effective use to create value for the stakeholders. Asset management, therefore, needs a data-enabled integrated view of asset life cycle to realize effective, planning, execution, and management of the life cycle of each asset. There is growing need to treat asset as an organizational asset and implement data governance so that lifecycle data is aligned with the strategic asset management orientation; data is value adding and aimed at continuous improvement of the business; risks posed to information and information resources are managed; information acquisition, exchange, processing, and management resources are properly planned and managed; and necessary audits are carried out to manage and manage information ownership, accountability, security, and quality. This paper develops a governance framework for asset lifecycle data. It begins with a discussion on the role of information in enabling an engineering asset life cycle. The paper then develops the case of data governance, followed by a data governance framework for asset lifecycle management. This framework is of particular interest to large-sized public sector organizations that own, operate, and manage assets and are interested in developing an integrated and data-enabled view of asset life cycle. It should also be noted that this paper uses information and data interchangeably.

27.2 Asset Management Information Systems

In theory, information systems in asset management have three major roles; firstly, information systems are utilized in collection, storage, and analysis of information-spanning asset lifecycle processes; secondly, information systems provide decision support capabilities through the analytic conclusions arrived at from analysis of data; and thirdly, information systems enable an integrated view of asset management through processing and communication of information and thereby allowing for the basis of asset management functional integration [2]. Information systems for asset management seek to enhance the outputs of asset management processes through a bottom–up approach. At the operational and tactical levels, information systems are required to provide necessary support for planning and execution of core asset lifecycle processes. For example, at the design stage, designers need to capture and process information such as asset configuration; asset and/or site layout design and schematic diagrams/drawings; asset bill of materials; analysis of maintainability and reliability design requirements; and failure modes, effects, and criticality identification for each asset. Planning choices at this stage drive future asset behavior, therefore the minimum requirement laid on information systems at this stage is to provide right and timely information, such that informed choices could be made to ensure availability, reliability, and quality of asset operation. An important aspect of asset design stage is the supportability design that governs most of the later asset lifecycle stages. The crucial factor in carrying out these analyses is the availability and integration of information, such that analysis of supportability of all facets of asset design and development, operation, maintenance, and retirement are fully recognized and defined. Nevertheless, effective asset management requires the lifecycle decision-makers to identify the financial and non-financial risks posed to asset operation, their impact, and ways to mitigate those risks. This can only be achieved if the information technologies are integrated with operational technologies, such that deviations from the normal operational pattern could be identified and analyzed in a holistic manner. This holistic analysis requires operational, control, financial, and administrative information. Therefore, asset management paradigm calls for realizing both hard as well as soft benefits of information systems to the organization by using quantitative as well as qualitative means and developing, maintaining, and nurturing their connection to organizational development. This, however, can only be attained if data is regarding as the most critical asset of the organization and effective data governance policies are in place aimed at enabling, managing, and maintaining business operations and at the same time providing a roadmap in terms of alternatives and choices to act as strategic advisory mechanisms that support business planning, decision-making, and management.

27.3 Data Governance

Data governance represents the enterprise policies or strategies that define the purpose for collecting data and governing the ownership and intended use of data [3]. It is a fundamental function of any large organization that provides for business intelligence and decision support by accounting for issues relating to data quality and data management within the organization. Data governance, therefore, is top driven and is a critical component of a CIO's agenda. It is, however, a subset of IT governance where many models and approaches remain the same though a more limited domain is governed—only activities and issues relating to data [4]. Ironically, data governance is largely absent from many pieces of literature and frameworks on IT governance; however, it should not be taken less seriously as data is the basis for system interoperability, business rules and processes, and application design [5]. As data is used as a foundation of those functions, if it is not kept 'clean' and the overall data quality is reduced, it ends up affecting the bottom line of the organization. This negative effect on the bottom line is becoming more and more evident, for example, in North America over $600 billion in lost revenue has been attributed to the data quality among companies [6]. A survey of two hundred and fifty-seven organizations by Waddington [7] found that the main driver for the implementation of data governance was to support business intelligence and data warehousing initiatives. Business intelligence and data warehousing have become important tools leading to significant benefits and profitability of businesses. However, the most significant aspect of data governance is data ownership and accountability, which often start with the CEO, CIO, or the senior management of an organization. From this top layer, stewardship is assigned to individuals or groups. Data stewardship entails management of data in a way that complies with the data policies set in place from data governance, ensuring that data is used for its intended purpose [3]. Stewardship is often assigned to the current managers or committees that are subject matter experts and have application-level access to the given action or area of data. In asset management industry, data stewards are particularly important, where error in data can cause major problems for asset operators, maintainers, as well as the service that the assets enable. Asset-managing organizations especially require critical error checking and analysis of all data, as any discrepancies or problems that may occur will affect the reputation of the organization and possibly be threatening for the economy of the country.

Data governance, however, is not just about improving quality of data alone. It is about managing the intellectual 'capital' of the organization so as to enable a continuous improvement-based learning progression. Therefore, data governance includes technical, organizational, as well as people aspects related to optimum utilization of technical resources to produce, maintain, manage, and use data and information resources for the internal and external growth and maturity of the organization.

27.4 Approaches to Data Governance

There are three common approaches used when implementing governance programs and frameworks, i.e., decentralized, centralized, and hybrid approaches [8]. A decentralized approach (bottom–up model) is popular for lower-level managers of an organization having decision-making powers over their IT segments, including data. An example of this approach is a governance team creating their own naming conventions for data they use on an everyday basis; however, coordination with other teams is still needed to coordinate naming standards. The approach encourages lower-level managers to be innovative in their unit's performance. A centralized approach (top–down model) involves decision-making regarding organizational data to be done by the information systems management, which is a more central authority when compared to a decentralized approach. These decisions are made at the corporate level where senior information systems management takes their full responsibility and are then pushed down into the data application level of the organization. There can also be a hybrid mix of these approaches that involves mixing elements of centralized and decentralized governance.

Scope of asset management spans many business activities, where most of these activities are cross-functional or cross-enterprise. For example, maintenance processes influence many areas, such as asset operations; asset design/redesign; and safe workplace assurance. The outputs from maintenance are further used to predict asset remnant lifecycle considerations, asset redesign/rehabilitation, and planning for the maintenance support and spares supply chain management. This highlights that the foundation of effective asset management is embedded in how information is used to drive asset life cycle; how learnings and knowledge are synthesizing from this information; how these learnings are preserved and made available to all functions involved in the management of asset life cycle. In crux, asset management is policy driven, information intensive, and value adding and is aimed at achieving cost-effective peak asset performance. The core objective of asset management processes is to preserve the operating condition of an asset to near original condition. However, in order to achieve, this asset-managing organizations need to have a long-term vision of how they generate, process, and manage data to enable a continuous improvement regime for asset solution as well as lifecycle support. A one size fit all approach to data governance for asset management is, therefore, not possible. Each organization needs to its tailor data governance approach according to its technical, organizational, and social capability and maturity. What is critical to all approaches are the data governance functions that complement each other to create a data governance regime that enables a data-driven integrated view of asset management.

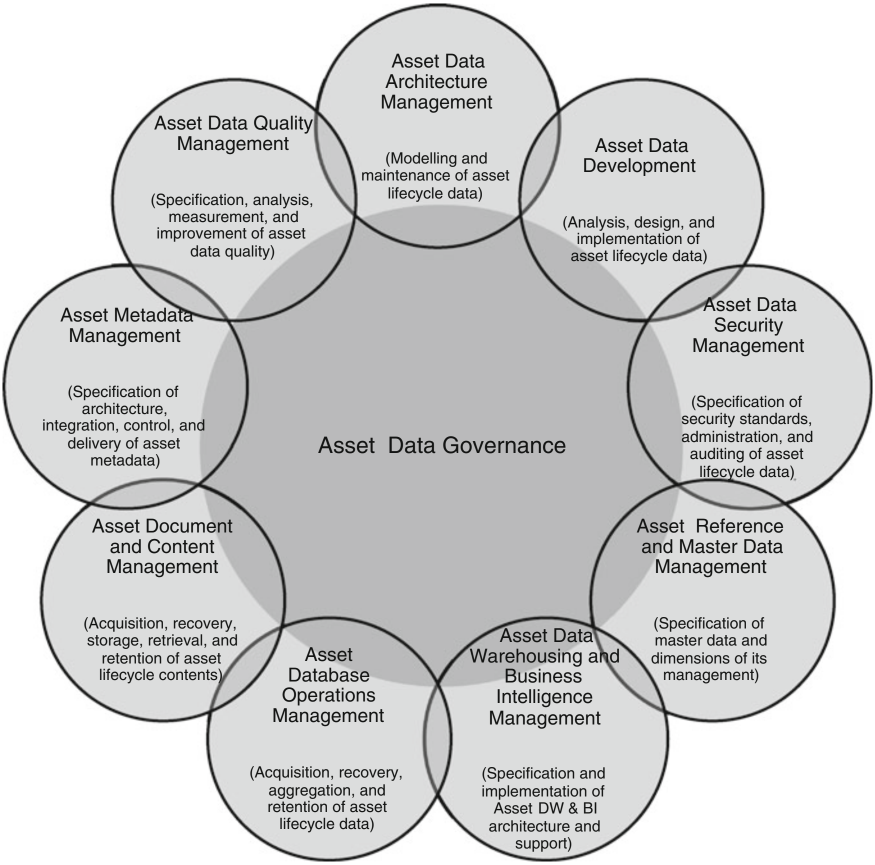


Fig. 27.1 Asset data governance functions (adopted and modified from [9])

27.5 Data Governance Functions

Data governance is a monolithic whole of a number of different dimensions such as the ones relating to strategic alignment of data with organizational orientation; value adding characteristic of information so that it enables continuous improvement of the business; mitigation of risk posed to information and information resources; information lifecycle management; and carrying out of necessary audits manage information ownership, accountability, security, and quality. These building blocks of data governance describe the broader scope of data governance dimensions relating to planning, controlling, and delivering data within an asset-managing organization (Fig. 27.1).

1. Asset data architecture management encompasses development and maintenance of overall data architecture of asset lifecycle data.

2. Asset development specifies data focused activities, such as modeling, analysis, design, implementation, and testing of asset data.
3. Asset data security management deals with the risk mitigation of asset lifecycle data, by specifying the support activities that ensure the secrecy, integrity, confidentiality, and authenticity of data.
4. Asset reference and master data management provides for specification and maintenance of semi-structured, structured, and unstructured asset lifecycle master data.
5. Asset data warehousing and business intelligence management stipulates planning, implementation and control processes for multi-dimensional data querying, analysis, and decision support.
6. Asset database operations management specifies how relational data is to be managed.
7. Asset metadata management specifies the schema of asset lifecycle data and its control, allowing for smooth data integration and interoperability.
8. Asset document and content management specifies how life cycle of unstructured asset lifecycle data like physical records, inspection sheets, documentation, videos, and images is to be maintained and managed.
9. Asset data quality management specifies planning and implementation of techniques and processes to cleanse and purify existing data, and to ensure that whatever data is entered in the organizational information systems conforms to quality specifications.

27.6 Data Governance for Asset Management

It has been established in earlier sections that asset-managing organizations need to have a long-term vision of how they generate, process, and manage data to enable a continuous improvement regime for asset solution as well as lifecycle support. Therefore, asset lifecycle management needs to be data/learnings focused, such that each lifecycle stage draws from and contributes to it to create a learning-based integrated view of asset life cycle. Information-enabled integrated asset lifecycle management implies that information requirements of asset management should dictate planning, execution, and management of asset life cycle. A data governance framework for asset management is illustrated in Fig. 27.2.

This framework divides asset life cycle into seven perspectives, i.e., competitiveness, design, operations, support, stakeholders, lifecycle efficiency, and learning perspective. It embeds aspects like data quality, integration, standardization, interoperability, and risk management into the frameworks through the connections between different perspectives. From top–down, the framework guides data governance policies and initiatives in mapping the organization's competitive priorities into asset design and reliability support infrastructure.

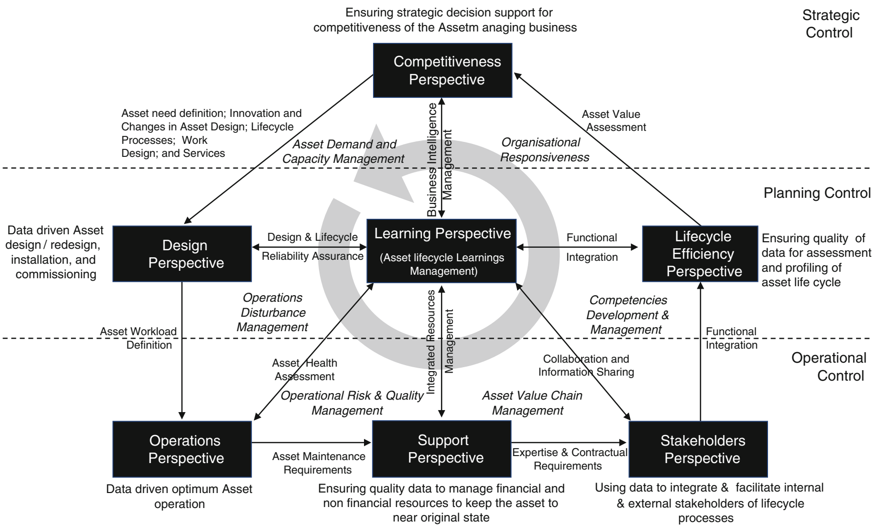


Fig. 27.2 Data governance model for asset management adopted and modified from [2]

The framework guides data governance initiatives through five further perspectives before informing the competitive priorities of the asset-managing organization. In doing so, the framework guides how data governance functions should be implemented, and at the same time can also assess the maturity of existing data governance initiatives. It thus shapes the role of data as strategic translator as well as strategic enabler of asset lifecycle management and enables generative learning. Instead of just being open-ended and laying down guidelines for implementing data governance initiatives, the framework has inbuilt assessment mechanism that provides gap analysis of the desired versus actual state of the maturity of organizational resources that enable data governance initiatives in the organization. Thus, the introduction, sustenance, and maturity of data governance in asset management depend upon three dimensions, i.e., asset lifecycle processes; functional elements of data governance; and the critical factors that contribute to organizational competence that facilitate data governance. Figure 27.3 illustrates the data governance cube to be applied to each of the seven perspectives of the framework described in Fig. 27.2.

To apply the data governance cube to asset lifecycle management framework described above, lifecycle processes under each perspective of the framework need to be identified. After the identification of processes, data governance functions for each process needs to be completed as per the different dimensions of organizational competence. For example, under the 'operations perspective' in Fig. 27.2, there is an asset lifecycle management process of 'developing asset operation workload schedule.' Viewing this process through the lens of 'data architecture modeling' highlights overall the type, object modeling, and relation of process data within the process as well as with other organizational data. These dimensions

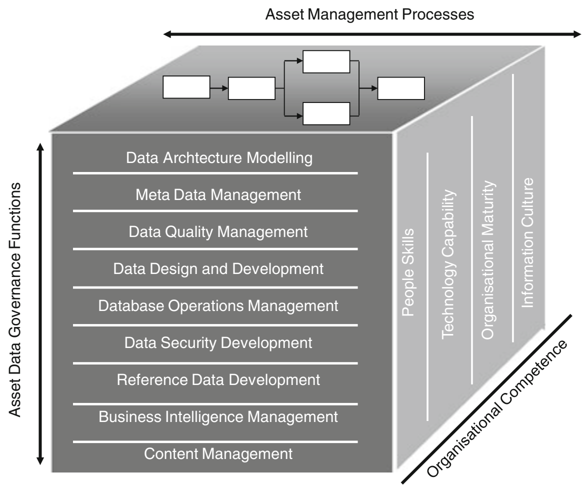


Fig. 27.3 Data governance cube

could then be measured on the four dimensions described under 'organizational competence.' Similar measurements for all the processes under the same perspective in Fig. 27.2 would enable the organization to figure out how competent the organization is in instituting overall enterprise data architecture modeling in terms of the skills of the employees to do so, availability of technology support, operational maturity, and culture of continuous improvement of the information and related resources. Doing so would provide the organization a gap analysis between the existing and optimum state of data governance in the organization. Similar measurements for each perspective in Fig. 27.2 would provide the complete picture of asset data governance within the organization. Comparison of the measurement results with the desired optimum state would highlight the gaps in organizational performance on instituting data governance. This gap analysis provides the roadmap for introduction, sustenance, maturity, and continuous improvement of data governance functions and related organizational resources.

27.7 Conclusion

Data governance is an important topic for any organization that acknowledges the importance of their business data as a foundation to their success. It is an area of corporate management that looks at decision-making and authority for data-related matters. This paper has presented a case of data governance for asset lifecycle management by highlighting as to what it is, how is it relevant to asset management, and how it could be implemented. Asset-managing organizations need to devise data governance policies that protect their data assets without hampering legitimate access. Data governance policies need to ensure that the quality of an organization's data is high, as it will be used for decision-making, and low-quality data will nearly always result in ill-informed decisions. As well as the quality of the data, the quantity of the data is important, with asset-managing organizations recording too much data and not having adequate management strategies in place. Technology can help with some of these issues, but it brings with it its own pitfalls. Organizational competence as a whole needs to be taken into account for effective introduction and maintenance of data governance in the organization.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centers Programme. The author gratefully acknowledges the financial support provided by the CRC.

References

1. Haider A, Koronios A (2005) ICT Based Asset Management Framework. In: Proceedings of 8th international conference on enterprise information systems, ICEIS, Paphos, Cyprus. vol 3, pp 312–322
2. Haider A (2013) Information systems for engineering and infrastructure asset management. Springer, Wiesbaden, Germany
3. Fernandes L, O'Connor M (2009) Data governance and data stewardship. Critical issues in the move toward EHRs and HIE. J Am Health Info Manage Assoc 80(5):36–39
4. Ariyachandra T, Fuller C, Watson H (2004) Data warehouse governance: best practices at blue cross and blue shield of North Carolina. Decis Supp Syst 38(3):435
5. Betz CT (2007) Architecture and patterns for IT service management, resource planning, and governance. Morgan Kaufmann Publishers, San Francisco, USA
6. Friedman T, Logan D, Newman D (2008) Governance in an essential building block for enterprise information management. Gartner Research, USA
7. Waddington D (2010) Data governance, MDM and data quality. DM Direct 20(5):14
8. Ross J, Weill P (2005) A matrixed approach to designing IT governance. MIT Sloan Manage Rev 46(2):26–34
9. Mosley M, Brackett M, Earley S, Henderson D (2009) The DAMA guide to data management body of knowledge (DAMA-DMBOK guide). Technics publications, USA

Chapter 28
Institutionalization of Information Technologies for Engineering Asset Management: An Australian Case

Abrar Haider

Abstract Information technologies (IT) implementation in asset managing organizations does not follow a linear path. It is primarily driven by cost concerns, rather than an approach that takes into account the existing technological infrastructure, business requirements, available skill base, social and cultural environment, and operational and strategic value of technology investment. This paper presents a case study of information technologies implementation in Australian asset managing organizations. It concludes that technology for asset management needs to be physically adopted, and socially and organizationally composed, to create consensus on what the technology is supposed to accomplish and how it is to be utilized.

28.1 Introduction

Traditionally, asset managers focus on developing the technical foundation for asset life cycle management around operational technologies and leave the selection, adoption, and maintenance of information technologies (IT) to IT managers. This may be attributed to the propensity of asset managers to view information systems utilization in general as a secondary or support activity to execute business processes. Their emphasis is more on the substitution of labor through technology utilization rather than business automation and functional integration aimed at internal efficiency and overall strategic advantage. Since the level of input from asset managers regarding choice of information systems has a narrow focus, these systems do not contribute to the organization's responsiveness to internal and external challenges. As a result, role of IT in managing engineering

A. Haider (✉)
Mawson Lakes Campus, University of South Australia, Adelaide, SA 5095, Australia
e-mail: abrar.haider@unisa.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_28,

assets has not fully institutionalized. Institutionalization of IT for asset management, however, is strongly underpinned in the political, economic, and cultural context of the organizations, which bring together individuals and groups with particular interests and interpretations and help them in creating and sustaining information systems as socio-technical systems. This research presents a study of Australian engineering and infrastructure asset managing organizations and focuses how they implement IT to manage the life cycle of their assets.

28.2 Scope of Asset Management IT

In theory, IT in asset management have three major roles; firstly, IT are utilized in collection, storage, and analysis of information spanning asset life cycle processes; secondly, IT provide decision support capabilities through the analytic conclusions arrived at from analysis of data; and thirdly, IT enable an integrated view of asset management through processing and communication of information and thereby allowing for the basis of asset management functional integration [1]. IT for asset management seek to enhance the outputs of asset management processes through a bottom-up approach. At the operational and tactical levels, IT are required to provide necessary support for planning and execution of core asset life cycle processes. For example, at the design stage, designers need to capture and process information such as asset configuration; asset and/or site layout design and schematic diagrams/drawings; asset bill of materials; analysis of maintainability and reliability design requirements; and failure modes, effects, and criticality identification for each asset. Planning choices at this stage drive future asset behavior, therefore, the minimum requirement laid on IT at this stage is to provide right and timely information, such that informed choices could be made to ensure availability, reliability, and quality of asset operation. An important aspect of asset design stage is the supportability design that governs most of the later asset life cycle stages. The crucial factor in carrying out these analyses is the availability and integration of information, such that analysis of supportability of all facets of asset design and development, operation, maintenance, and retirement are fully recognized and defined. Nevertheless, effective asset management requires the life cycle decision makers to identify the financial and non-financial risks posed to asset operation, their impact, and ways to mitigate those risks. This can only be achieved if the IT are integrated with operational technologies, such that deviations from the normal operational pattern could be identified and analyzed in a holistic manner. This holistic analysis requires operational, control, financial, and administrative information. Therefore, asset management paradigm calls for realizing both hard as well as soft benefits of IT to the organization by using quantitative as well as qualitative means and developing, maintaining, and nurturing their connection to organizational development. This, however, can only be attained if data are regarding as the most critical asset of the organization and effective data governance policies are in place aimed at enabling, managing, and maintaining business operations and at the same time providing a road map in terms of alternatives and choices to act as strategic advisory mechanisms that support business planning, decision making, and management.

28.3 Research Methodology

This exploratory research employs an interpretive epistemology with a qualitative perspective. The aim of this research was to explore the issues that influence institutionalization of IT for asset management. In order to address the research question for this study, 25 middle managers representing various roles associated with asset life cycle management were interviewed in asset managing organizations that manage below track rail assets, water pumping station assets, electricity assets, and road infrastructure assets. These interviews were conducted over a one–one and half hours period and included the following job descriptions, asset designers, maintenance engineers, network access manager, business development manager, operations and maintenance manager, manager projects, manager assets management, project officer assets, finance manager, and IT manager. The interview questions were open-ended, and interviewees had freedom to describe their experiences and problems beyond the scope of the questions. In addition, researchers were provided access to all documentation concerning asset life cycle management, as well as access to sites of asset operation. The interviews were transcribed, and data from all sources were analyzed using NVivo software using its features of themes, descriptions, and assertions. The conclusions drawn in the following case, thus, represent interpretations of the evidence as understood by the author. To comply with confidentially agreements with the case study participants, theirs as well as their organization's identity has been suppressed.

28.4 Institutionalization of IT for Asset Management: Case Study

IT in Australian asset managing organizations have evolved in response to various industrial, legislative, and competitive pressures. The following sections provide an analysis of how IT have been implemented and assimilated in case organizations.

28.4.1 IT Implementation Strategy

All the organizations investigated for this study lacked a centralized technology adoption policy and do not conform to a common information model. As a result, different functions have implemented their own customized spreadsheets and databases to aid their day-to-day operations. Although these applications aid in the execution of work within the department, however, they are not integrated with other information systems and are of little value to other departments who could use this information for better asset life cycle management. In general, asset managing organizations have a strong cost focus, but the same is not matched as far as the commitment from senior management in terms of adoption of new technology in concerned. For example, Business IT systems Manager of the water asset management organization stated that '*last year (while the company was considering implementation business intelligence infrastructure) a vendor walked through the door and asked what you want. Management did not know what are business intelligence systems, what are their capabilities, and what kind of reports can they generate……… I am not a qualified engineer so I cannot quality assure if the system is capable of providing what the engineers want ………(for implementation of technology) we consider the total cost of ownership. The actual implementation may be cheaper but when we consider costs incurred on learning a new technology, and costs relating to adoption of technology you have to make a decision on the total cost of ownership and not just the initial implementation cost.*' More or less the same sentiment is reflected in the response of the IT manager of electricity asset management organization, who summarizes the technology adoption approach of the company as, '*we are not early adopters, and we are not explorers and we are not easily influenced or driven by whatever the latest thing on the market is. It is need driven and business case driven. Basically in past our motivating factors have been tactical needs of individual areas, so it hasn't been strategic at all but its moving towards being more strategic mainly for information integration. We now have stronger governance and cost focus, since we are now viewing ourselves as a market player as we are expanding nationally and are moving into more commercial roles.*'

28.4.2 IT Implementation at Operational Level

Technical infrastructure of the case organizations consists of various off the shelf proprietary, legacy, customized systems, and a number of ad hoc solutions in the forms of spreadsheets and databases. Off the shelf systems are developed on customized guidelines and support proprietary data formats; whereas legacy systems are technologically weak even though they have evolved with the organization. These systems have been in operation for more than 20 years, developed in old technologies, and are not compatible with new technologies. Ad hoc solutions developed by employees on their own and that do not conform to any quality or technical standard are naturally isolated from the mainstream technology-based logical and physical operating model of the organization. As a result of these anomalies, asset life cycle information is hard to aggregate, lacks interoperability, and has tight coupling with technology and cannot be reused. Information systems in case organizations are nothing more than isolated pools of data that may serve the needs of individual departments but do not contribute toward an integrated information enabled view of asset life cycle management. This means that the existing technical infrastructure in general and information systems in particular is not aligned with the strategic asset management considerations, does not contribute to functional integration, and does not conform to a unique enterprise information model.

In terms of asset design, water asset managing organization presented an interesting scenario. In the formal process of asset design/redesign, chief engineers visit different regions and talk to designers and operations to discuss the design and operational requirements. Deign/redesign process of an upgrade or refurbishment is carried out through consultation with designers and is fully reviewed by the technical service engineers in that region. The company thus ensures that they have consensus from all the parties involved. In so doing, there is heavy reliance on the tacit knowledge held by staff, whereas there is no system for preserving the same or making it available to other functions within the organization. In the words of the design manger, '*we use AutoCAD but that just gives us an electronic version of a drawing. Ideally we would like to have access to information to analyze how good our design is. The information available to us has been input by the people who really don't have sufficient technical background to understand the key things that need to be there. We've got long way to go I think before our systems are going to be sufficiently up to date and have sufficient useful information that our guys could pretty quickly get a hand on. Once we cannot get our hands on this information we have no choice but to rely on the knowledge held with our filed staff.*' The design exercise has a strict focus on the design of the asset, and there is not enough consideration given to the support asset design throughout its life cycle. Although it can also be attributed to the fairly stable nature of water infrastructure assets, however the major cause for inability to carry out a comprehensive set of supportability analysis is the non-availability and lack of access to requisite information. Apart from this, the company being a semi government organization, retains the hierarchical silo approach that resists information exchange and collaboration, which is evident from the following response of a design engineer, '*getting the historical information in order to get the design right at the first place is very difficult. In order to get this information costly onsite information extraction is required, that is by talking to the local people. We would like to see accurate maintenance information in the field, like what is happening, why it happened, what are the failures* etc. *On top of this, there is no interaction between electrical and mechanical engineers. They just log information in their own systems for someone else to look at and make decisions. If you want to know performance of a particular system, your best bet is to go and speak to a field* *engineer.*' This argument is further supplemented by a design engineer in the infrastructure asset managing organization, which further amplifies the manifestation of a silo approach to asset life cycle management, who commented, '*we are not into risk assessments in a big way. Obviously all of systems are subject to corporate risk assessment. From design point of view we look at what condition we need to implement to manage risks posed to an asset, and that's where all the issues are that we want to monitor. For example, what space do we need to back the system up, what redundancies do we need to guarantee* etc. *So all those sort of things we do as a matter of course in the design exercise. Once the asset is in operation it's for the operations and maintenance people to do risk assessments. They never ask us for information on any previous assessments and even if they do we cannot provide it to them easily since we perform our assessments manually.*'

28.4.3 IT Implementation at Planning/Management Level

At planning and management level, the major IT concern is how it must be implemented to meet the planning and control of asset life cycle management? Whereas the aim of implementation of these technologies is to fulfill asset life cycle planning and control requirements aimed at continuous asset availability, through performance analysis based on analysis of various dimensions of asset information such as design, operation, maintenance, financial, and risk assessment and management. However, the situation on ground is far from being perfect. For example, the case of managing water infrastructure asset operation that we investigated. Water is sourced from specific supply points and thus cannot be pumped from anywhere, which means that the infrastructure is static and the environmental and operational constraints on the asset are relatively easily predictable. This also means that the water asset infrastructure has to operate at a certain level and the usual principles of load apportionment do not work in this situation. This organization's assets base consists of a variety of asset types and is spread anywhere between 30 and 100 km or even longer. Asset monitoring, therefore, is not only costly but is also time-consuming. Although the company makes use of geographical information system (GIS) and SCADA system to monitor asset operation, yet asset operation and condition assessment are largely manual. The information captured through SCADA systems is only used for alarms generation and failure reporting, it is not used or aggregated with other information for analysis such as failure root cause. The operations manager of the company suggests that '*condition assessment is manual exercise at the moment, since we are struggling to integrate different systems with our major asset management system (SAP). When we are required to do condition assessment, our guys will go and do that and in the process if they identify something that in their opinion presents an undesirable outcome they will flag that.*'

The organizations investigated for this research differentiate between maintenance and asset ownership, i.e., work is executed by maintenance crew, whereas asset ownership is the mandate of another function. Consequently, there are multiple versions of the same information within the organization. Furthermore, these versions have their own bias and standard of quality. Although the organizations are aware of the potential of quality information, yet there is little emphasis made on recording and capturing correct and complete information. For example, a maintenance engineer of the rail asset managing organization summarized the situation as, '*maintenance crew is not technically qualified or capable to operate an IT system. They consider it as an add-on to maintenance work, something that just has to be recorded. At the end of the day they will not be judged on what information they entered. Their performance is evaluated on the quality of their maintenance work.*' Maintenance information, however, is crucial for asset life cycle management, as it provides the basis for life cycle cost benefit analysis, remnant life cycle calculations, as well as for asset refurbishment, upgrade, and overhauls. However, like other functions, maintenance information is not exchanged with other life cycle functions. In addition, the main focus is on capturing maintenance execution information with little provision for integrating this information with financial information. Consequently, there is no way of calculating the cost of failure as well as real costs incurred on maintaining the asset.

The rail asset managing organization caters for metropolitan as well as countryside track assets and therefore is not only concerned with the traffic on the tracks but also the weight borne by these tracks over the period of time. Traffic is managed by state-of-the-art software that manages as well as allocates traffic on the tracks; whereas the condition of the track is monitored through sensors and is also inspected manually. Human inspections actually constitute majority of condition monitoring due to the remoteness of the track assets. The company has an extensive network of track inspectors, which includes a substantial number of indigenous Australians who are well known for their knowledge and familiarity with outback terrain and geography. The organization relies heavily on their tacit knowledge, and these track inspectors have also proven to be extremely reliable sources of track information. However, there has been no effort made to record information collected through these manual inspections. Whereas there are certain aspects of asset operation that seem to be over automated, as is evident from the response of operations manager, '*for a case of a broken rail, essentially it's about train coming off. One system records broken rail, which goes to the network controller who can stop trains from going on the track. Another system records the same incident the same information in a track incidence system to raise signal alarms. Yet another one of the systems records the same incident in the rail defect management system, such that a request could be generated to fix it. Now you have the same information available in three different systems. There is not only duplication, in fact triplication of information. Information in each system has a bias towards a particular function, so which version is more credible?*' This symbolizes the typical behavior of a public sector large organization. In this case, each function trusts its own information and does not believe in sharing the same. As a result, there is significant wastage of effort and finances, and quality of information is undermined due to information redundancy and lack of integration.

28.4.4 IT Implementation at Strategic Level

IT resources represent the combination of IT infrastructure, human IT resources, and soft assets, such as the shared performance and prospect development potential of a corporation. Organizations improve externally and internally by making informed choices which may affect the learning, acquiring, and operation of IT resources. The closeness between the Chief Executive Officer, Chief Information Officer, and Chief Technology Officer can improve the organization by bringing technology and supporting organizational changes together, which are vital for successful performance as well as effective management of related resources. The visions of business and asset managers must address the role, which IT present in their organizations; managers must also address directions to employ resources to gain more benefits for their business. The important point here is that IT infrastructure needs to be designed to provide a holistic view of the asset life cycle. In such a setting, it is imperative to preserve life cycle learnings and the organization should use the same to evolve, grow, and mature.

In the case organizations, IT are being used for information keeping or what could be best described as recording what the organization has done. This information is seldom used for more high-profile purposes, such as organizational integration, planning, and executive decision support. The prevailing silo approach has thus affected departmental efficiency as well as functional integration. User training has traditionally been a weak area, as little training is provided and even whatever is provided is aimed at training managers and supervisors rather than the staff who use the system on daily basis. The idea is that the supervisory staff (with the help of IT department) train functional staff. In these circumstances, it is obvious that staff do not feel comfortable with using major platforms such as SAP and business units A are more inclined toward using internally developed spreadsheets and databases. Since there is little information exchange internally as well as with business partners, there is substantial information and knowledge drain. Senior management is not technology savvy and therefore does not rely on IT for asset life cycle decision support. Even otherwise, information lacks quality and there is no way of managing the important asset life cycle learning. For example, the water asset managing organization depends a lot on the tacit knowledge (particularly for maintenance) and at present more than 65 % of the staff at the company is within 10 years of retirement age, which means substantial amount of intellectual capital loss.

28.5 Discussion

IT implementation for asset management in case organizations has narrow focus and scope, which emphasizes technical aspects and does not give due attention to organizational, social, and human dimension of technology implementation. This approach to technology implementation at best serves to treat IT as process automation tools and does not contribute to the cultural, organizational, and technical maturity of the organization. Technology is a passive entity, and its use is shaped by its interaction with organizational, environmental, and human factors. Implementation exercises that do not account for the cause and effect relationship that shapes technology are unable to institutionalize technology in the organization. There is an evident lack of commitment from top management in case organizations to institutionalize technology. As a result, IT implementation in general and information systems implementation in particular have been disorganized and is not driven by the strategic business considerations. Most of these technologies have been implemented due to the pressure from regulatory agencies. Thus, these technologies have been pushed into the IT infrastructure of the organization, without considering the fit between business processes and technology. This lack of cultural, organizational, and technical alignment and lack of user or stakeholder's involvement in technology adoption hamper development of a collaborative, creative, and quality-conscious organizational culture and impede organization-wide coordination and horizontal integration. IT implementation, thus, is heavily predisposed toward a technology push rather than technology pull strategy. The case study, thus, affirms that asset managing organizations follow technology determinism for implementation of IT supporting asset life cycle. In summary, IT have not been institutionalized properly in the organization.

Institutionalization of IT is strongly underpinned in the political, economic, and cultural context of the organizations, which bring together individuals and groups with particular interests and interpretations and help them in creating and sustaining information systems as socio-technical systems. Institutional isomorphism is a process in which organizations aim to excel in their practice of social rules, ideals, and practices by aligning themselves with the environmental conditions. These institutional pressures push organizations to adopt shared notions and routines. Thus, the interpretation of intention to adopt technology and the prevailing context of the organization is affected by its perception of these pressures. Coercive, normative, and mimetic are three isomorphic mechanisms which influence organizations in gaining operational efficiency, similarity with peers, and success [2]. Regulative, cultural-cognitive, and normative are three institutional views representing theses isomorphic pressures which are not mutually exclusive and are interdependent (Fig. 28.1). It is important for the Australian asset managing organizations to strike a balance between these mechanism, in order to be able to create the shared understanding of the use and value of IT and to align it with the social, cultural, and organizational institutions that constitute the context of asset life cycle management.

The coercive isomorphism occurs by organizational desire to conform to laws, rules, and sanctions established by institutional actors or sources. The existing backdrop of IT in the case organizations represents a fragmented approach aimed at enabling individual processes in functional silos. These organizations are aiming to mature technologically along the continuum of stand-alone technologies to integrated systems (like SAP ERP system), and in so doing are aiming to achieve

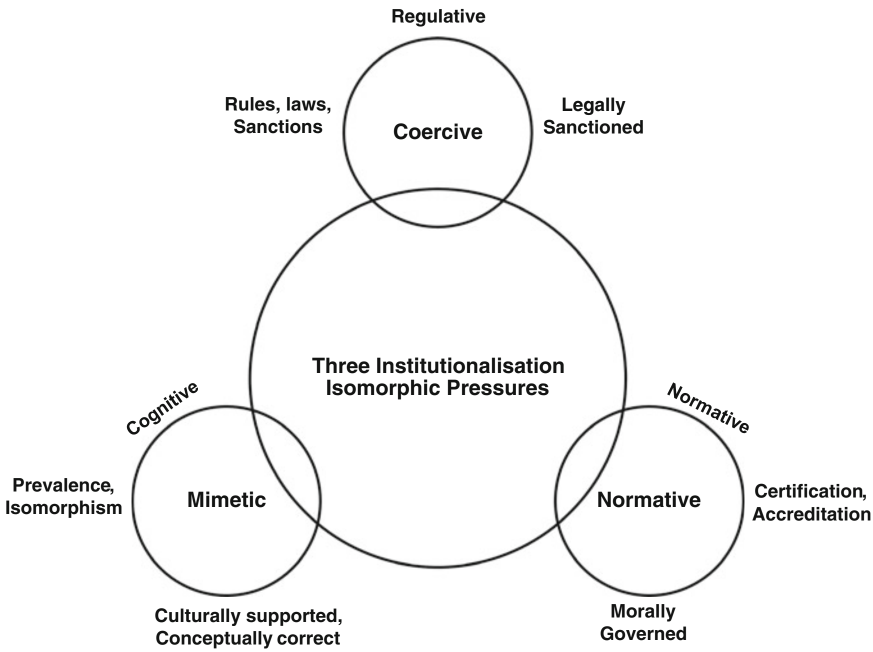


Fig. 28.1 Institutional isomorphism mechanisms [3]

the maturity of processes enabled by these technologies and the skills associated with their operation.

The case study reveals that most of the asset management specific technology initiatives have been in response to the legislative pressure from the government. Being government organizations that have changed to semigovernment, the case organizations are under significant pressure for compliance. There is no technology that uniformly covers every aspect of asset management; therefore, the coercive pressure to adopt particular technology creates asymmetry of power within the organization, where some functions are well automated and some and not. However, since the case organizations adopted technology without accounting for their information requirements, contextual strengths and weaknesses, and other factors such as maturity of existing technical infrastructure, they spent a lot of resources in fire fighting rather than utilizing technology for their optimum advantage. In actual affect, in most cases there were two set of technologies working in parallel in the organization, where one was forced upon the organization by external pressure, and the other set of technologies that the users felt comfortable with. A good example is utilization of SAP, and the same time scores of ad hoc spreadsheets developed in Microsoft Excel.

The normative mechanism concerns the moral and pragmatic aspect of legitimacy by assessing whether the organization plays its role correctly and in a desirable way. It can refer to the positive pursuit of valued ends, as well as negative deviations from goals and standards [4]. The disparity in the way technology is used at each stage of asset life cycle explains the normative influences. For example, maintenance has traditionally been the focus of asset life cycle management. It is not surprising that in the case study organization maintenance function is the most technology intensive. However, the normative pressure created by the maintenance function for technology enablement has not transcended to other function, due to the fact that the case organization is a hierarchical and operates in functional silos. There is little interaction between different functions of asset life cycle; consequently, themes relating to success and effectiveness of technology have seldom crossed the functional boundaries that could stimulate the same of view of technology by decision makers in other functions. This is explained in the knockout quote from the group manager of Rail asset managing organization who while explaining lack of information ownership in the organization stated that '*some would argue that we manage assets, not intellectual property!*' In such circumstances, top management has been unable to create a vision or a technology road map by mapping organizational needs with capabilities of technology to fulfill the demands of each stage of asset life cycle. Since the direction from top is missing, people view technology as a passive activity rather than an active enabler of primary activities in asset management. Compliance with these norms with respect to strategic concerns (as viewed by the stakeholders) has led to the existing form of legitimacy and control within the organization, which has strengthened the status quo of a chaotic strategy to the utilization of IT within the organization.

The mimetic isomorphism is a cause of organizational tendency to remain similar to its peers in order to get a positive evaluation from the organizational environment. This mechanism results in reducing uncertainty, improving predictability, and benchmarking organizations that are performing at or near optimum level [5]. None of the case organizations engage in taking stock of their technical infrastructure and the business processes enabled by it. As a result, these organizations are unable to find how well their business processes are performing, how effectively these processes are coupled with technology, and what are the information gaps or requirements that technology has not fulfilled. Technology is 'pushed' into the technical infrastructure of these organizations based on its reputation rather than its applicability or usefulness. Consequently, there is task technology mismatch that gives rise to issues such as lacking information integration and interoperability across these organizations.

28.6 Conclusion

This paper has presented a case study on the state of IT implementation in asset managing organizations. It concludes that issues of IT implementation range from technical issues to social, cultural, managerial, and organizational issues. However, the origin of these issues can be traced back to two factors: inadequate planning for institutionalization of technology in the organization, and disregard of organizational and social change associated with technology adoption. IT for asset management call for consideration of organizational, technical, structural, and people dimensions of IT to create the 'shared understanding' and 'meaning' of the use and value of IT. It is the way that technology is institutionalized and consequently legitimized, which defines the responsiveness and maturity of the asset managing organizations.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established, and supported under the Australian Government's Cooperative Research Centers Program. The author gratefully acknowledges the financial support provided by the CRC.

References

1. Haider A (2007) Information systems based engineering asset management evaluation: operational interpretations. PhD thesis, University of South Australia, Adelaide, Australia
2. Greenwood R (2008) The SAGE handbook of organizational institutionalism. Sage Publications, Los Angeles
3. DiMaggio PJ, Powell WW (1983) The Iron cage revisited: institutional isomorphism and collective rationality in organizational fields. Am Sociol Rev 48(2):147–160
4. Scott WR (2008) Institutions and organizations; ideas and interests (3rd edn). Reference and research book news. Sage Publications Inc., Ney York
5. Teo HH, Wei KK, Benbasat I (2003) Predicting intention to adopt interorganizational linkages: an institutional perspective. MIS Q 27(1):19–49

Chapter 29
Experimental Measurement and Theoretical Modeling of Internal Combustion Engine Valve Leakage by Acoustic Emission Method

Seyed Mohammad Jafari, Hamid Mehdigholi and Mehdi Behzad

Abstract This paper presents an experimental study on the vibration signal patterns associated with a simulated piston slap test of a four-cylinder diesel engine. It is found that a simulated worn-off piston results in an increase in vibration RMS peak amplitudes associated with the major mechanical events of the corresponding cylinder (i.e., inlet and exhaust valve closing and combustion of Cylinder 1). This then led to an increase of overall vibration amplitude of the time-domain statistical features such as RMS, crest factor, skewness, and kurtosis in all loading conditions. The simulated worn-off piston not only increased the impact amplitude of piston slap during the engine combustion, but it also produced a distinct impulse response during the air induction stroke of the cylinder attributing to an increase of lateral impact force as a result of piston reciprocating motion and the increased clearance between the worn-off piston and the cylinder. The unique signal patterns of piston slap disclosed in this paper can be utilized to assist in the development of condition-monitoring tools for automated diagnosis of similar diesel engine faults in practical applications.

29.1 Introduction

As part of a broader project being undertaken by the CRC for infrastructure and engineering asset management, an online diesel engine-monitoring system is being developed in collaboration with ASC Pty Ltd. The application of this system is initially focused on its application to a Collins Class Submarine test engine; however, such a system and its associated technologies have the potential for a much broader application.

S. M. Jafari · H. Mehdigholi · M. Behzad (✉)
Sharif University of Technology, Azadi St, Tehran, Iran
e-mail: M_behzad@sharif.edu

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_29,

Piston slap is one of the most common combustion-related diesel engine faults, which is caused by excessive clearance between the piston and the cylinder wall or liner due to wear or inappropriate operation of an engine. Excessive piston slap can induce severe engine damage such as scuffing on the piston and cylinder wall/liners. Vibration and noise radiation from diesel engines due to piston slap has been studied for several decades.

29.2 Theoretical Model

29.2.1 Acoustic Emission

Acoustic emission technique (AET) in simple terms is defined as the transient elastic stress waves generated due to dynamic events like fluid leakage, crack propagation, plastic deformation, corrosion, and fracture in materials. Acoustic emission technique is considered quite unique among the nondestructive testing methods [1].

The feasibility of using AET for leak detection in general depends on three main factors: energy radiated from the leaks, attenuation between the source and the sensor, and ambient noise. The minimum detectable leakage depends on the type of fluid, the geometry of discontinuity, and the sensitivity of the AE system. Signals due to leakage are wideband in general extending from low audible range to 1 MHz [2].

AE signals of the leakage are continuous AE signals. The most used parameter for continuous waves is average energy. Average energy of the raw AE signals is calculated by taking a root mean square value over a time, t, as follows:

$$\mathrm{AE}_{\mathrm{rms}} = \sqrt{\frac{1}{T}\int_{t_0}^{t_0+T} v^2(t)dt} \tag{29.1}$$

where $v(t)$ is the amplitude of the AE waveform in volts (V), t_0 the initial time, T the integration time of the signal in seconds, and AErms average energy in (V) [3].

29.2.2 Flow Area Through Valve

Intake and exhaust valves have important role for inlet air to the working cylinder and discharge of burned air/fuel combination after combustion. Therefore, in the case of valve not closed completely, engine torque or power would be affected directly due to valve leakage. The most obvious factor that affects the gas flow and output of an engine is the flow area.

Considering typical engine designs, at low lifts, the flow area normal to the seat faces is shown in the Fig. 29.1 which is followed by Eq. 29.2.

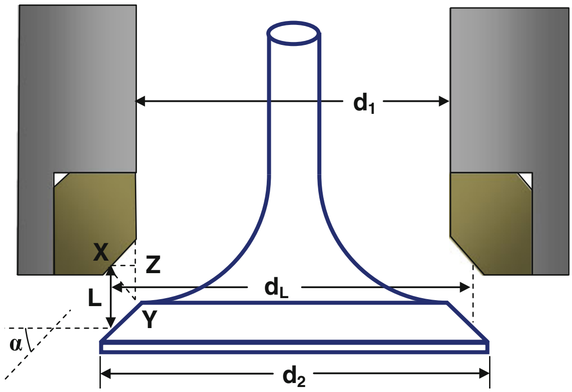


Fig. 29.1 Valve area with conical seats [4]

$$d_L = d_1 + 2\overline{XZ} = d_1 + 2\overline{XY}\sin\alpha = d_1 + 2L\cos\alpha\sin\alpha \tag{29.2}$$

where L is the valve lift and α is the valve seat angle.

Now, the surface area of a right circular cone is

$$A = \frac{\pi}{2}\overline{XY}(d_1 + d_L). \tag{29.3}$$

This gives the valve-opening area per valve:

$$A_V = \pi L\cos\alpha(d_1 + L\cos\alpha\sin\alpha) = \pi d_1^2\cos\alpha\left(1 + \frac{L}{d_1}\cos\alpha\sin\alpha\right)\frac{L}{d_1} \tag{29.4}$$

where A_V is the perpendicular area to the flow path. A non-dimensional lift L/d_1 is convenient when comparing flow characteristics [4].

29.2.3 Flow Characteristics and Velocity

At low lift, the air jet fills the gap and adheres to both the valve and insert seat. In this state, it is common practice to assume incompressible flow. By increasing lift step by step, the flow will break away from one of the surfaces gradually and compressible flow is a good assumption. The ideal velocity for the incompressible flow would be [4]

$$v = \sqrt{2RT_0\frac{r-1}{r}} \tag{29.5}$$

And for compressible flow, it would be

$$v = \sqrt{\frac{2\gamma R T_0}{\gamma - 1}\left(1 - \frac{\gamma - 1}{r\gamma}\right)} \tag{29.6}$$

where

v Frictionless velocity, m/s
R Gas constant, J/kg K
T_0 Upstream temperature, K
r Pressure ratio $r = P_0/P_d$
P_0 Upstream pressure or cylinder pressure, N/m^2
P_d Downstream, N/m^2
γ Ratio of specific heats = 1.4 for air

29.2.4 Theoretical Formula Between AE Energy and Valve Leakage Characteristics

It is possible for turbulent flow such as valve leakage to induce vibrations in their structures. The frequency of oscillation depends on the stiffness of the structure. The amplitude depends on mean flow velocity and on how close the forcing frequency is to the natural frequency of the structure.

Lighthill's formulation of the aeroacoustic problem showed that the mixing region of a jet could be equated to a volume of quadrupoles with strength proportional to the stress tensor in the moving fluid. Such a mathematical formulation allows radiated jet acoustic power levels to be estimated and shown to be proportional to v8 [5, 6]

$$P_s = \frac{\rho v^8 d_1^2}{\alpha^5} \tag{29.7}$$

where P_s is the sound power (W), ρ is the fluid density (kg/m^3), v is the average turbulence jet velocity (m/s), d_1 is the valve size (m), and α is the sound velocity in the fluid (m/s).

Valve leakage generates continuous AE signal so the most commonly used AE parameters are the AE energy because AE activity is attributed to rapid releases of energy in the material; the energy content of the AE signal is related to this energy release. From the literature [5], it was shown that the direct proportional relationship between AE energy and the sound power of fluid flow P_s is

$$E = C_0 P_s \tag{29.8}$$

where C_0 is the proportional constant. It is the function of fluid variables with some factors neglected to simplify the relationship, such as AE sensor, instrumentation gain, reference voltage, signal attenuation, AE acquisition system, and valve material.

Substituting Eq. 29.7 into Eq. 29.8 results

$$E = C_0 \frac{\rho v^8 d_1^2}{\alpha^5}. \tag{29.9}$$

Then, air density is estimated from the ideal gas law [7]:

$$\rho = \frac{P_0}{RT_0}. \tag{29.10}$$

Substitution of Eqs. 29.5, 29.6, and 29.10 into Eq. 29.9 and rearranging yield For incompressible flow,

$$E = C_{\text{inc}} \frac{P_0 (RT_0)^3 d_1^2}{\alpha^5} \left(\frac{r-1}{r} \right)^4; \tag{29.11}$$

For compressible flow,

$$E = C_{\text{com}} \frac{P_0 (RT_0)^3 d_1^2}{\alpha^5} \left(\frac{\gamma}{\gamma - 1} \right)^4 \left(1 - \frac{\gamma - 1}{r\gamma} \right)^4 \tag{29.12}$$

where C_{inc} and C_{com} are the constants for incompressible flow and compressible flow, respectively. These constants are function of fluid regime with ignored parameters, such as AE instruments and data acquisition system and wave transmission path characteristics. Equations 29.11 and 29.12 show that AE energy extremely depends on pressure ratio (r) and flow Mach number (α). Further parameters are constant. Valve dimensions are exhaust valve $d_{1,\text{ex}} = 30.52$ mm and intake valve $d_{1,\text{in}} = 36.24$ mm.

29.3 Experimental Setup

29.3.1 Test Rig and Measurement System

The experimental test rig is shown in Fig. 29.2. Test setup consists of an aluminum cylinder head of a small four-cylinder spark-ignited engine. Bottom of the CH was plugged by a rigid plate and a rubber gasket. Pressurized air interred into

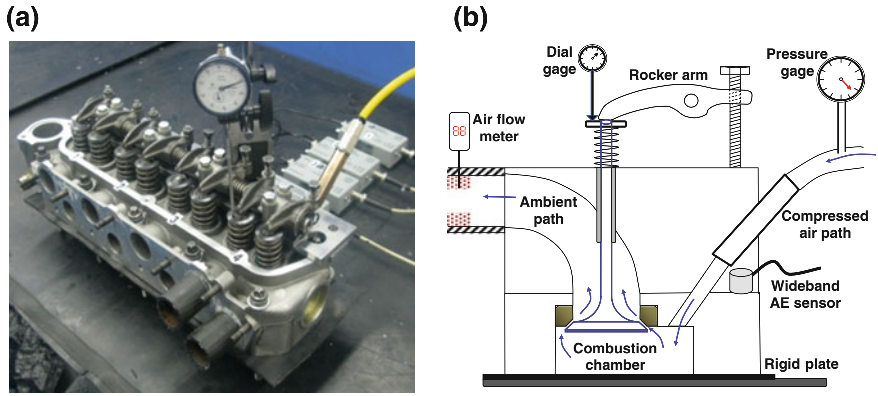


Fig. 29.2 Experimental test rig: **a** cylinder head of a small four-cylinder spark-ignited engine, **b** Schematic of test rig including system measurement and sensor position

combustion chamber through spark plug hole by an intermediate tool. The black rubber tubes in the Fig. 29.2 at the end of the intake/exhaust ports have two functions. The first function is to prepare a fully developed air flow through a larger path, and the second one is to seal air flow meter device. Air flow is measured by a digital air flow meter with an accuracy of 1 lit/min. Valve lift is measured by a dial gage with an accuracy of 0.01 mm.

29.3.2 AE Instruments and AE Sensor Positions

AE signals were acquired using four PAC WSα wideband sensors with an operating frequency range from 100 kHz to 1 MHz. The AE sensors were coupled to the cylinder head by means of thin layers of vacuum grease. The signals were then amplified using four PAC 2/4/6 preamplifiers before being recorded. Then, the signals were digitized by using a PAC PCI-2 AE data acquisition system. Raw data were acquired at sampling rates of 2 MHz via 4 channels. AE sensors were mounted on cylinder head near spark plug hole on every four cylinders as shown in Fig. 29.3.

29.4 Test Procedure and Measurements

Inlet pressurized air interred just only in cylinder no. 4. Therefore, all data such as valve leakage rate, inlet pressure, and valve lift refer to cylinder no. 4. AE sensor no. 4 on cylinder no. 4 was the nearest sensor to all measurements, and sensor no. 1 on cylinder no. 1 was the farthest one.

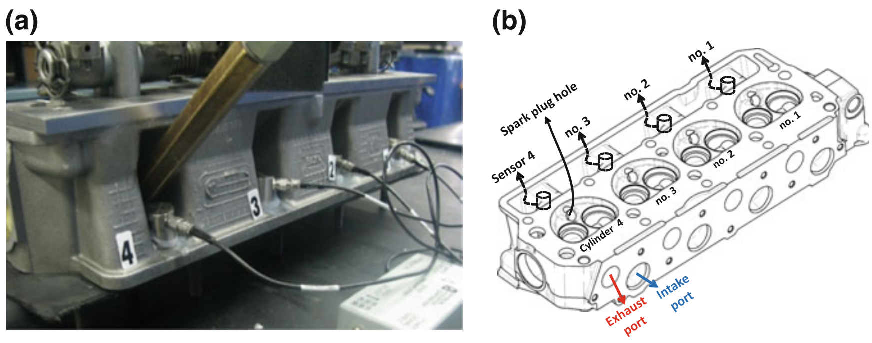


Fig. 29.3 AE sensor positions on cylinder head, **a** overall view of experimental AE sensor positions, **b** schematic of cylinder head including intake/exhaust ports, cylinder arrangement, spark plug hole, and AE sensor position

The leakage of valve has been artificially simulated by very small valve lift. This artificial valve lift is similar to valve wrong clearance state. When a valve clearance is not truly adjusted, valve is open at its closing time.

Valve lift was edited by a bolt's vertical movement very slowly and regularly. After each valve lift step, all parameters were recorded (4 AE sensor, valve leakage rate, valve lift, and inlet air pressure). Test was done for intake and exhaust valve lift individually at three inlet pressures: 2, 4, and 6 bar. At each pressure, valve lift was increased from zero lift state in 10 steps.

29.5 Results and Discussion

29.5.1 Valve Leakage Results

Figure 29.4 shows valve leakage rate of cylinder 4 versus non-dimensional valve lift (L/d1) at 2, 4, and 6 bar inlet pressurized air for exhaust/intake valves. The results show a jump in valve leakage rate at each inlet pressure due to various flow states including incompressible flow at first valve lifts and compressible flow at last valve lifts.

29.5.2 Time-Domain Signal Analysis Results

It was expected that the analysis in time domain reveals the overall signal amplitude, periodic features, and AE signal type (burst or continuous signal).

Figure 29.5 shows the recorded raw AE waveform caused by exhaust valve leakage from sensor no. 4 at 6 bar inlet pressurized air. Frequency sampling is

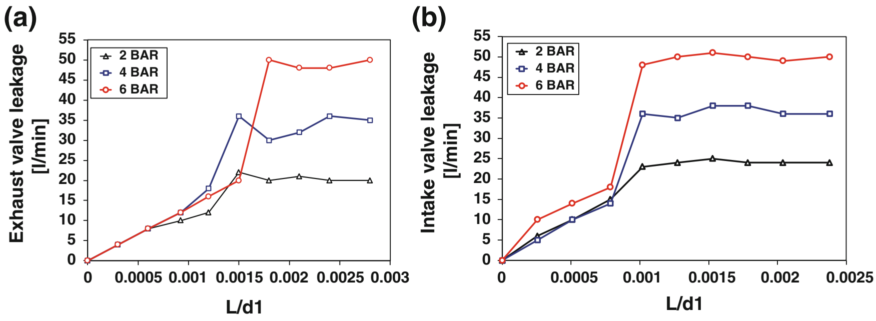


Fig. 29.4 Valve leakage rate of cylinder 4 versus non-dimensional valve lift (L/d_1) at 2, 4, and 6 bar inlet pressurized air. **a** Exhaust valve leakage and **b** intake valve leakage

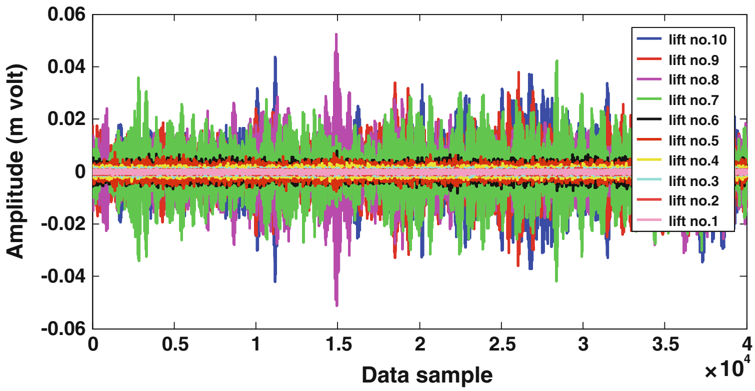


Fig. 29.5 Raw AE signal due to exhaust valve leakage versus time at 6 bar inlet pressure in different valve lifts by sensor no. 4

2 MHz, and Fig. 29.5 presents the AE signal acquired from 0.02 s for 10 exhaust valve lift steps.

As it is clear from Fig. 29.5, there are obvious differences at amplitude and waveform between first 6 steps and last 4 steps. Their differences are due to different fluid flow, incompressible flow, and compressible flow. A similar event is seen at all measurement conditions at different inlet pressures and intake/exhaust valve.

29.5.3 Frequency-Domain Signal Analysis Results

Frequency analysis of the valve leakage AE signal clearly shows the energy distribution of the signal in frequency domain. Figure 29.6 shows the power spectral

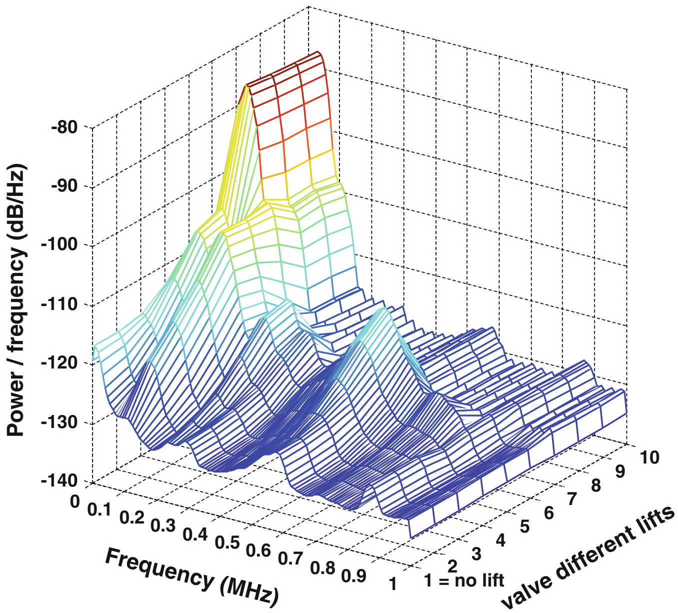


Fig. 29.6 Power spectral density of raw AE signal due to exhaust valve leakage at 6 bar inlet pressure in different valve lifts by sensor no. 4

density of the signals of Fig. 29.5 (AE signals caused by exhaust valve leakage from sensor no. 4 at 6 bar inlet pressurized air).

The results show the relationship between power spectral density (PSD) and valve lifts in frequency domain. Figure 29.6 shows peaks at operating frequency ranges of 40, 110, 230, 540, and 870 kHz for first 6 valve lift steps (incompressible flow). As soon as the flow changes to compressible flow, frequencies of 230 and 540 kHz in PSD of signals do not show important peak amplitudes, but frequencies of 40 and 110 kHz show more significant peak amplitudes. The frequency of 870 kHz has the same amplitude at all tests, which could probably be assumed as a background noise.

29.5.4 AE Average Signal Level of Valve Leakage

Average signal level (ASL) is a measure of the continuously varying and "averaged" amplitude of the AE signal. The ASL parameter is wisely related to the magnitude of source event. The ASL of exhaust and intake valves is shown in Fig. 29.7a and b, respectively. Figure 29.7 shows that the ASL of intake valve has a sharper jump through changing flow state from incompressible flow to compressible flow. This may be due to the different valve geometry and less friction

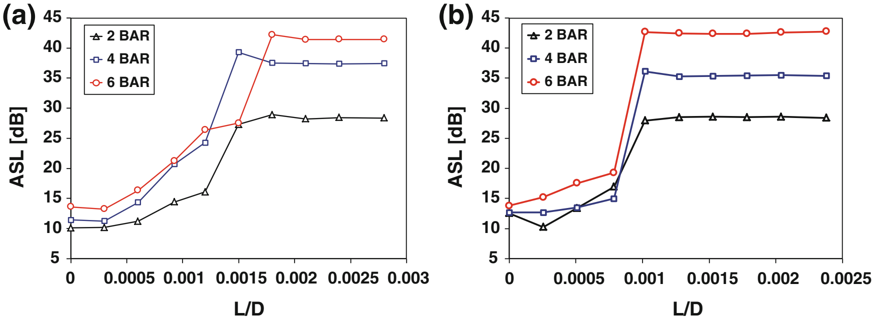


Fig. 29.7 Average signal level of valve leakage versus valve lift by sensor no. 4. **a** exhaust valve leakage and **b** intake valve leakage

force with air jet at intake valve which is due to less contact area. So change in flow state is significant. A common result between intake and exhaust valve is that both of them have approximately fixed ASL after their jump. Such fixation is because of low valve leakage rate changes.

29.5.5 Comparison of Theoretical AE Average Energy and Experiment Results

The AE_{rms} is preferred for understanding the magnitude of source event over counts because it is sensitive to the amplitude as well as the duration, and less dependent on the voltage threshold and operating frequencies [8]. Figure 29.8 shows the results of theoretical model of AE average energy and experimental data for both intake and exhaust valve in different inlet pressures. Theoretical data were calculated from Eqs. 29.11 and 29.12.

Figure 29.8 confirms strong correlation between AE signal energy and valve lift (pressure ratio). Also it shows that compressible flow generates more AE energy than incompressible flow. Another important result is that AE energy generation has a high value in compressible flow state and an approximately decreasing tendency because of lower effects of boundary condition of valve gap on fluid flow.

29.5.6 AE Sensor Location Investigation

As mentioned in part no. 4, all the tests were done on cylinder no. 4 (Fig. 29.2). Four AE wideband sensors were installed on every cylinder (Fig. 29.3). Figure 29.9 shows AE wave transmission between four AE sensors. Also it

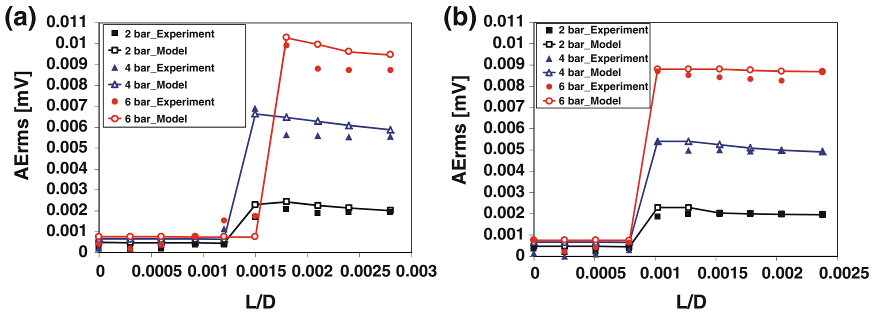


Fig. 29.8 Theoretical and experimental comparison of average AE energy (AE_{rms}) of valve leakage versus valve lift by sensor no. 4. **a** exhaust valve leakage and **b** intake valve leakage

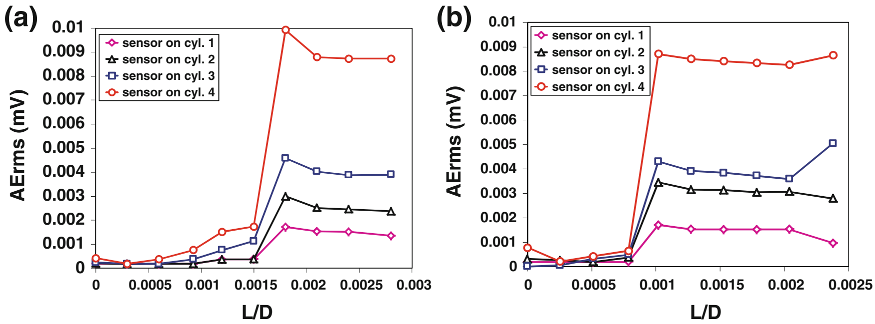


Fig. 29.9 Experimental comparison of average AE energy (AErms) between different sensor locations at 6 bar inlet pressure in cylinder no. 4. **a** Exhaust valve leakage and **b** intake valve leakage

illustrates the effect of sensor location on valve leakage detection on aluminum cylinder head for intake and exhaust valve leakage in cylinder no. 4 with 6 bar inlet pressure. It can be found from Fig. 29.9 that there is a clear difference between sensors at more valve leakage. This is because of more strength of AE signals at more valve leakage. Therefore, the sensor that is nearer to valve leakage position detects more direct signals, while the further ones are affected by more noise due to signal reflections and attenuation.

29.6 Conclusions

This paper presents an experimental investigation of measuring AE signals generated by air leakage through exhaust and intake valves of an internal combustion engine. Average AE energy (AE_{rms}) of the signal presented for valve leakage in cylinder no. 4 versus valve lifts. The results show that AErms and power spectral density (PSD) of AE signal extremely depend on valve leakage rate (valve lift) and flow state (incompressible and compressible flow). Also a new theoretical model was developed for both incompressible and compressible flow regimes. Such model estimates AErms through valve lift. Theoretical model had a good correlation with experiment data.

Acknowledgments This paper was financially supported by IRANKHODRO POWERTRAIN COOPERATION (IPCO). The authors gratefully acknowledge the IPCO engine laboratory researchers for their helps.

References

1. Beattie AG (1983) Acoustic emission, principles and Instrumentation. J Acoustic Emis 2:95–128
2. Murthy CRL, Rao AK, Seth VK, Krishnan A (1983) Acoustic emission monitoring of a nuclear reactor end shield for leak detection. In Proceedings of the non-destructive examination for pressurised components; Monterey, California; USA; pp 29–30 Aug 1983
3. Nivesrangsan P, Steel JA, Reuben RL (2005) Acoustic emission mapping of diesel engines for spatially located time series. Mech Sys Sign Proc 19:1034–1054
4. Wang Y (2007) Introduction to engine valvetrains. SAE international, ISBN: 978-0-7680-1079-4, 2007
5. Kaewwaewnoi W, Prateepasen A, Kaewtrakulpong P (2007) Investigation of the relationship between internal fluid leakage through a valve and the acoustic emission generated from the leakage. Meas 43:274–282
6. Morse PM, Ingard KU (1968) Theoretical acoustics. McGraw Hill, New York
7. White FM (2003) Fluid mechanics. McGraw-Hill, ISBN: 0-07-240217-2
8. Grosse CU, Ohtsu M (2008) Acoustic emission testing. Springer, Berlin, Heidelberg. ISBN 978-3-540-69895-1
9. Shack WJ, Ellingson WA, Youngdahl CA (1980) Development of a noninvasive acoustic leak detection system for large high pressure gas valves. ISA Tran 19:65–71
10. Nakamura T, Terada M (1985) Development of leak monitoring system for pressurizer valves. Prog Nucl Ener 15:175–179
11. Püttmer A, Rajaraman V (2007) Acoustic emission based online valve leak detection and testing. In: Proceedings of IEEE ultrasonics symposium, New York, USA, 28–31 Oct 2007
12. Mathew J, Ma M, Tan A, Weijnen M, Lee J (2011) Engineering asset management and infrastructure sustainability. In: Proceedings of the 5th world congress on engineering asset management (WCEAM 2010), ISBN: 978-0-85729-301-5, Springer, London, 1st Edition
13. Fog TL, Brown ER, Hansen HS, Madsen LB, Sorensen P, Hansen ER, Steel JA, Reuben RL, Pedersen PS (1998) Exhaust valve leakage detection in large marine diesel engines. In: Proceedings of the 11th international conference on condition monitoring and diagnostic engineering management (COMADEM '98) Monash University 1
14. El-Ghamry M, Reuben RL, Steel JA (2003) The development of automated pattern recognition and statistical feature isolation techniques for the diagnosis of reciprocating machinery faults using acoustic emission. Mech Syst Sig Proc 17:805–823
15. Sharkey AJC, Chandroth G, Sharkey NE (2000) Acoustic emission, cylinder pressure and vibration: a multisensory approach to robust fault diagnosis. In: Proceedings of the IEEE international joint conference on neural networks, vol 6. pp 223–228

16. Lin TR, Tan ACC, Joseph M (2011) Condition monitoring and diagnosis of injector faults in a diesel engine using in-cylinder pressure and acoustic emission techniques. In: Proceedings of the 14th Asia pacific vibration conference (APVC 2011) The Hong Kong Polytechnic University, 5–8 Dec 2011
17. Steel JA, Reuben RL (2005) Recent developments in monitoring of engines using acoustic emission. Invited Pap IMechE J Str Anal (special issue on AE) 40:45–57
18. Gill JD, Reuben RL, Scaife M, Brown ER, Steel JA (1998) Detection of diesel engine faults using acoustic emission. In: Proceedings of a 2nd international conference on planned maintenance, reliability and quality, Oxford, vol 1. pp 57–61, 2–3 Apr 1998
19. El-Ghamry M, Steel JA, Reuben RL, Fog TL (2005) Indirect measurement of cylinder pressure from diesel engines using acoustic emission. Mech Sys Sign Proc 19:751–765
20. Douglas RM, Steel JA, Reuben RL (2006) A study of the tribological behaviour of piston ring/cylinder liner interaction in diesel engines using acoustic emission. Trib I 39:1634–1642
21. Gill JD, Reuben RL, Steel JA (2000) A study of small HSDI diesel engine fuel injection equipment faults using acoustic emission. In: Proceedings of the 24th European conference on acoustic emission testing (EWGAE 2000), Paris May 2000
22. Noorul Haq A, Tamizharasan T (2007) Ring wear monitoring in IC engines: an acoustic emission approach. I J Adv Man Tech 31:1148–1155
23. Elamin F, Fan Y, Gu F, Ball A (2009) Detection of diesel engine valve clearance by acoustic emission. In: Proceedings of the computing and engineering researchers' conference, University of Huddersfield, December 2009
24. Buksa A, Tudor M, Martinovic D (2005) Research of the failure incidences in the diesel engine propulsion system. In: Proceedings of the 7th international conference on engine room simulation. Portoroz, Slovenia. 14–15 Nov 2005
25. Fog TL, Brown ER, Hansen HS, Madsen LB, Sorensen P, Hansen ER, Steel JA, Reuben RL, Pedersen PS (1999) On condition monitoring of exhaust valves in marine diesel engines. In: Proceedings of the IEEE workshop on neural networks for signal processing IX, Piscataway, New Jersey, pp 554–563
26. Holford KM, Carter DC (1999) Acoustic emission source location. Key Eng Mat 167–168

Chapter 30
A Comparative Study of Prognostic Model for Bearing Failure

Hack-Eun Kim, Jun-Seok Oh, Andy C. C. Tan and Joseph Mathew

Abstract An effective prognostics program will provide ample lead time for maintenance engineers to schedule a repair and to acquire replacement components before catastrophic failures occur. This paper presents a technique for accurate assessment of the remnant life of machines based on health state probability estimation technique. For comparative study of the proposed model with the proportional hazard model (PHM), experimental bearing failure data from an accelerated bearing test rig were used. The result shows that the proposed prognostic model based on health state probability estimation can provide a more accurate prediction capability than the commonly used PHM in bearing failure case study.

30.1 Introduction

The ability to accurately predict the remaining useful life of machine components is critical for machine's continuous operation and can also improve productivity and enhance system's safety. Recent advances in computing and information technology have accelerated the production capability of modern machines with only reasonable progress that has been achieved in machine fault diagnostics but not in prognostics. Although expert diagnostic engineers have significant information and experience about machine failure and health states by continuously monitoring and analyzing machine condition in industry, unfortunately, well-understood systematic methodologies and supporting systems on how to predict machine remnant life are still not available. Consequently, the prognostic task still

H.-E. Kim (✉) · J.-S. Oh · A. C. C. Tan · J. Mathew
Korea Gas Technology Corporation, 839 Bongsan, Yuseong, Daejeon 305-506, Korea
e-mail: hackun@kogas-tech.co.kr

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_30,

relies on human expert knowledge and experience. Therefore, there is an urgent need to continuously develop and improve prognostic models and algorithms that can be implemented in intelligent maintenance systems with minimum human involvement.

An effective prognostic model requires performance assessment, development of degradation models, failure analysis, health management and prediction, feature extraction and historical knowledge of faults [1]. In general, the machine failure prognostics has a large margin of uncertainty because machine degradation is dynamic and undergoes stochastic process that usually consists of a series of degradation states. Liu et al. [2] suggested the similarity-based method for manufacturing process diagnostics and performance prediction. In their paper, similarities with historical data were used to predict the probabilities of a failure over time by evaluating the overlaps between predicted feature distributions and feature distributions related to unacceptable equipment behavior for long-term prediction of the process performance. However, most of the similarity-based methods only considered two degradation processes namely, a normal process behavior and a faulty process behavior. An accurate and precise prognostic of the time to failure of a failing component or subsystem needs to consider the critical-state variables associated with the changes of physical conditions.

Vachtsevanos et al. [3] suggest that the desire for prognostics has evolved from an increase in diagnostic capability. They strongly emphasize that diagnostics is a prerequisite for prognostics, declaring that "the task of a prognostic module is to monitor and track the time evolution (growth) of the fault". This implies that the fault must be identified prior to attempting prognostics (i.e., diagnostics). An effective prognostic requires performance assessment, development of degradation models, failure analysis, health management and prediction, feature extraction and historical knowledge of faults. In general, each machine system has its inherent characteristics that could be used to identify the source of failure. Therefore, prior analysis of the machine and knowledge of failure pattern could lead to more accurate prediction of remnant life.

Long-term prediction of a fault evolution that may result in a failure requires a tool to manage the inherent uncertainty. Depending on the criticality of the system or subsystem being monitored, various levels of data, models and historical information are required to develop and implement the desired prognostic model. Many accomplishments have been reported but major challenges for long-term prediction of RUL still remain to be addressed. In order to provide long-term and accurate forecasting, an integrated prognostics system that includes full utilization of system degradation data, a well-established failure model, and event history has the potential for practical application in industry. Indeed, uncertainty representation and management is at the core of performing successful prognostics. Long-term prediction of the time to failure entails large-grain uncertainty that must be represented effectively and managed efficiently. For example, as more information about past damage propagation and about future use becomes available, means must be devised to narrow the uncertainty bounds.

In this paper, to achieve long-term prediction of the remnant life of machine, the authors propose a machine prognostic model based on health state probability estimation using a modified SVM classifier. In this model, prior historical knowledge is embedded in the closed-loop prognostic system together with the classification of faults and health state estimation. The historical knowledge includes prior knowledge of the machinery degradation process, failure patterns and maintenance history. This model proposes an integrated approach because an accurate prognostic model requires good diagnostics information. Diagnostics also provides information on obtaining reliable event data and acquiring feedback for system redesign. Therefore, by using an integrated system of diagnostics and prognostics, predetermined dominant fault obtained in the diagnostic process can be used to improve the accuracy of prognostics in estimating the remnant life.

The health state probability estimation is carried out through exploring a full failure degradation process of the machine from new to final failure stages. In terms of the historical knowledge, historical failure data and events are applied to identify the failure patterns. This approach produces an effective feature extraction and the construction of fault degradation steps for impending faults.

To validate the feasibility of the proposed model, two sets of experimental data of bearing run-to-failure were analyzed and employed to predict the remnant life of a pump based on estimation of health state probability using the SVM classifier. In addition, model comparison study was also conducted using the commonly applied PHM. The results show that the proposed prognosis model has the potential to be used as an estimation tool for bearing remnant life prediction.

The remaining part of the paper is organized as follows: Sect. 30.2 presents the proposed diagnostics and prognostics model based on health state probability estimation. Section 30.3 describes the designed experimental test rig for accelerated bearing failure test and how these experimental data are used for validating the prognostic model including prediction results. The comparative prognostics test and results using experimental bearing failure data are addressed in Sect. 30.4. We conclude the paper in Sect. 30.5 with a summary and future research work.

30.2 Integrated Diagnostics and Prognostics Model Based on Health State Probability Estimation

In this research, an innovative diagnostics and prognostics model based on health state probability estimation with embedded historical knowledge is proposed. For an accurate assessment of machine health, a significant amount of a priori knowledge about the assessed machine or process is required because the corresponding failure modes must be known and well described in order to assess the current machine or process performance [4]. To obtain the best possible prediction on the machine remnant life, the proposed prognostics model is integrated with fault diagnostics and empirical historical knowledge. The outcome of a diagnostics module provides reliable information for the estimation of machine health state and system redesign by employing the precise failure pattern of the impending fault. Therefore, by using an integrated system of diagnostics and prognostics, knowledge of a predetermined dominant fault obtained in the diagnostic process can be used to improve the accuracy of prognostics in predicting the remnant life.

In this model, through prior analysis of the historical data and events, major failure patterns that affect the entire life of the machine are identified for diagnostics and prognostics. The historical knowledge provides the key information on diagnostics and prognostics of this system such as empirical training data for the classification of impending faults and historical failure patterns for the estimation of current health state. Moreover, it also could be used to determine appropriate signal processing techniques and feature extraction techniques for effective diagnostics and prognostics.

Figure 30.1 presents the flowchart of the integration of historical knowledge, diagnostic and prognostics system for health state estimation. The proposed system consists of three subsystems, namely, historical knowledge, diagnostics and prognostics. The entire sequence includes condition monitoring, classification of impending fault, health state estimation and prognostics and is performed by linking them to case-based historical knowledge.

Through prior analysis of historical data, the historical knowledge provides useful information for the selection of suitable condition-monitoring techniques, such as sensor (data) type and signal processing techniques, which are dependent on machine fault type. In the proposed model, the feature extraction and selection techniques in the diagnostics module are linked with the historical knowledge. The predetermined discrete failure degradation of the machine located in the historical knowledge module can be used to estimate the health state of the machine located in the prognostics module. The final output of the prognostics module of certain impending faults can also be accumulated to update the historical knowledge. This accumulated historical knowledge can then be used for system updating and improving of the prognostics model by providing reliable posterior degradation features for diverse failure modes and fault types. In this proposed model, the health state probability estimation of discrete failure degradation can be performed using classification algorithms. The authors employed the SVMs classifier for the health state probability estimation because SVMs show outstanding performance in the classification process compared with the other classifiers in our previous work [5].

After identifying the impending fault in the diagnostic module, the discrete failure degradation states determined in prior historical knowledge module are employed in the health state estimation module as depicted in Fig. 30.1. The historical failure patterns also can be used to determine the optimum number of health states for the prediction of the machine remnant life. In estimating the health state, predetermined discrete degradation states were trained before being used to test the current health state. Through prior training of each failure degradation state, current health condition is obtained in terms of probabilities of each

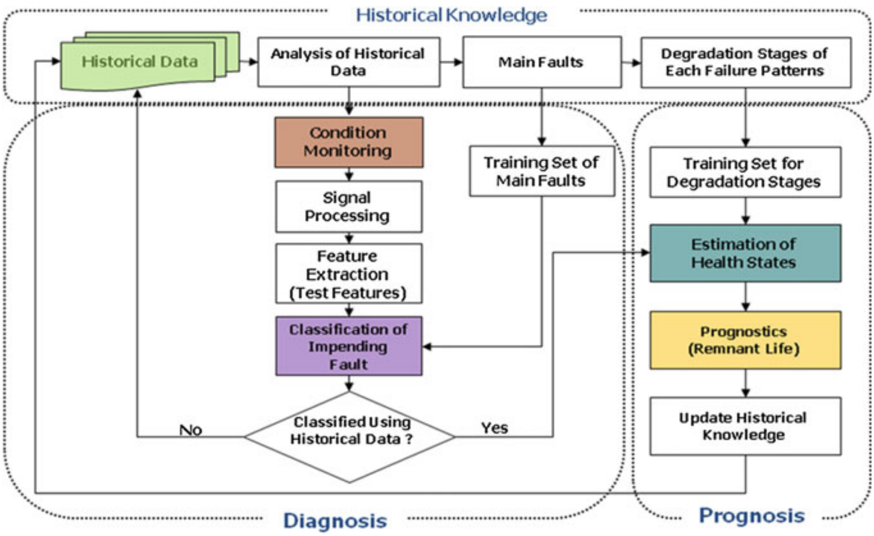


Fig. 30.1 Flow chart of the diagnostic and prognostic system based on health state probability estimation

health state of the machine using the capability of multi-classification. At the end of each prognostics process, the output information will also be used to update the historical knowledge. The following section provides a brief summary of the proposed health state estimation methodology and the RUL prediction using the SVM classifier.

SVM is based on the statistical learning theory introduced by Vapnik and his co-workers [6, 7]. SVM is also known as maximum margin classifier with the abilities of simultaneously minimizing the empirical classification error and maximizing the geometric margin. Due to its excellent generalization ability, a number of applications have been addressed with the machine learning method in the past few years. The theory, methodology and software of SVM are readily available in references [6–9]. Although SVMs were originally designed for binary classification, multi-classification can be obtained by the combination of several binary classifications. Several methods have been proposed, such as "one-against-one", "one-against-all" and directed acyclic graph SVMs (DAGSVM). Hsu and Lin [9] presented a comparison of these methods and pointed out that the "one-against-one" method is suitable for practical use than the other methods. Consequently, in this work, the authors employed the "one-against-one" method to perform the estimation of discrete health states.

In given observations $\vec{x}_t = (x_{t1}, x_{t2}, \ldots, x_{tm})$, m is the number of observations and t is the time index. Let y_t be the health state (class) at time (t) and $y_t = 1, 2, \ldots, n$, where n is the number of health states. For multi-classification of

n-health state (class) event, the "one-against-one" method have $n(n-1)/2$ classifiers, where each classifier is trained on data from two classes. For training data from the ith and the jth classes, SVM solves the following classification problem:

$$\text{minimize: } \frac{1}{2}\left\|w^{ij}\right\|^2 + C\sum_{\mathrm{t}} \xi_t^{ij}\left(w^{ij}\right)^T \tag{30.1}$$

$$\begin{aligned}
&\text{subject to: } (w^{ij})^T\phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij}, && \text{if } y_t = i,\\
&(w^{ij})^T\phi(x_t) + b^{ij} \leq -1 + \xi_t^{ij}, && \text{if } y_t = j\\
&\xi_t^{ij} \geq 0, && j = 1, 2, \ldots, l
\end{aligned}$$

where the training data $\vec{x}_t$ is mapped to a higher dimensional space by function, ϕ, $\phi(x_t)$ is kernel function, (x_t, y_t) is the ith or jth training sample, $w \in R^n$ and $b \in R$ are the weighting factors, ξ_t^{ij} is the slack variable, and C is the penalty parameter. For detailed explanations on the weighting factors, slack variable and penalty parameter can be seen in [6].

There are different methods that can be used in future testing after all the $n(n-1)/2$ classifiers are constructed. After a series of tests, the decision is made using the following strategy: if sign $\left((w^{ij})^T\phi(x_t) + b^{ij}\right)$ says x is in the ith class, then the vote for the ith class is added by one. Otherwise, the jth value is increased by one. Then, the ith class is predicted using the largest vote. The voting approach described above is also called max-win strategy [10]. From the above SVM multi-classification result (y_t), we obtain the probabilities of each health states (S_i) using the smooth window and indicator function (I_i) as following:

$$\text{Prob } (S_t = i|\vec{x}_t, \ldots, \vec{x}_{t+u-1}) = \sum_{j=t}^{t+u-1} I_i\left(y_j\right)/u \tag{30.2}$$

$$I_i(y) = \begin{cases} 0 & y \neq i \\ 1 & y = i \end{cases}$$

where S_t is the smoothed health state and u is the width of the smooth window.

In the given smooth window subset, the sum of each health state probabilities is shown in Eq. 30.3

$$\sum_{i=1}^{m} \Pr(s_t = i|\vec{x}_t, \ldots, \vec{x}_{t+u-1}) = 1 \tag{30.3}$$

After the estimation of current and each health states in term of the probability distributions, the RUL of machine is obtained according to the probabilities of each health state (s_t) and historical operation time (τ_i) at each state and can be expressed as Eq. 30.4 accordingly.

$$RUL(T_t) = \sum_{i=1}^{m} \Pr(s_t = i | \vec{x}_t, \ldots, \vec{x}_{t+u-1}) \cdot \tau_i \tag{30.4}$$

where τ_i is the average remaining life at state i.

30.3 Case Study Using Experimental Bearing Failure Data

30.3.1 Experimental Set-Up and Acquisition of Accelerated Bearing Failure Data

For validation of the proposed prognostic model, bearing run-to-failure tests were performed under controlled load conditions on a specially designed test rig that can facilitate accelerated bearing life tests. This test rig will simultaneously host four test bearings on a shaft driven by an AC motor. Coupling used so that when a bearing fails, it can be extracted and replaced easily without having to move the other bearings on the shaft. A spring was designed to spring load the two middle bearings. The load can be adjusted accordingly by tightening or loosening screw on the spring mechanism. The schematic of the test rig is depicted in Fig. 30.2.

Prognostic experiments with test bearings that are induced with a prominent crack or hole are less likely to develop natural defect propagation in the early stages. Therefore, the accelerated bearing run-to-failure tests were conducted with defect-free condition of bearings and excessive overloading conditions. In this experimental test, SMT 61806 single-row deep groove ball bearings were used for the run-to-failure test at constant 1,300 rpm of rotation speed. Table 30.1 summarizes the bearing specifications.

In this bearing run-to-failure test, two sets of bearing failure data were collected with identical condition for the proposed model validation. The data sampling rate was 250 kHz, and data collections were conducted by a National Instruments LabVIEW program. The two collected vibration data sets are summarized in Table 30.2.

30.3.2 Feature Calculation and Selection

Using vibration and AE data from the experimental test, a total of 28 features were calculated from the time domain and the frequency domain. For outstanding performance of fault classification and reduction of computational effort, effective features were selected using the distance evaluation technique of feature effectiveness introduced by Knerr et al. [11, 12] as depicted below.

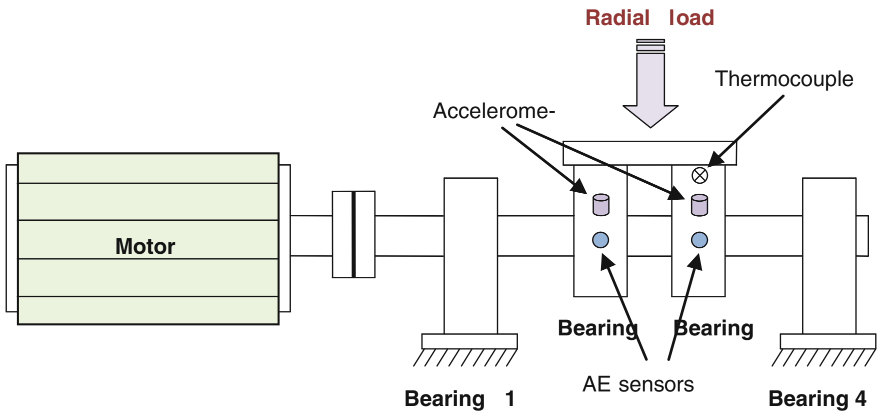


Fig. 30.2 Schematic of the bearing test rig

Table 30.1 Test bearing specifications for experiment

Inner diameter (mm)	Outer diameter (mm)	Width (mm)	Dynamic load rating (kN)	Static load rating (kN)	Fatigue load limit (kN)	Reference speed rating (rpm)
30	42	7	4.29	2.9	0.146	3200

Table 30.2 Experimental bearing failure data set

Test no	Number of sample	Bearing position	RPM	Sampling frequency (K)	Total operation time (min)
1	912	3	1300	250	683
2	810	3	1300	250	579

The average distance $\left(d_{i,j}\right)$ of all the features in state i can be defined as follows:

$$d_{i,j} = \frac{1}{N \times (N-1)} \sum_{m,n=1}^{N} \left| P_{i,j}\, m - P_{i,j}\, n \right|. \tag{30.5}$$

The average distance $\left(d'_{i,j}\right)$ of all the features in different states is

$$d'_{i,j} = \frac{1}{M \times (M-1)} \sum_{m,n=1}^{M} \left| P_{ai,m} - P_{ai,n} \right| \tag{30.6}$$

where $m,\ n = 1,\ 2, \ldots,\ N,\ m \neq n,\ P_{i,j}$ eigenvalue, i data index, j class index, a average, N number of feature, and M number of class. When the average distance

$\left(d_{i,j}\right)$ inside a certain class is small and the average distance $\left(d'_{i,j}\right)$ between different classes is big, these averages represent that the features are well separated among the classes. Therefore, the distance evaluation criteria (α_i) can be defined as:

$$\alpha_i = d'_{ai}/d_{ai}. \tag{30.7}$$

The optimal features can be selected from the original feature sets according to the large-distance evaluation criteria (α_i). Figure 30.3 shows the distance evaluation criterion (α_i) of 28 features in this work. In order to select the effective features, the authors defined a value <1.9 of normalized distance evaluation criterion, $\left|{}^{\alpha_i}/_{\alpha_N}\right| > 1.9$.

From the results, four features were selected as effective features compared with the other features. The four selected features were RMS, entropy estimation value, histogram upper value from vibration data and peak value from AE data. Figure 30.4 presents the trends of each of the selected four features. The trends of the four selected features show the dynamic and stochastic process of the real bearing failure.

30.3.3 Health State Estimation and Prediction of RUL

Through the prior analysis of failure patterns, six discrete degradation stages were determined as the number of health states of bearing failure in this experimental test because they indicated discrete health states relating to bearing failure over the time of test. The prediction tests of bearing failure were performed using the four selected features above. The training data sets for health state estimation are summarized in Table 30.3.

The polynomial function was used as the basic kernel function of SVM. In multi-class classification method using SVMs, the OAO method was applied to perform the health state probability estimation of bearing failure. In this experimental test of bearing failure, closed and open tests were also conducted.

30.3.3.1 Closed Test of Experimental Data

Once the six health states were trained using the four selected features from experimental data 1 as depicted in Table 30.3, the full data sets of data 1 (912 samples) were tested to obtain each health state's probabilities.

Figure 30.5 shows the probabilities of each state of the experimental data 1 that was also used for training of the health states. The probability variation of health state was perceived after 278 samples because an abnormal condition of bearing was detected at this point of time. In general, the abnormal condition of the bearings suddenly occurred at the early stage of defect development and degraded

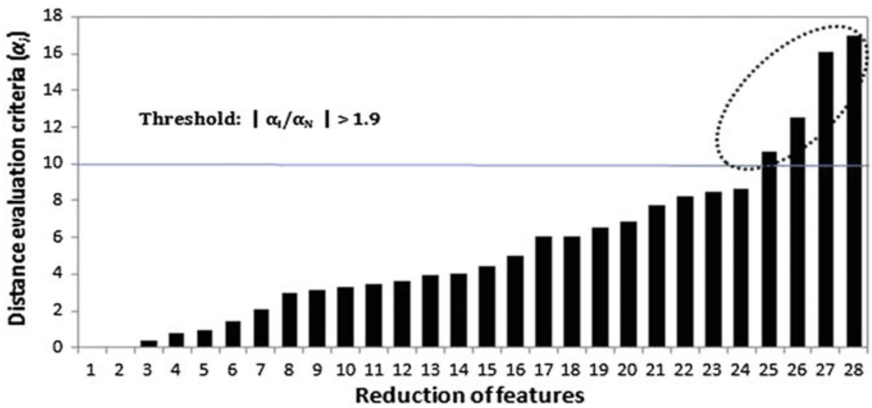


Fig. 30.3 Feature selection using distance evaluation criterion

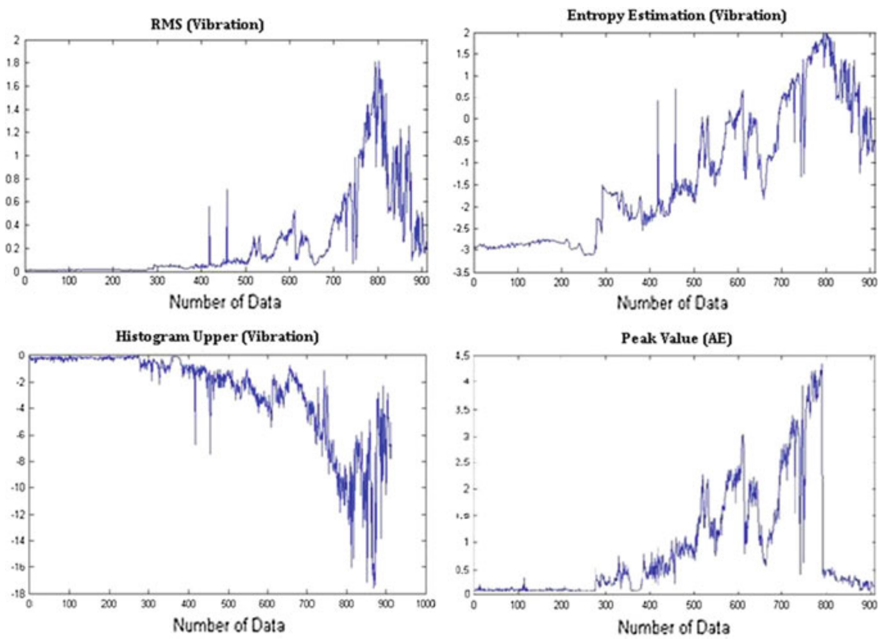


Fig. 30.4 Trends of selected features

rapidly. The probability distribution of the bearing health state effectively presented the transition of bearing conditions as shown in Fig. 30.5. The entire probabilities of each stage explain the sequence of six degradation states after starting at the abnormal condition and are distinctly separated as shown in Fig. 30.5. The training error value was about 1.7 % for the six health states.

Table 30.3 Training data sets for health state probability estimation of experimental test

State no.	No. of samples (u)	Average operation time (τ_i, min)	RUL (%)	No. of features
1	1 ~ 10	9	98.7	4
2	301 ~ 310	571	16.4	4
3	501 ~ 510	608	11.0	4
4	701 ~ 710	645	5.6	4
5	801 ~ 810	663	2.9	4
6	903 ~ 912	682	0.1	4

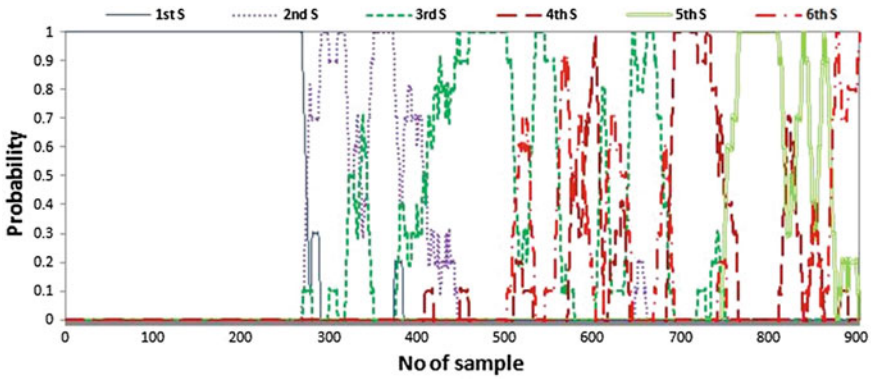


Fig. 30.5 Probability distribution of each health state (closed test)

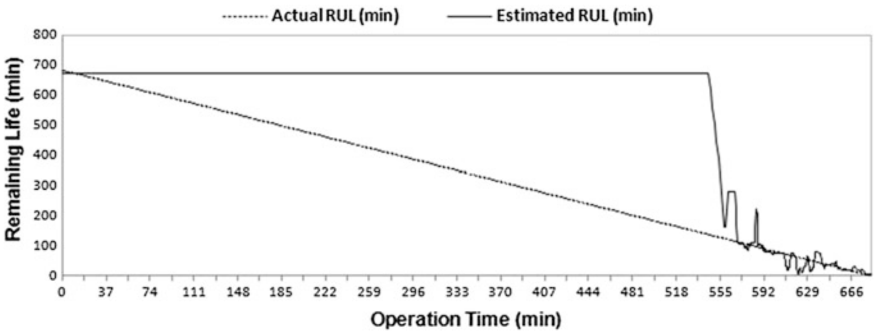


Fig. 30.6 Comparison of actual RUL and estimated RUL (closed test)

The expected life was also calculated by using the time of each training data set (τ_i) and their probabilities of each health state as has be expressed in Eq. 30.4. Figure 30.6 shows the closed test result with comparison between actual RUL and estimated RUL. As shown in Fig. 30.6, there were high margins of error between the actual remaining useful life and the estimated life in the initial state because of the long duration time of the normal condition. However, the estimated life closely followed the actual remaining life after the beginning of abnormal condition

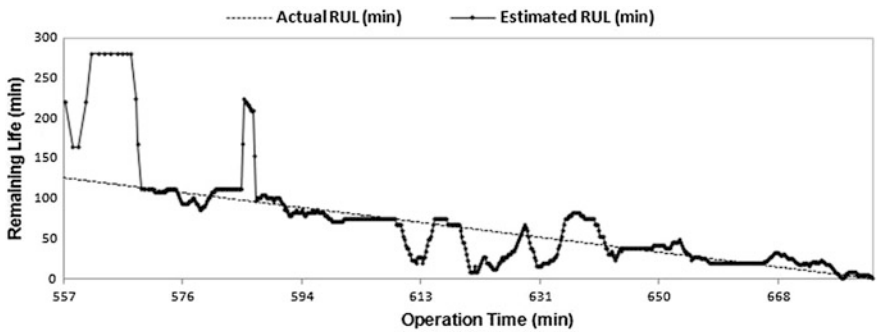


Fig. 30.7 Close view of the period of bearing fault condition (closed test)

(540 min). The accuracy of prediction was also gauged using the Eq. 30.8. The average prediction value was 86.32 % over the entire range of data set.

$$\text{Average prediction accuracy } (\%) = \left(1 - \frac{\sum_{i=1}^{N} \left|\mu_i' - \mu_i\right|}{N}\right) \times 100 \qquad (30.8)$$

where N is number of sample, μ_i' is actual RUL (%) and μ_i is estimated RUL (%).

Figure 30.7 shows the close view of the period of bearing fault condition with comparison between actual remaining useful life and estimated life. After the start of the bearing fault, the estimated RUL was closely matched with the actual remaining life. The average prediction value after the beginning of the abnormal condition (from 540 min) was 97.67 %.

30.3.3.2 Open Test of Experimental Data

The second experimental test data consisted of the 810 sample sets employed for the open test using identical training data as depicted in Table 30.3. Figure 30.8 shows the test results of probabilities of each health state.

As shown in Fig. 30.8, the probability variations began after around 600 samples because an abnormal condition started at the time of about 600 samples in the case of the second bearing test. Compared with the former result (closed test), the probability of five state indicated relatively low values and was hard to find out in the probability distribution. However, the probability distribution of each health state effectively represented the dynamic degradation process of the bearing health state after the beginning of the abnormal condition.

Figure 30.9 shows the comparison between the actual RUL and the estimated RUL. The estimated life of open test (data 2) also started to follow the actual remaining life after the beginning of the abnormal bearing condition. The average prediction value was 38.93 % over the entire range of data set.

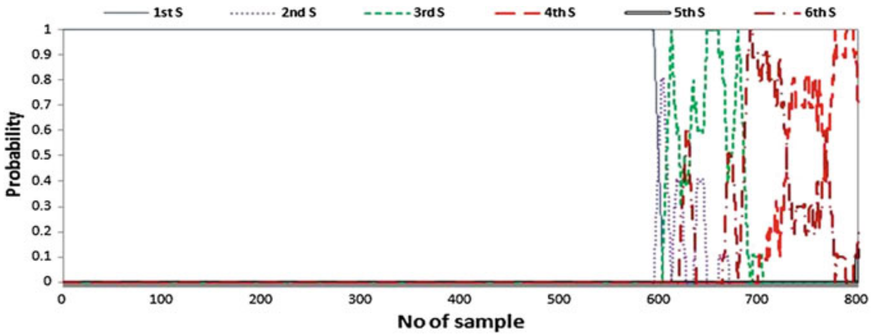


Fig. 30.8 Probability distribution of each health state (open test)

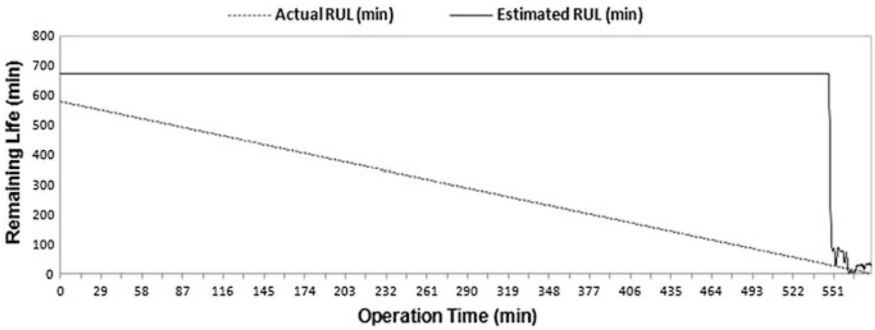


Fig. 30.9 Comparison of actual RUL and estimated RUL (open test)

Figure 30.10 shows the close view of the period of bearing fault condition for the open test. Compared with the result of the closed test as shown in Fig. 30.7, the prediction result showed some low accuracy until after the starting of the abnormal condition at 540 min. Furthermore, the difference between the actual RUL time and the estimated RUL time at initial health state originated from the different life time between training data (first data set, 683 min) and test data (second data set, 579 min) as described in Table 30.3. These results indicate that accurate estimation of health states is achievable for prediction of machine remnant life. Moreover, the proposed model also has the capability to indicate abnormal machine conditions.

30.4 Model Comparison Using PHM

30.4.1 Proportional Hazard Model

The proportional hazard model (PHM), which was originally proposed in the medical research field, can model the uncertain relationships between multiple indicators and time-dependent failure rate. Cox's PH model [13] is a widely

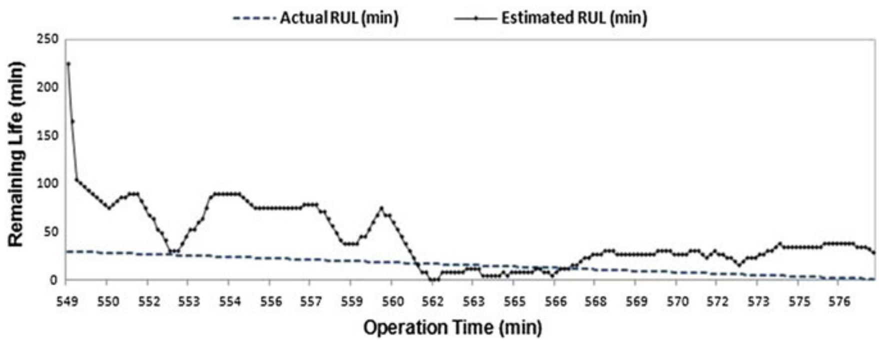


Fig. 30.10 Close view of the period of bearing fault condition (open test)

accepted semi-parametric model for analysis of failures with covariates. It has been successfully used for survival analyses in medical areas and reliability predictions in accelerated life testing. In this case study, to compare the performance of the proposed model, a model comparison was conducted using the commonly used PHM because this model is also performed based on historical failure data.

PHM is developed based on the hazard rate function and assumes that the hazard rate $\lambda(t|\underline{Z})$ under the covariate $\underline{Z}$ is the product of an unspecified baseline hazard rate $\lambda_0(t)$ and a relative risk ratio, $\exp(\underline{\beta}'\ \underline{Z})$, where $\underline{\beta}$ is the regression coefficient vector. The model can be generally expressed as:

$$\lambda(t|\underline{Z}) = \lambda_0(t)\exp(\underline{\beta}'\underline{Z}). \tag{30.9}$$

The significant flexibility of PHM is that the regression coefficients can be estimated by maximizing the corresponding partial likelihood function without specifying the baseline $\lambda_0(t)$. On the other hand, if the baseline hazard $\lambda_0(t)$ is specified, the usual maximum likelihood approach can be carried out to estimate the parameters in the model.

Considering the hard failure by the baseline function $\lambda_0(t)$ and degradation simultaneously, the hazard rate in the form of PHM can be expressed as:

$$\lambda(t|\underline{Z}(t)) = \lambda_0(t)\exp(\underline{\beta}'\underline{Z}(t)) \tag{30.10}$$

where $\underline{Z}(t) = (Z_1(t), Z_2(t), \ldots Z_n(t))$ consists of n degradation features at given time t. Note that the conditional hazard rate $\lambda(t|\underline{Z}(t))$ in Eq. 30.9 is a function of time only. The corresponding reliability function conditional on the history of degradation features up to time t is:

$$R(t|\{\underline{Z}(\tau) : 0 \le \tau \le t\}) = \exp\left(-\int_0^t \lambda(\tau|\underline{Z}(\tau))d\tau\right). \tag{30.11}$$

For failure time distribution, the Weibull distribution is widely used. In a special case, assuming the baseline hazard has the form of two parameters, Weibull yields:

$$\lambda(t|\underline{Z}(t)) = \frac{\gamma}{\eta}\left(\frac{t}{\eta}\right)^{\gamma-1} \exp(\underline{\beta}'\underline{Z}(t)) \tag{30.12}$$

where $\gamma > 0$ and $\eta > 0$ are the shape and scale parameters of Weibull, respectively. The model is referred to as the Weibull PH model. This model is utilized in this case study.

In order to estimate the parameters in the PHM, it is necessary to have the historical data collected under the given operating conditions. The data consist of aging times, feature sample paths and indicators of events (failure versus censored). Then, the likelihood function of the collected data is given by:

$$\begin{aligned} & l\left(\gamma, \eta, \underline{\beta}\middle| Data\right) \\ = & \prod\nolimits_{j\in\Omega_F} \lambda\left(t_j|\underline{Z}\left(t_j\right)\right) \times \prod\nolimits_{k\in\{\Omega_F\cup\Omega_C} R\left(t_k\middle|\underline{\{Z}(\tau) : 0 \le \tau \le t_k\}\right) \end{aligned} \tag{30.13}$$

where Ω_F is the set of failure times, Ω_C is the set of surviving times, t_j is the failure time of the jth unit and t_k is either the failure time or the surviving time of the kth unit. The log-likelihood function can be expressed as:

$$L\left(\gamma, \eta, \underline{\beta}\middle| Data\right) = \sum_{j\in\Omega_F} \Lambda\left(t_j|\underline{Z}\left(t_j\right)\right) - \sum\nolimits_{k\in\{\Omega_F\cup\Omega_C\}} \int_0^{t_k} \lambda(\tau|\underline{Z}(\tau))d\tau \tag{30.14}$$

where $\Lambda\left(t_j|\underline{Z}\left(t_j\right)\right) = \ln\lambda\left(t_j|\underline{Z}\left(t_j\right)\right)$ is the log-hazard rate and the integration is implemented using the adaptive Simpson quadrature rule. The $\hat{\gamma}, \hat{\eta}$ and $\underline{\beta}$ in maximum likelihood estimate (MLE) can be obtained by maximizing the log-likelihood function using Nelder–Mead's algorithm. Then, the MLEs of the reliability indices of interest can be obtained by substituting the MLEs of the model parameters.

30.4.2 Prediction of Remnant Life Using PHM

This comparative study was conducted using the PHM algorithm developed in [14]. Two vibration and AE data sets collected from the bearing test rig as shown in Table 30.3 were also used for the model comparison. For the comparison under identical conditions, the four selected features in the above section such as RMS, entropy estimation value, histogram upper value from vibration data and peak value from AE data were also used for the prediction of RUL using PHM.

The parameters of the PHM were identified using the likelihood function given by Eq. 30.14. In order to obtain a better fit, the features were transformed by taking nature logarithm and denoted by $Z_1(t) = \ln(\text{RMS}(t))$, $Z_2(t) = \ln(\text{Entropy Estimation}(t))$, $Z_3(t) = \ln(\text{Histogram Upper}(t))$ and $Z_4(t) = \ln(\text{Peak}(t))$, respectively. For the PHM:

$$\lambda(t|\underline{Z}(t)) = \frac{\gamma}{\eta}\left(\frac{t}{\eta}\right)^{\gamma-1} \exp(\beta_1 Z_1(t) + \beta_2 Z_2(t) + \beta_3 Z_3(t) + \beta_4 Z_4(t)) \tag{30.15}$$

The MLEs of the parameters of the PHM are presented in Table 30.4.

Using these parameters, the RULs of the bearing failure were estimated, respectively. In the closed test, Table 30.5 presents the prediction results both for the PHM and the proposed model including comparison with the actual RUL after stating abnormal condition of bearing (570 min).

In this table, it can be seen that the estimated RUL from the proposed model are in accordance with the actual remaining life of bearing and outperform the ones from the PHM. Although the estimated RUL from PHM approached the actual RUL closely according to the degradation of the bearing, the prediction of the RUL still has significant difference between the actual RUL and the estimated RUL compared to the results of the proposed model.

Table 30.6 shows the open test result of the second experimental data using identical training data (data 1) after stating of bearing faulty condition. In the case of the second experimental test, it had different bearing degradation pattern with long duration of normal condition (around 540 min) and rapid failure after the start of the faulty condition compared with the first experimental data.

As shown in Table 30.6, the PHM cannot provide accurate prediction results compared with the actual RUL and the results of the proposed model. Although the estimated RULs of PHM matches with the actual RULs as the final bearing failure approaches, the prediction of the RUL still has a significant difference between the actual remaining life and the estimated life shown in Table 30.6. For instance, the PHM still has high estimation error value (109 min) compared with the estimation error of the proposed model (33 min) at the final bearing failure stage (578 min). In this case study, it can be seen that the proposed model provides a more accurate prediction capability than the PHM in these bearing failure cases.

The above prediction result of PHM originates from insufficient historical events in this case study. For better prediction using PHM, extensive data on a substantial failure are required. However, in this case study, only one failure data were available to be used for the prediction test. Moreover, the test data that have considerably different life time from that of the training data can result in large estimation error value in prediction of RUL.

The estimation of survival life using the PHM is based on the prediction of degradation indicators. It is required to predict the degradation features and consider the degradation characteristics [14]. For better prediction result of PHM, stochastic process fitting methods are required for the dynamic and stochastic

Table 30.4 Estimated parameters of PHM using experimental data 1

γ	η	β'_1	β'_2	β'_3	β'_4
2.3988	541.7	0.4717	0.5874	0.0025	3.678e−014

Table 30.5 Comparison of RUL prediction between PHM and proposed model (closed test)

Time (min)	570	580	590	600	610	620	630	640	650	660	670	680	683
Actual RUL	113	103	93	83	73	63	53	43	33	23	13	3	0
Estimated RUL—PHM	357	234	232	210	184	152	120	110	89	66	51	46	47
Estimation error—PHM	244	131	139	127	111	89	67	67	56	43	38	43	47
Estimated RUL—proposed model	112	97	101	71	68	38	45	68	42	20	25	8	3
Estimation error—proposed model	1	6	8	12	5	25	8	25	9	3	12	5	3

Table 30.6 Comparison of RUL prediction between PHM and proposed model (open test)

Time (min)	549	552	555	558	561	564	567	570	573	576	578
Actual RUL	29	26	23	20	17	14	11	8	5	2	0
Estimated RUL—PHM	540	366	280	238	204	178	157	140	126	114	109
Estimation error—PHM	511	340	257	218	187	164	146	132	121	112	109
Estimated RUL—proposed model	337	82	90	79	53	12	12	27	19	34	33
Estimation error—proposed model	308	56	67	59	36	2	1	19	14	32	33

degradations of machine failure. In this comparison study, each bearing degradation features indicated a long and flat region of normal bearing condition before the degradation initiation as shown in Fig. 30.4. Therefore, it is inappropriate to fit the degradation features globally using nonlinear functions of time. The local fitting of degradation features after degradation initiation appears to be more appropriate for nonlinear model fitting in this case.

30.5 Conclusions

In this paper, the proposed prognostics model was validated using experimental bearing run-to-failure data. In addition, model comparison study was also conducted using the commonly applied PHM. For the experimental validation of the proposed model, the accelerated bearing failure test rig was designed and developed. Then, two bearing run-to-failure tests were conducted to obtain the progressive bearing failure data for prognostics. To increase the performance of the SVM classifier and the selection of sensitive degradation features for the health state estimation, effective features were selected using an evaluation method of feature effectiveness.

The results from an experimental case study indicate that accurate estimation of health states is achievable and also provides long-term prediction of machine remnant life. In addition, the results of the experimental test show that the proposed model also has the capability to provide early warning of abnormal machine conditions.

Through the comparison study using PHM, it was verified that the proposed prognostic model based on health state probability estimation can provide a more accurate prediction capability than the commonly used PHM in this bearing failure case study. However, knowledge of failure patterns and physical degradation from historical data for different types of faults still needs further investigation.

Acknowledgments This research was concluded within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centres Program, and also supported by Korea Gas Technology Corporation.

References

1. Lee J, Ni J, Djurdjanovic D, Qiu H, Liao H (2006) Intelligent prognostics tools and e-maintenance. Comput Ind 57:476–489
2. Liu J, Djurdjanovic D, Ni J, Casoetto N, Lee J (2007) Similarity based method for manufacturing process performance prediction and diagnosis. Comput Ind 58:558–566
3. Vachtsevanos G, Frank Lewis, Roemer M, Hess A, Wu B (2006) Intelligent fault diagnosis and prognosis for engineering systems. Wiley, New Jersey
4. Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech Sys Sign Proc 20:1483–1510
5. Kim HE, Tan ACC, Mathew J (2010) Integrated approach for HP-LNG pump diagnostics and prognostics based on health state probability estimation. In: Proceedings of the 5th world congress on engineering asset management (WCEAM-ICF/IQ-AGIC)
6. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York
7. Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Networks 10:988–999
8. Cristianini N, Shawe-Taylor NJ (2000) An introduction to support vector machines. Cambridge University Press, Cambridge
9. Hsu CW, Lin CJ (2002) A comparison of methods for multiclass support vector machines. IEEE Trans Neural Networks 13:415–425
10. He LM, Kong FS, Shen ZQ (2005) Multiclass SVM based on land cover classification with multisource data. In: Proceedings of the fourth international conference on machine learning and cybernetics, pp 3541–3545
11. Hwang WW, Yang BS (2004) Fault diagnosis of rotating machinery using multi-class support vector machines. Korea Soc Noise Vib Eng 14:1233–1240
12. Knerr S, Personnaz L, Dreyfus G (1990) Single-layer learning revisited: a stepwise procedure for building and training a neural network. Springer, New York
13. Cox DR (1972) Regression models and life tables (with discussion). J R Stat Soc 34:187–220
14. Liao H, Zhao W, Guo H (2006) Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model. In: Reliability and maintainability symposium RAMS '06

Chapter 31
Sustainability Improvements Through Efficient Asset Management Networks

Helena Kortelainen, Markku Reunanen and Teuvo Uusitalo

Abstract Asset management supports long-term planning of physical assets and offers a structured approach for more sustainable manufacturing. Asset management aims to improve profitability of the production by supporting life cycle decision making by developing and applying optimal maintenance, capacity planning, and investment strategies. So far the impacts of networked business models and demands for sustainability have not been widely dealt with. The paper presents results from a road map exercise that aimed to identify future needs for sustainable manufacturing models, and focuses on the use of company networks for efficiency of physical assets. The main research questions addressed here are as follows: How does asset management contribute to the demands for more sustainable manufacturing and what implications the demand for more sustainable manufacturing has to asset management strategies?

31.1 Introduction

Asset management supports long-term planning of physical assets and offers a structured approach for more sustainable manufacturing. Asset management aims to improve the profitability and efficiency of the production by supporting life cycle decision making by developing and applying optimal maintenance, capacity planning, and investment strategies. Asset management is closely linked to other governance and management processes as it deals with the creation, maintenance, and improvement of profit-making capacity of production assets, maintaining and optimization of the net asset value of production assets, and improvement of

H. Kortelainen (✉) · M. Reunanen · T. Uusitalo
VTT Technical Research Centre of Finland, P.O. Box 1300, 33101 Tampere, Finland
e-mail: helena.kortelainen@vtt.fi

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_31,

sustainability and safety of asset solutions [1]. In manufacturing industry, the lifetime of production assets is influenced by a variety of exogenous and internal factors as illustrated in Fig. 31.1.

With increasing operation time of a production facility numerous rebuilds, modifications, replacement, and expansion investments typically take place. All of these decisions together with the chosen maintenance strategies and operational models affect the condition and productivity of assets. The life cycle view is reflected, e.g., in the definition given by the European Federation of National Maintenance Societies [3], which states that asset management is "the optimal life cycle management of physical assets to sustainable achieve the stated business objectives."

Sustainable manufacturing objectives are expressed in well-known sustainability's dimensions, i.e., in economic, environmental, and social dimensions. Integrating these factors of sustainability into strategic decision making is an essential prerequisite for moving toward sustainable development. Sustainable manufacturing can be defined as the ability to address limitedness of natural resources, to mitigate the excess of environmental load of manufacturing activities, and to enable an environmentally benign life cycle of products and services, while continuing to improve the quality of human life [4, 5]. When emphasizing long-term profitability, asset management offers a structured approach for more sustainable production [6].

In traditional manufacturing industry, suppliers, lead producers, and customers are seen as independent sequential operators in a value chain, whereas novel collaboration models can rather be described as networks [7, 8]. This development calls for new models for both network governance and performance of manufacturing operations, including also asset management procedures [9]. These models should enable clear identification of the network actors and stakeholders who influence and can be influenced by the sustainability of the product in the course of its life cycle [10]. Moreover, new models should help in integrating sustainability into companies' as well as entire networks' core strategies and thus support the alignment of business processes according to sustainability objectives [11].

Asset management procedures are commonly applied to improve production efficiency in manufacturing industry. What is the significance of asset management as a part of efforts toward more sustainable manufacturing networks? And what are the implications to asset management strategies? The paper presents results from a road map exercise that aimed to identify future needs for sustainable manufacturing models and focuses on efficiency of use of physical assets in company networks. The road map exercise was carried out in the SustainValue project (SustainValue is the acronym for the FP7-funded research project entitled "Sustainable Value Creation in Manufacturing Networks," see www.sustainvalue.eu) that aims to develop industrial models, solutions and performance standards for new sustainable and more performing production and service networks.

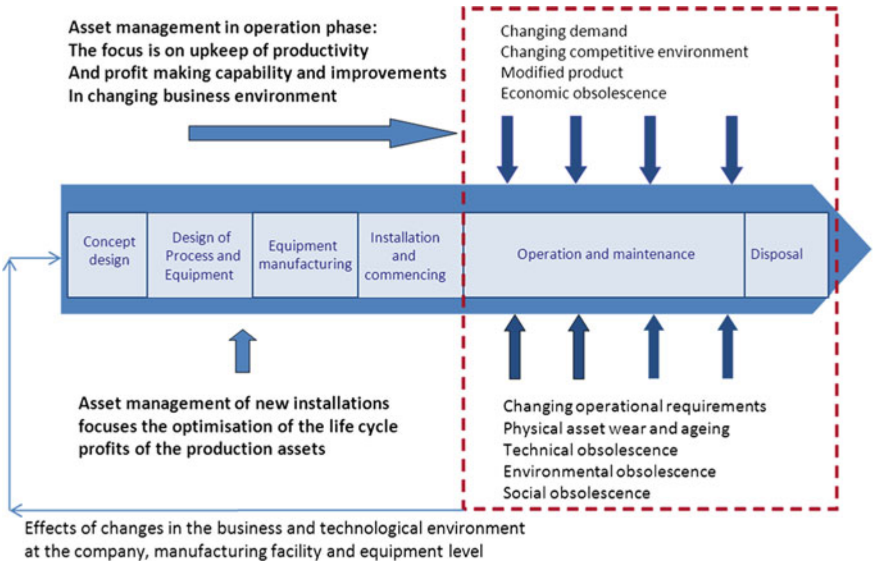


Fig. 31.1 Driving forces to asset management (modified from [2])

31.2 Roadmapping Process

Road maps are regarded as strategic tools for creating deeper understanding and setting agendas for development and change. Visionary socio-technical road maps are visualizations of knowledge based on expert assessment. They combine economic, societal, and technological issues with explicitly stated visions of the future. The road map process is planned to identify elements and issues of development that have strong potential for producing the outcomes that the vision presents. Road maps are not intended to forecast the future in a deterministic way, but they are based on the assumption that future development is likely to include some elements that are presented in these road maps [12, 13].

As the SustainValue project deals with business models and broad concepts like sustainability, the visionary road map process was chosen to guide the work. The chosen time periods were short, middle (5 years), and long term (10 years). The future development was assessed from five perspectives, namely stakeholders, business ecosystem, success criteria, benefits/value, and catalysts/obstacles. The road map process starts by defining a vision which serves as a target against which the current state is compared and the needed changes discussed. The vision for sustainability within the manufacturing industry was stated as *New forms of business models and value networks together enable knowledge-based transformation of the manufacturing industry and improve all three dimensions of sustainable value* (*economic, environmental, and social*) [10].

In the workshop, participants were asked to define the current state of sustainable manufacturing from the different perspectives and then to identify and discuss which changes are required to reach the vision. Furthermore, catalysts and obstacles of the development toward the vision were discussed. A total of 14 experts participated in the workshop on the spot, and two further experts through online from remote locations. The experts presented a wide spectrum of knowledge from industrial enterprises (4), universities and research organizations (11), and a standardization organization. After the expert workshop, the raw data were organized and reworked by the core group of the study, and the material was also iterated and commented within the rest of the expert group.

In the workshop, a Group Support System (GSS) tool consisting of networked computers and dedicated software was used to collect and display participants' responses. With the tool, the input could be used to stimulate new ideas and discussion and they could also be organized and evaluated through prioritizing and voting. One of the strengths of the GSS method is that the tool allows a simultaneous collection of ideas, thus making it possible for every participant to contribute equally. Moreover, when using a computerized GSS, all responses are fully documented in the system, and the data can be easily used also after the workshop (see [14]).

31.3 Findings from the Roadmapping Process

The SustainValue road map exercise aimed to identify future needs for sustainable manufacturing models and address also the efficiency of use of physical assets in company networks. As the road map (see [10]) covers all aspects of sustainability, only a part of identified elements are relevant from asset management point of view. The collected data were analyzed by the authors from this perspective in order to identify those development items and issues that are significant from physical asset's point of view and which could have implications to asset management strategies. The findings are collected in Tables 31.1 and 31.2.

31.3.1 Stakeholders

The willingness to change toward sustainable manufacturing remains an important issue throughout the road map period. Rising energy and raw material prices and increasing awareness of limited natural resources are drivers for a change from price-centerd thinking toward long-term models, especially in investment decisions. It was recognized that all stakeholders should be included in strategic company decisions. The consumption trajectory is expected to change toward

Table 31.1 sustainvalue sub-road map for asset management-related items

Road map level	Timescale		
	Present/Short term	Middle term (5 years)	Long term (10 years)
Stakeholders	Willingness for change Long-term thinking	Willingness for change Proliferation in standards and certifications Flexibility in the use of resources Individual and efficient solutions New cooperation/interfaces with stakeholders Product and service bundles Consequences for all stakeholders should be considered in strategic decision-making	Changing customer behaviour Advanced remote collaboration solutions
Business ecosystem	Cooperation of stakeholders Methods that support the value network's target definition Analyzing and updating current business models	Value network innovation is needed, i.e., innovation with consumers and customers Taking manufacturing processes closer to the end users (e.g., licensing, subcontractors) Adoption of new business models in manufacturing networks more widely	Development of value networks is a way to achieve sustainability goals Adoption for closed-loop business models

(continued)

Table 31.1 (continued)

Road map level	Timescale		
	Present/Short term	Middle term (5 years)	Long term (10 years)
Success criteria	Seeing the benefits of sustainable business	Partnering	Business success would be measured by taking into account all sustainability pillars together with the product value chain and life cycle approach, i.e., measuring performance using triple bottom line approach (economical, environmental, and social value)
	Need for a new, dynamic relationships with customers	Standardized KPIs	
	KPIs/certificates/new performance measures	Sustainability integrated into business functions and into product and business model development in an early stage	
	Changing the focus from products to services		
Benefits/ value	Promoting resource conservation at every stage of the life cycle	New business and markets, more employment	Efficient processes and new technologies for recycling and materials reuse
	Methodology and tools for life cycle data management (e.g., integrating participants in value chain)	Longer lifetime, good spare part, and lifetime service	
		Indicators of operational performance during the development process	

Table 31.2 Identified catalysts and obstacles for a change

Catalysts for a change	Integrated product–service solutions, increasing know-how of new technologies
	Systemic thinking methodology, requirement for transparency, and accountability of solutions
	Scarcity of natural resources and environmental regulations to limit exploitation of natural resources, increased variation in prices of energy and raw material
	Standardization trials in many industries
	End-of-life requirements of products
	Developed markets demand products with a high degree of process-integration
	Education and dissemination
	Rising professionalism in many industries leads to application of new methods and technologies + cross-industry knowledge transfer
	Economic recession, new competition from the emerging new economies
Obstacles for a change	ICT solutions to support transparent efficiency data for all stakeholders during production processes are not available
	Current integration of technical products and services does not support sustainability goals
	Just in time principles have to take sustainability goals into account
	Industrial customers can have problems to comprehend the real value or life cycle perspective when selecting a product
	Global dispersion of manufacturing operations—cultural and managerial differences exist
	Attitudes and resistance against any changes to current methods and business models
	Partnerships although significant for manufacturing companies can create hindrances due to differences in individual objectives
	Short-term perspective on sustainability, profitability, and sustainability collide

sustainable and transparent life cycle of production assets. The demand for services as well as performance-based contracts is expected to increase as customers will consider also the intangible part of a product. As industrial customers pay more attention to all elements of sustainability pillars, they will acquire consulting services to make a better use of the asset, apply advanced maintenance services, and support for the environmentally sustainable disposal of products.

Manufacturing industry uses advanced remote collaboration solutions, and the development of value networks is perceived, an improved way to achieve sustainability goals for the business ecosystem. This could be promoted by standards and certifications which create more transparency and catalyst the development of value networks.

31.3.2 Business Ecosystem

The manufacturing industry is starting to build dynamic relationships with customers, suppliers, and other stakeholders. The expert group expects that manufacturing networks have adopted new business models like product–service systems more widely. The suppliers add information services or decision-making support to their (tangible) products, thus helping to use products in a more sustainable way. Companies are starting to take manufacturing processes closer to the end users. This is implemented by using subcontractors, licensing, and franchising. The development of ICT technology supports this trend. According to the expert group, more profound understanding of business models will be important for gaining mutual benefits. New methods to support the value network's target definition in the long term are required; service contracts as an example. In a long run, development of value networks is perceived as a good way to achieve sustainability goals for the business ecosystem. Industrial actors will create a more holistic approach for looking the whole value chain from a life cycle perspective. Closed-loop business models are implemented. A closed-loop business model includes up-front design of products that can be manufactured using materials reclaimed throughout the manufacturing process and at the end of a product's life.

31.3.3 Success Criteria

There is a need to make sustainability measureable and to define clear and relevant success criteria for each stakeholder. Thus, also the benefits of sustainable business will be more visible. New performance measures and KPIs should be standardized, and they should include, e.g., safety, quality, and environmental impact, use of natural resources, energy, and life cycle cost. In the future, business success is measured by taking into account all sustainability pillars. Companies are able to calculate the value of sustainable image, and sustainability is embedded in all products, services, and lifestyles.

In the supplier industry, focus has changed from products to services supporting the use and maintenance phases of the product life cycle. This change is enhanced by more close and dynamic relationship with the customer and a systemic view and creates a need to change the distribution of knowledge and the management of systems to be maintained and systems needed for maintenance, i.e., systems for logistics support. This creates a need for strategies for the transformation, new business models like leasing, sharing, cooperation, and new revenue streams. This is expected to bring success as product–service providers are specialists and can provide competences, technologies, and resources.

Sustainability demands can be measured during the development, and decision-making tools that take the quantitative evaluation of sustainability into consideration have been developed. This means that the definition of which inputs are required in order to measure the sustainability of the outputs of the different processes of a company is available. Also the life cycle cost and profit decision-making tools are being used during the development process of solutions.

31.3.4 Benefits and Value

Reduction in waste and emissions is considered one of the current benefits of the sustainability, but companies have also realized that sustainable life cycle solutions offer opportunities for more efficient operation and use of resources, new disposal concepts, and cooperation between stakeholders. These factors need to be embedded in the business model in order to avoid considerable short-term costs. Methodology and tools for life cycle data management, consumption of natural resources, and calculation of economic benefit of the environmental friendliness have been introduced integrating different stakeholders.

Equipment life cycles are longer and more performing in respect to raw material and energy consumption. Efficiency can be increased by making the current performance transparent and using indicators of operational performance during the development and innovation process of new solutions. Good spare part service and lifetime service is implemented. The operational efficiency of physical assets has increased. In long term, efficient processes and new technologies for recycling and materials reuse in the end of life of products are well established.

31.3.5 Catalysts and Obstacles

During the workshop session, the catalysts and obstacles for the change toward sustainable manufacturing were also identified (see [10]). The findings related to asset management are listed in Table 31.2.

31.4 Discussion and Conclusions

Asset management is, in its basic nature, an iterative process. The independent development of each sub-area is possible, but the issues are strongly interlinked. As a consequence, it is difficult to link various topics within the timescale. The visionary road map process proved to suit well for studying future trends, challenges, and possibilities of asset management in striving toward sustainable manufacturing. The paper presents results from a road map exercise that aimed to identify future needs for sustainable manufacturing models. It collects the views and opinions of those experts that took part in the workshop and the following rounds of commenting. Therefore, a challenge relating to the method used and to the results is the number and competencies of the experts that participated in the study. As the number of experts was rather small, the results may be regarded as preliminary.

Road map exercise raised up several issues dealing with the key elements of asset management. The trends toward sustainable manufacturing and increasing networking seem to have significant effects also to current asset management practices. 'Cradle to grave' thinking and management of life cycle of physical assets gain increasing importance, new business models restructure the asset management practices in manufacturing companies, services and collaborative models of operation necessitate common targets and standardized KPIs also to asset-related issues, and distributed operations call for novel maintenance solutions. All these changes are expected to create needs for new governance and management methodologies and tools for asset-related decision making.

One of the key issues of asset management and sustainability considerations is *the management of asset life cycle*. Sustainable development also emphasizes recycling, reuse, and remanufacturing of production assets, which makes life cycle data collection even more important than today. Sustainability demands lead to extension of asset life cycles and higher performance and operational efficiency demands. From the asset management point of view, this will pose increasing demands on spare part and lifetime services and probably also require revisions on maintenance strategies and resources.

The emerging trends like *distributed operation models* and developing offering of *intangible service products* offer opportunities for more dynamic customer relationship but create challenges in organizing efficient and effective maintenance function. Technological means like advanced remote maintenance solutions and intelligent logistics systems support local operations requiring skilled workforce and spare parts. In the long run, suppliers will include advanced maintenance services, information services, or decision-making support to their offering. As new business models spread, asset ownership may be totally or partly replaced by performance-based contracts, leasing, sharing, or other forms of collaboration. The arising collaboration models call for common definitions and target setting for asset performance, asset management strategies, and application of new risk sharing models.

With increasing awareness, companies pay more attention to all elements of sustainability. In manufacturing networks, the performance measurement has to be based on commonly accepted and applied measures of *key performance indicators* which are transparent and make it possible to define clear and relevant success criteria for each stakeholder. As asset management strategies typically aim to improve the economically profitable use of assets, key indicators may conflict with this. *Methodologies and tools* for life cycle data management, life cycle cost, consumption of natural resources, and calculation of economic benefit of the environmental friendliness should be introduced in a way that integrates different stakeholders. This leads to profound discussion on the common objectives in the manufacturing networks. More comprehensive methodologies and tools to support asset management-related decisions such as investment, renovation, and maintenance strategy decisions are required.

The main research questions addressed here are as follows: How does asset management contribute to the demands for more sustainable manufacturing and what implications the demand for more sustainable manufacturing has to asset management strategies? Asset management aims to improve profitability of the production by supporting life cycle decision making by developing and applying optimal maintenance, capacity planning, and investment strategies. In this context, demand for sustainable manufacturing seems to pose challenges to current asset management practices. On the other hand, asset management offers structured approaches to achieve sustainability targets.

Acknowledgments The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7, 2007–2013) under grant agreement 262931. The authors wish to acknowledge the European Commission for their support.

References

1. Komonen K, Kortelainen H, Räikkönen M (2012) Corporate asset management for industrial companies: an integrated business-driven approach. In: Van der Lei T, Herder P, Wijnia Y (eds) Asset management. The state of the art in Europe from a life cycle perspective. Springer Science Business Media B.V, Berlin, pp 47–63
2. Komonen K, Kortelainen H, Laakso K, Rosqvist T, Räikkönen M (2005) Physical asset management, a challenge for finnish capital-intensive industry. In Industrial systems review. Espoo. Finland: VTT Industrial Systems. Technical Research Centre of Finland, Tampere, pp 78–84
3. EFNMS (2009) A definition of asset management. http://www.efnms.org/ Accessed 1 Sept 2012
4. Garetti M (2009) Sustainable manufacturing: industry needs and research challenges, quaderni della XV summer school "Francesco Turco" Impianti Industriali Meccanici, Porto Giardino (Italy), 15th–19th Sept 2009
5. IMS2020 (2009) Supporting global research for IMS2020 vision. Intelligent manufacturing systems (IMS) program, funded by the European Commission, http://www.ims2020.org. Accessed 19 Aug 2011
6. Van der Lei T (2012) Towards a research agenda for strategic engineering asset management. In: Van der Lei T, Herder P, Wijnia Y (eds) Asset management. The state of the art in Europe from a life cycle perspective. Springer Science Business Media B.V, Berlin, pp 169–172
7. Normann R, Ramirez R (1994) Designing interactive strategy: from the value chain to the value constellation. Wiley, Chichester
8. Peppard J, Rylander A (2006) From value chain to value network: insights for mobile operators. Eur Manag J 24(2–3):128–141
9. Tywoniak S, Rosqvist T, Maradismo D, Kivits R (2008) Towards an integrated perspective on fleet asset management : engineering and governance considerations. In: Jinji G, Lee J, Ni J, Ma L, Mathew J (eds) 3rd world congress on engineering asset management and intelligent maintenance systems conference (WCEAM-IMS 2008). Beijing, China, 27–30 Oct 2008
10. Palomäki K, Reunanen M, Valkokari K, Valkokari P (2011) Sustainability gaps and stakeholder requirements. Sust Value Deliverable D1:1

11. Valkokari K, Valkokari P, Reunanen M, Palomäki K, Amirmostofian A (2012) Towards sustainability governance in manufacturing networks. Sust Value Deliverable D1:2
12. Ahlqvist T, Carlsen H, Iversen J, Kristiansen E (2007) Nordic ICT foresight: futures of the ICT environment and applications on the Nordic level. VTT Publications, Espoo 653. 147 p. Espoo
13. Ahola J, Ahlqvist T, Ermes M, Myllyoja J, Savola J (2010) ICT for environmental sustainability. Green ICT Roadmap. VTT Tiedotteita—Res Notes 2532:51 p (Espoo)
14. Porthin M, Rosqvist T (2009) Computerised risk analysis workshops. In: Rouhiainen V (ed) Scientific activities in safety and security 2009. Finland VTT, Espoo

Chapter 32
Linkage Between Mobile Materials Management System and Information Quality Dimensions

Sang Hyun Lee, Tae Silk Kim and Abrar Haider

Abstract The importance of materials management systems (MMS) is critical for an asset managing organization because of the role that these systems play in asset life cycle management. However, it is essential that these systems acquire, exchange, and process quality information. It is the quality of information held in MMS that eventually decides the credibility of the decision making regarding materials management, maintenance, and asset life cycle support. At the same time, asset management paradigm has moved toward mobile environment, where the demand of practitioners is that of real-time operation and availability of information. Information quality (IQ), however, has a number of technical, organizational, and people components. Each of these components affects the various IQ dimensions, such as accuracy, consistency, timeliness, and ease of operation. In order to improve IQ, it is important to know the relationship of these dimensions with the characteristics of the system. Therefore, this paper categorizes IQ dimensions according to the requirements associated with major functions of mobile MMS and identifies the types of IQ dimensions that are the most relevant in mobile environments.

32.1 Introduction

In recent years, the significance of materials management system (MMS) has increased for asset managing organizations, as these organizations are using the systems to support asset life cycle at various stages. These organizations are required to cope with the wide range of changes in the business environment, continuously reconfigure manufacturing resources so as to perform at accepted

S. H. Lee (✉) · T. S. Kim · A. Haider
University of South Australia, F1-51A, Mawson Lakes, Adelaide, Australia
e-mail: leesy116@mymail.unisa.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_32,

levels of service, and be able to adjust themselves to change with modest consequences on time, effort, cost, and performance [1].

Since the awareness of critical impact of poor information in information system has been raised throughout all across industries [2], asset managing organizations have been taking information quality (IQ) initiatives to improve their performance and reduce time and money caused by rework in business process [3]. MMS is one of the key components that enable organizations to have sustainable business activities. As for using MMS in asset managing organizations, the biggest issues are handling the materials with correct information and storing the information that is able to be used just in time. The major challenge of MMS faced is maintaining a consistent flow of materials for production. There are many factors that inhibit the accuracy of inventory which results in production shortages, premium freight, and often inventory adjustments. The major issues that all materials managers have are incorrect information of materials, inaccurate cycle counts, unreported scrap, shipping errors, receiving errors, and production reporting errors. Therefore, materials managers have striven to determine how to manage these issues in the business sectors.

In order to provide high quality of MMS, this research conducted case study to identity MMS issues and map them to IQ dimensions to discover root causes of MMS problems. In Sect. 32.2, details of case study and identified MMS problems are explained. Section 32.3 shows how those MMS problems can be mapped with relevant IQ dimensions based on literature review. Section 32.4 describes proposed MMS model to improve quality of information and provides solid environment for MMS. This paper concludes with the discussion of the research and highlights its importance for asset managing organizations.

32.2 Case Study and Identifying MMS Problems

In order to identify the problems in relation to MMS that organizations have, we conducted a case study in a world-class chemical organization. It equips a manufacturing plant focusing on producing polyvinyl chloride (PVC) and developing solar and secondary battery materials. In this research, the organization is referred as company A due to ethical confidentiality.

Case study is an ideal methodology when a holistic, in-depth investigation is required [4]. Yin [5] identified six primary sources of evidence for case study research (i.e., documentation, archival records, interviews, direct observation, participant observation, and physical artifact). As the interview is one of the most important sources of case study information to acquire perceived inferences, we conducted interviews with the focused group who are responsible to materials management in company A from March to July 2010.

The identified problems from company A as to MMS can be fold in three broad categories: technology, human, and operation. While the technology problems are related to IT systems, the human problems are associated with manual or improper

Table 32.1 Categorized materials management system problems in company A

Category	Problem
Technology	Analysis of current materials status
Human	Inconsistency
Operation	Gap between online and off-line system
	Defective information

process. The operation problems are linked with materials flow or business process. The problems in each category are shown in Table 32.1.

32.2.1 Analysis of Current Materials Status

When analysis of their actual materials status is required for future demands, company A has been faced following two issues.

- Inaccurate information between purchasing system and real inventory status
- Gap between real inventory status and MMS status

Those two issues are technology-oriented problems. As purchasing system tends to focus on materials that are stocked, real inventory status often does not match with the purchasing system. This is because the purchasing system and MMS are managed by different departments, respectively, and when they installed the purchasing system, company A did not concern about synchronization issues between the two systems.

32.2.2 Inconsistency

While processing materials, manual information generating at work site and automatic MMS information updating are applied concurrently. This causes following problem.

- Inconsistency between off-line documents and MMS information caused by manual handling

As information is entered to the off-line documents manually, this is more vulnerable to information deficiency than MMS's procedures. This means that if information in inventory has poor information, the results of monitoring or auditing contain wrong or improper information. As a consequence, inefficient materials monitoring process and erroneous decision making are conducted. Moreover, it directly affects company A's profit due to redundant or lack of materials.

One of the reasons that manual information handling generates inconsistency information is on difficult table fields in a materials process form. For example, the table field name of materials code starts with 'EKPO-MATNR.' This code comes from legacy systems, and it is difficult to hand in manual operation.

32.2.3 Gap Between Online and Off-line System

MMS should provide an integrated view of materials management through processing of information. However, the gap between online and off-line system in company A causes following problems.

- Complicated process throughout materials flow line
- Mismatch between process order and MMS information
- Inefficient materials monitoring process

Monitoring and auditing of current materials status are mandatory to all different types of businesses. While company A conducts monitoring or auditing their materials, they have relied on information that MMS provides. Here, the gap between real inventory status and MMS status is occurred because information is generated by manually in the real inventory site. In contrast, MMS stores information automatically using a synchronized device that is provided by materials provider.

32.2.4 Defective Information

The chance of defective information generating exists throughout materials processes. As to the poor information in MMS, particular process generates defective information more. Company A points major reason as follows.

- Management difficulty of return materials information

When the return of materials task is occurred in company A, the materials have to be returned to the original inventory through complex return processes. Due to the process complexity, materials are exposed to be placed in wrong places. Due to the complicated materials flow and defective information between off-line documents and MMS, it is difficult to set up standardization of process and incorrect decision making is taken and rework.

32.3 Linkage Between Mobile-MMS and IQ Dimensions

With the results of open-ended questions from the interview participants, the authors discovered that the root causes of MMS-related problems are on IQ. In order to resolve these problems, this section reviews IQ dimensions and maps them to MMS-related problems.

32.3.1 IQ Dimensions

Wang and Strong [6] define an "IQ dimensions" as a set of IQ attributes that represent a single aspect or construct of IQ. The research about IQ dimensions and their attributes has been conducted by many studies [7–9]. The IQ literature provides a thorough classification of IQ dimensions; however, there are a number of discrepancies in the definition of most dimensions due to the contextual nature of quality. The six most important classifications of quality dimensions are provided by Wand and Wang [10]; Wang and Strong [6]. According to Wang and Strong [6], each of IQ dimensions has their own IQ category such as intrinsic IQ, contextual IQ, representational IQ, and accessibility IQ. Intrinsic IQ indicates that information has quality in its own right. Contextual IQ contains the requirements that IQ must be considered within the context of the task at hand for adding value. Representational IQ emphasize that information system should present information clearly to information users. Accessibility IQ highlights the importance of access and security toward information.

Stvilia et al. [11] derived taxonomy of IQ dimensions for IQ measurement framework. In their research, IQ dimensions broadly were categorized to intrinsic, relational, and reputational IQ dimensions. The taxonomy is comprised of 22 IQ dimensions. Firstly, intrinsic IQ category includes IQ dimensions that are related to IQ attributes which can be assessed by measuring IQ characteristics according to given constraints rather than context dependency. Secondly, since information could be affected by external constraints in a given context, relational IQ category represents IQ dimensions that are measured by relationships between information and some aspects of its usage context. Thirdly, reputational IQ category contains one dimension, authority that measures the position of an information entity in a cultural or activity structure.

Redman [12] argued that defect-free IQ dimensions are generally not enough for IQ. Information must embrace the extra dimensions beyond defect-free-related dimension. In his research, he points out that high quality of information means they are fit for their intended uses in operations, decision making, and planning. Kahn and Strong [13] introduced the product and service performance model for IQ (PSP/IQ) integrating definitions from the Total Quality Management literature especially classify quality as conformance to specifications and quality as meeting or exceeding customer expectations, with the product and service aspects of

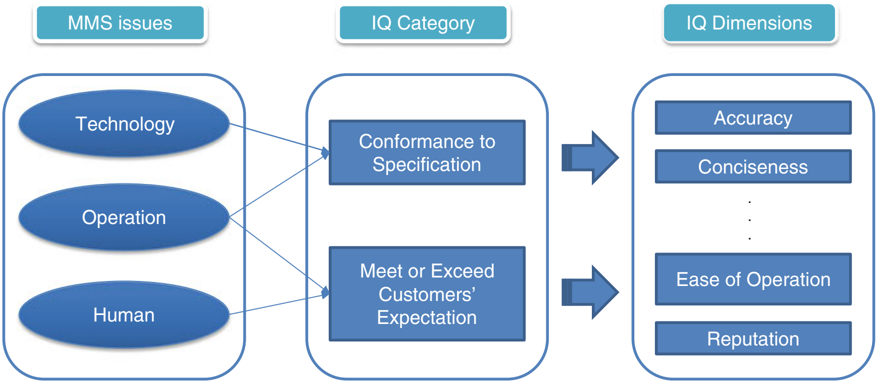


Fig. 32.1 Mapping with MMS issues with corresponding IQ dimensions

information. The IQ dimensions categorized in their model represents four distinctive information characteristics: sound information, useful information, dependable information, and usable information.

Even though IQ dimension-related literature provides guidance of how IQ dimensions can be categorized and identified according to context or business demand, still there is no general agreement exists either on which set of dimensions defines the quality of information, or on the exact meaning of each dimension [6].

32.3.2 Mapping with MMS Issues with Corresponding IQ Dimensions

The fundamental problems of MMS in company A are categorized to technology, operation, and human-related issues as shown in Fig. 32.1. Those MMS issues can be mapped to two IQ categories, i.e., 'conformance to specification' and 'meet or exceed customers' expectation' to embrace objective and subjective perspective of IQ.

The IQ dimensions are categorized by the 'conformance to specifications,' i.e., more product quality-oriented and measureable objectively; and the 'meets or exceeds customers' expectation,' i.e., more service quality-oriented and capturing the essence of 'fitness for use,' respectively. In order to map corresponding IQ dimensions, the dimensions are based on PSP/IQ Model [13] and its mapping results are described in Table 32.2.

Table 32.2 Corresponding IQ dimensions

MMS issues	MMS issues—details	IQ category	Corresponding IQ dimensions
Technology	Gap between real inventory status and MMS status	Conformance to specification	Timeliness
	Inaccurate information between purchasing system and real inventory status		Accuracy, timeliness
Human	Inconsistency between off-line documents and MMS information caused by manual handling	Meet or exceed customers' expectation	Interpretability, consistency, ease of operation
Operation	Complicated process throughout materials flow line	Conformance to specification, meet or exceed customers' expectation	Ease of operation
	Mismatch between process order and MMS information		Timeliness, relevancy
	Inefficient materials monitoring process		Ease of operation
	Management difficulty of return materials information		Ease of operation, conciseness

32.4 Proposed Model of MMS

After mapping between MMS issues and corresponding IQ dimensions, the authors could extract corresponding defective IQ dimensions (e.g., timeliness, accuracy, interpretability, consistency, ease of operation, relevancy, and conciseness). The selected IQ dimensions are identified according to the relationship between definitions of IQ dimensions derived by Wang [6] and detailed MMS issues.

When we look at the overall listed IQ dimensions, 'timeliness' and 'ease of operation' have the most frequency. The 'timeliness' dimension is information age that is appropriate for the task at hand [6]. This concerns life cycle of information and how long the information values remain valid. The issues regarding 'timeliness' are on the gap between real inventory status and MMS status, inaccurate information between purchasing system and real inventory status, and mismatch between process order and MMS information. Those problems' root causes are on different information updating time, and it brings synchronization issues with other systems (i.e., purchasing system and MMS, off-line documents, and MMS). Due to the different system operations, it affects 'ease of operation' IQ dimension as well. In addition, the authors could find out that other dimensions (i.e., accuracy, interpretability, consistency, relevancy, and conciseness) are also affected by synchronization issues. Therefore, the authors derived three solutions to fulfill the requirements of each mentioned IQ dimensions. The derived three solutions are mobile environment setting, introduction of bar-code system, and synchronized MMS building.

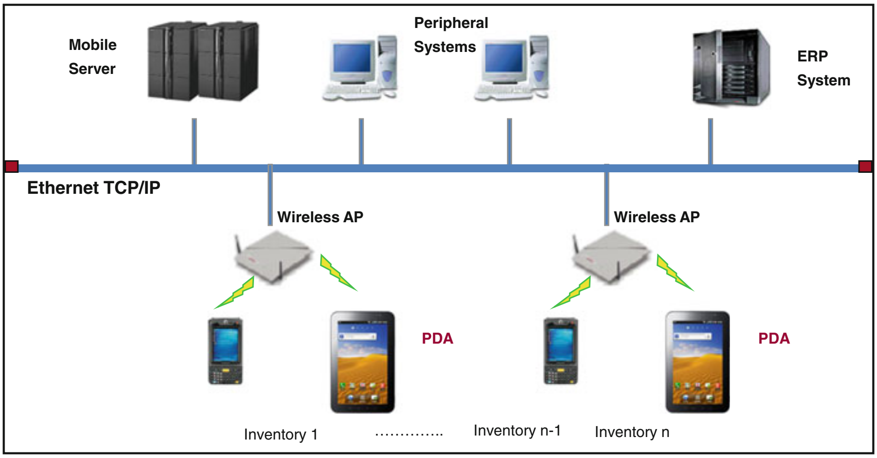


Fig. 32.2 H/W configuration

32.4.1 Mobile Environment Setting

Off-line documents that are used at a materials inventory are generated manually and processed within particular interval. Therefore, real-time information updating is impossible. Moreover, they are needed to fulfill electronic documents requirements for information storing.

Wan Satirah et al. [14] point out that in digital environment, off-line documents need to possess certain dimensions to be transformed by electronic documents. The required dimensions are 'static,' 'authority,' 'unique,' and 'authentic.' Off-line documents at a materials inventory often violates 'static' dimension as it is manipulated by on-site workers to correct their mistaken information entering. In order to minimize manual information entering and support real-time information updating, we designed mobile environment for MMS. Figure 32.2 illustrates the hardware configuration for mobile-MMS environment.

To minimize the risk of manual information entering, we applied two types of on-site mobile devices: Android- or iOS-based mobile devices and industrial PDA (personal digital assistant) device. The on-site mobile devices are operated under Wi-Fi technology, and they are connected with mobile servers and peripheral systems through Ethernet TCP/IP. Under mobile environment, real-time information updating from different sites can be ensured and improve quality of 'timeliness' and 'ease of operation' dimensions. The details of software configuration are shown in Fig. 32.3.

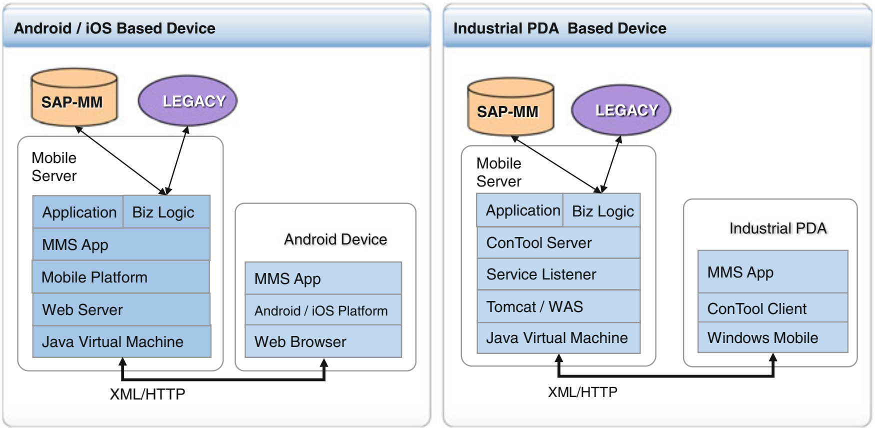


Fig. 32.3 S/W configuration

32.4.2 Introduction of Bar-Code System

By adding a bar code on purchasing order form, the purchasing system is able to synchronize with MMS throughout materials flow. Bar-code information stored for each materials attributes includes item number, date and time of issues, and appropriate places to be loaded. In comparison with existing system, flow of materials can be traced by bar code. Thus, inconsistency of information between purchasing system and MMS can be prevented, and materials status in inventory can be updated by scanning of the bar code. This also prevents incorrect manual information entering. Figure 32.4 shows actual purchasing order form with the bar code and its application to materials inventory.

32.4.3 Synchronized MMS Building

Under mobile-MMS environment using mobile devices and bar code, real-time-based information updating is possible. This means that all relevant system with MMS can be tightly linked each other and those linked systems enable simultaneous information updating. More importantly, this synchronized MMS provides intuitive operation to those who work in on-site areas. Due to the complexity of materials process, person in charge of materials often generates wrong information. In comparison with this, the Synchronized MMS is operated by personal hand-held mobile device by scanning bar code. Therefore, they do not need to concern about difficult information entering as well as required information updating interval. The synchronized MMS operation is illustrated in Fig. 32.5.

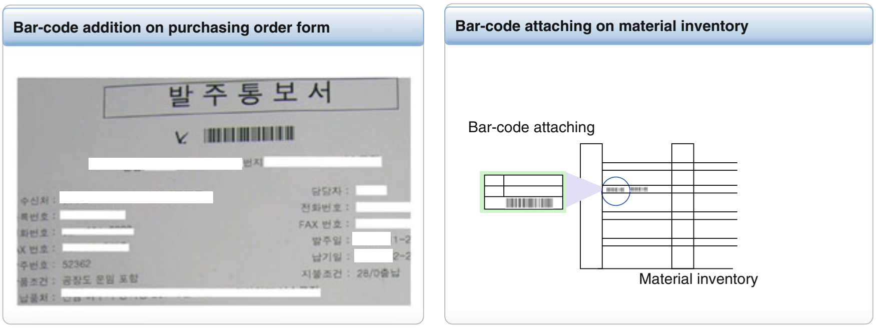


Fig. 32.4 Bar-code addition for mobile-MMS

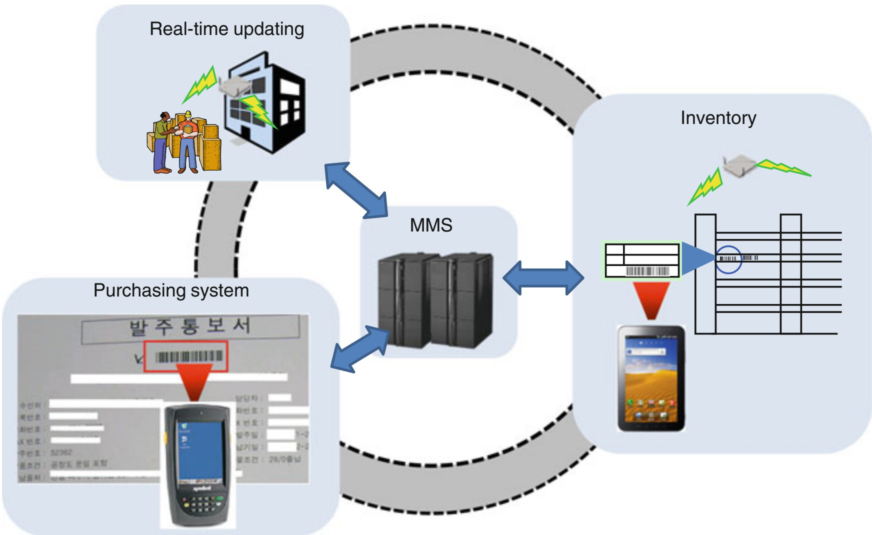


Fig. 32.5 Synchronized MMS operation

32.5 Conclusion

In this paper, we have identified the MMS problems in company A. The identified MMS problems are then attempt to have IQ perspectives and mapped those issues to correspond IQ dimensions. From the case study, we discovered that major reasons of MMS problems are on synchronization issues and manual handling of information. In order to prevent the deficiency, the authors proposed three solutions, i.e., mobile environment setting, introduction of bar-code system, and synchronized MMS building, and those are implemented to company A. As to the future research direction, since the whole implementation of IQ dimensions was not conducted because of wide range of IQ dimensions, prioritizing IQ dimensions according to MMS requirements is required for extracting more considerable IQ dimension and sophisticated mapping.

References

1. Haider A (2011) Information technologies implementation and organizational behaviour: an asset management perspective. In: Technology management in the energy-smart world: PICMET 2011 conference. Portland, Oregon, USA
2. Elizabeth P, Beverly K, Helina M (2006) A comparison of quality issues for data, information, and knowledge. In: Proceedings of the 2006 information resources management association conference, 17th IRMA international conference, Washington, DC, USA
3. Al-Hakim L (2007) Information quality management: theory and applications. Idea Group Pub, Hershey, PA
4. Feagin J, Orum A, Sjoberg G (1991) A case for case study. University of North Carolina Press, Chapel Hill, NC
5. Yin R (1994) Case study research: design and methods, 2nd edn. Sage Publishing, Thousand Oaks, CA
6. Wang RY, Strong DM (1996) Beyond accuracy: what data quality means to data consumers. J Manage Info Syst 12(4):5–33
7. Ballou DP, Pazer HL (2007) A framework for the analysis of error in conjunctive, multi-criteria satisficing decision processes. Decision Sci 21(4):752–770
8. Redman TC (1996) Data quality for the information age. Artech House Boston, MA
9. Huang KT, Lee YW, Wang RY (1998) Quality information and knowledge. Prentice Hall PTR, Upper Saddle River, NJ
10. Wand Y, Wang RY (1996) Anchoring data quality dimensions in ontological foundations. Commun ACM 39(11):86–95
11. Stvilia B, Gasser L, Michael B, Linda C (2007) A framework for information quality assessment. J Am Soc Inform Sci Technol 58(12):1720–1729
12. Redman TC (2000) Data quality: the field guide. Digital Press
13. Kahn BK, Strong DM, Wang RY (2002) Information quality benchmarks: product and service performance. Commun ACM 45(4):184–192
14. Wan Satirah WMS, Haider A, Ku Bahador KM (2012) Records management using information and communication technologies. In: 18th IBIMA conference on innovation and sustainable economic competitive advantage: from regional development to world economies

Chapter 33
Evaluating the Time-Varying Mesh Stiffness of a Planetary Gear Set Using the Potential Energy Method

Xihui Liang, Ming J. Zuo and Yangming Guo

Abstract Time-varying mesh stiffness is a periodic function caused by the change of the number of the contact tooth pairs and the contact positions of the gear teeth. In this study, we have derived the analytical equations of the time-varying mesh stiffness of a planetary gear set using the potential energy method. Three simulations are conducted with a common planetary gear set under fixed carrier, fixed ring gear, and fixed sun gear. The results indicate that the obtained time-varying mesh stiffness can reflect the stiffness variation, and the proposed approaches can be extended in the future to model the stiffness of a planetary gear set when faults are introduced.

33.1 Introduction

Planetary gears are widely used in aeronautic and industrial applications due to the properties of compactness and high torque-to-weight ratios [1]. Research on planetary gears always has gained attention. In the study of gear vibration, mesh stiffness variation is one of the major sources [2]. The finite element method (FEM) and the analytical method (AM) can be used to evaluate the time-varying mesh stiffness. FEM is time-consuming because one needs to model every meshing gear pair in order to know the mesh stiffness of a range of gear pairs. On the other hand, AM can offer a general approach to evaluate the mesh stiffness.

X. Liang · M. J. Zuo (✉)
Department of Mechanical Engineering, University of Alberta,
Edmonton, AB T6G2G8, Canada
e-mail: ming.zuo@ualberta.ca

Y. Guo
School of Computer Science and Technology, Northwestern Polytechnical
University, 127 Youyi West Road Street, Xi'an 710072, China
e-mail: yangming_g@nwpu.edu.cn

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_33,

In terms of AM, a square waveform has been used [1, 2] to approximate the variation of the time-varying mesh stiffness. The square waveform mesh stiffness was represented by the summation of a mean value component and a time-varying component. However, there was no general way presented in [1, 2] on how to get the values of the mean components and the magnitude of the time-varying components. These values were just assumed. Besides, the square waveform method ignored the variation of the mesh stiffness caused by the change of the contact point. As a result, the reported results generate unwanted frequency components due to the flatness of the varying stiffness. The approach to be used in this study aims to overcome these shortcomings.

In this study, the potential energy method is used to derive the time-varying mesh stiffness of a planetary gear set. The potential energy method is initially proposed by Yang and Lin [3] to calculate the mesh stiffness of a pair of fixed-shaft external–*external* spur gears. Later, Tian [4] improved this method by introducing an additional energy term called the shear energy. We believe that this same approach used by Tian [4] can be used to derive the meshing stiffness expression of a planetary gear set. In their studies [3, 4], all the gears are assumed to be involute spur gears and the deflection of the gear body is ignored. The same assumptions will be applied in this study.

For a planetary gear set, three different configurations are possible (most commonly used): fixed carrier, fixed ring gear, or fixed sun gear. The time-varying mesh stiffness for each case is developed in this study. A paper documenting detailed results has been submitted to a journal for possible publication [5].

33.2 Stiffness of Fixed-Shaft External–*External* Gears

Tian [4] calculated the mesh stiffness for a pair of fixed-shaft external–*external* gears. The total mesh stiffness is decomposed into four items: Hertzian contact stiffness k_h, bending stiffness k_b, axial compressive stiffness k_a, and shear stiffness k_s. For a pair of meshing gears whose contact ratio is between 1 and 2, the alternation of one pair and two pairs of teeth in contact is observed.

For the single-tooth-pair meshing duration, the total mesh stiffness is expressed as [4]:

$$k_t(\theta) = \frac{1}{\frac{1}{k_\mathrm{h}} + \frac{1}{k_\mathrm{b1}} + \frac{1}{k_\mathrm{s1}} + \frac{1}{k_\mathrm{a1}} + \frac{1}{k_\mathrm{b2}} + \frac{1}{k_\mathrm{s2}} + \frac{1}{k_\mathrm{a2}}} \tag{33.1}$$

where k_t is a function of the rotation angular displacement of the pinion (θ), and subscripts 1 and 2 represent the pinion and the gear, respectively.

For the double-tooth-pair meshing duration, there are two pairs of gears meshing at the same time. The total mesh stiffness is expressed as [4]:

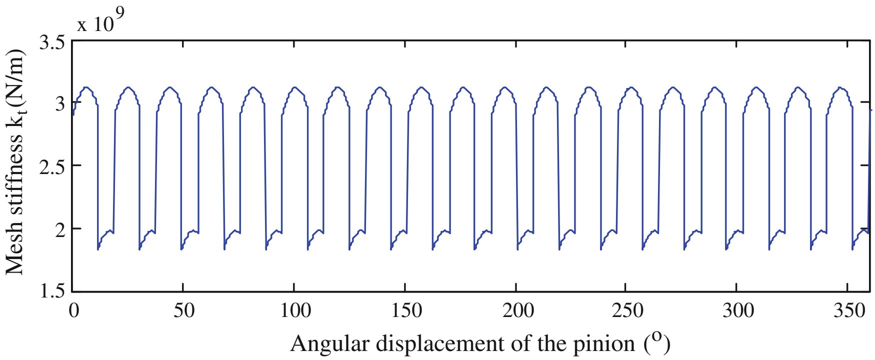


Fig. 33.1 Mesh stiffness of a pair of fixed-shaft external–*external* gears

Table 33.1 Physical parameters of a planetary gear set

Parameters	Number of teeth	Module (mm)	Pressure angle	Face width (m)	Young's modulus (Pa)	Poisson's ratio
Sun gear	19	3.2	20	0.0381	2.068×10^{11}	0.3
Planet gear	31	3.2	20	0.0381	2.068×10^{11}	0.3
Ring gear	81	3.2	20	0.0381	2.068×10^{11}	0.3

$$k_t(\theta) = k_{t1}(\theta) + k_{t2}(\theta) = \sum_{i=1}^{2} \frac{1}{\frac{1}{k_{h,i}} + \frac{1}{k_{b1,i}} + \frac{1}{k_{s1,i}} + \frac{1}{k_{a1,i}} + \frac{1}{k_{b2.i}} + \frac{1}{k_{s2,i}} + \frac{1}{k_{a2,i}}} \quad (33.2)$$

where $i = 1$ for the first pair of meshing teeth and $i = 2$ for the second pair.

Figure 33.1 illustrates the time-varying mesh stiffness of a pair of fixed-shaft external–*external* gears. The parameters of the pinion and the gear of the fixed-shaft external–*external* gear pair are the same as the sun gear and the planet gear (see Table 33.1). There are 19 times of gear meshing in one revolution of the pinion as the number of teeth of the pinion is 19.

These results reported in Ref. [4] will be used later in the derivation of the mesh stiffness of a planetary gear set.

33.3 Stiffness of a Pair of Fixed-Shaft External–*Internal* Gears

Although the mesh stiffness of a pair of fixed-shaft external–*external* gears has been analytically derived, the general mesh stiffness of a pair of external–*internal* gears has not been reported. Based on the potential energy method, the analytical expression of the mesh stiffness of a pair of fixed-shaft external–*internal* gears is derived in this section, in which, Hertzian contact, bending, axial compressive, and shear stiffness will be considered.

33.3.1 Hertzian Stiffness

According to Hertzian law, the elastic compression of two isotropic elastic bodies can be approximated by two paraboloidal bodies in the vicinity of the contact point [6]. For the planet–ring contact, we approximate the planet gear as a cylinder and the ring gear as a circular groove. Similar to the derivation of the Hertzian stiffness of a pair of external–*external* gear meshing [6], the Hertzian stiffness of a pair of external–*internal* gear meshing can be expressed as follows:

$$k_{\mathrm{h}} = \frac{\pi EL}{4(1 - v^2)} \tag{33.3}$$

where E, L, v present Young's modulus, the width of the tooth, and Poisson's ratio, respectively and k_{h} is the Hertzian stiffness that has the same expression as the external–*external* gear meshing [6].

33.3.2 Bending, Shear, and Axial Compressive Stiffness

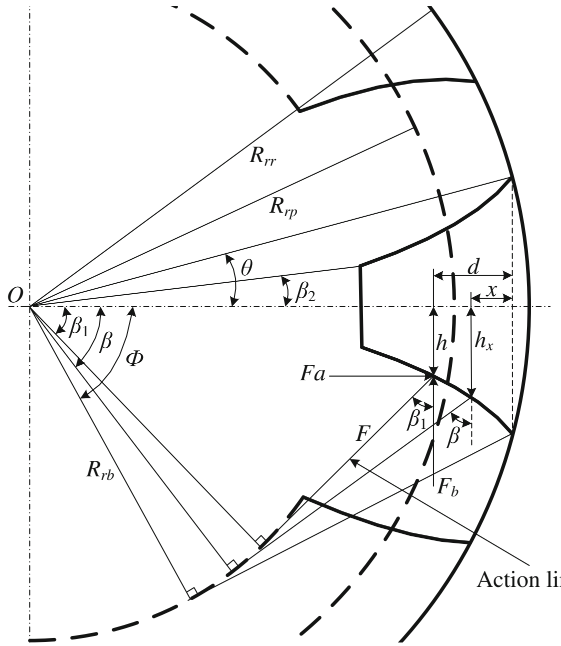
The expressions of the bending, shear, and axial compressive stiffness of the external gear were described in Ref. [4]. Here, based on the potential energy method, aiming at the external–*internal* gear pair, new time-varying mesh-stiffness-evaluating equations of the bending, shear, and axial compressive stiffness of the internal gear are provided below.

Based on the beam theory, the bending stiffness k_{b}, shear stiffness k_{s} and axial compressive stiffness k_{a} of an internal gear tooth can be expressed as follows:

$$\frac{1}{k_{\mathrm{b}}} = \int_{\phi}^{\beta_1} \frac{3\{1 + \cos\beta_1[(\beta_2 - \beta)\sin\beta - \cos\beta]\}^2(\beta_2 - \beta)\cos\beta}{2EL[\sin\beta + (\beta_2 - \beta)\cos\beta]^3}\mathrm{d}\beta \tag{33.4}$$

$$\frac{1}{k_{\mathrm{s}}} = \int_{\phi}^{\beta_1} \frac{1.2(1 + v)(\beta_2 - \beta)\cos\beta\cos^2\beta_1}{EL[\sin\beta + (\beta_2 - \beta)\cos\beta]}\mathrm{d}\beta \tag{33.5}$$

Fig. 33.2 Elastic force on an internal gear tooth



$$\frac{1}{k_{\mathrm{a}}} = \int_{\phi}^{\beta_1} \frac{(\beta_2 - \beta)\cos\beta \sin^2\beta_1}{2EL[\sin\beta + (\beta_2 - \beta)\cos\beta]} \mathrm{d}\beta \tag{33.6}$$

where $\beta_1 = \tan(\arccos\frac{N_2\cos\alpha_0}{N_2-2}) + \frac{\pi}{2N_2} - \mathrm{inv}\alpha_0 + \frac{N_1}{N_2}\psi$, β_2 is the half tooth angle on the base circle, N_1 and N_2 are the numbers of teeth of the external gear and the internal gear, respectively, and $\phi = \beta_2 + \tan(\arccos\ (R_{\mathrm{rb}}/R_{\mathrm{rr}}))$ (see Fig. 33.2).

33.3.3 Overall Effective Mesh Stiffness

Similar to the external–*external* gear meshing, the alternation of one pair or two pairs of teeth in contact will be experienced by the external–*internal* gear meshing. Equations 33.1 and 33.2 can also be used here to calculate the total mesh stiffness of the single-tooth-pair meshing duration and the double-tooth-pair meshing duration, respectively. Finally, the total mesh stiffness can be expressed as:

$$k_t = k_t(\theta) \tag{33.7}$$

where θ is the angular rotation displacement of the external gear and k_t is a function of θ.

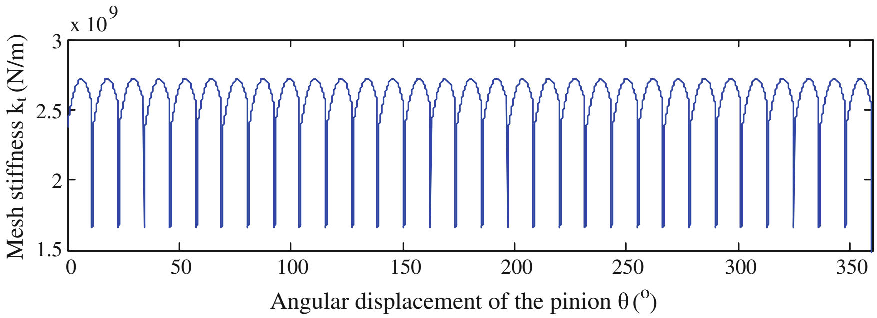


Fig. 33.3 Mesh stiffness of a pair of fixed-shaft external–*internal* gears

Figure 33.3 shows the mesh stiffness of a pair of fixed-shaft external–*internal* gears. The parameters of the external gear and the internal gear of the fixed-shaft external–*internal* gear pair are same as the planet gear and the ring gear (see Table 33.1). There are 31 times of gear meshing in one revolution of the external gear as the number of teeth of the external gear is 31.

33.4 Stiffness of a Planetary Gear Set

Because the planetary gear set may have different structures, its total mesh stiffness can be considered in three cases: fixed carrier, fixed ring gear, or fixed sun gear. In this section, we will discuss the time-varying mesh stiffness under these three cases. The simulations are conducted with a common planetary gear set, which has one sun gear, one ring gear, and four equally spaced planet gears.

Parker and Lin [7] pointed out that the sun–planet meshes are not necessary in phase with each other while each of them has the same shape of mesh stiffness variation. Similar comments apply to the ring–planet meshes in this paper. Here, firstly, we use the results in Ref. [7] to calculate the relative phases of the planetary gear set with the same parameters in Table 33.1 and the results are shown in Table 33.2. The value of γ_{sn} ($n = 1, 2, 3, 4$) is the relative phase between the nth sun–planet pair with respect to the first sun–planet pair. The value of γ_{rn} is the relative phase between the nth ring–planet pair with respect to the first ring–planet pair. The value of γ_{rs} is the relative phase between the nth ring–planet mesh with respect to the nth sun–planet mesh. If the value of γ_{rs} equals to 0, it means that the sun–planet gear pair and the ring–planet gear pair mesh at the pitch point simultaneously.

Table 33.2 Relative phases of the planetary gear set

γ_{s1}	γ_{s2}	γ_{s3}	γ_{s4}	γ_{r1}	γ_{r2}	γ_{r3}	γ_{r4}	γ_{rs}
0	0.75	0.5	0.25	0	0.25	0.5	0.75	0

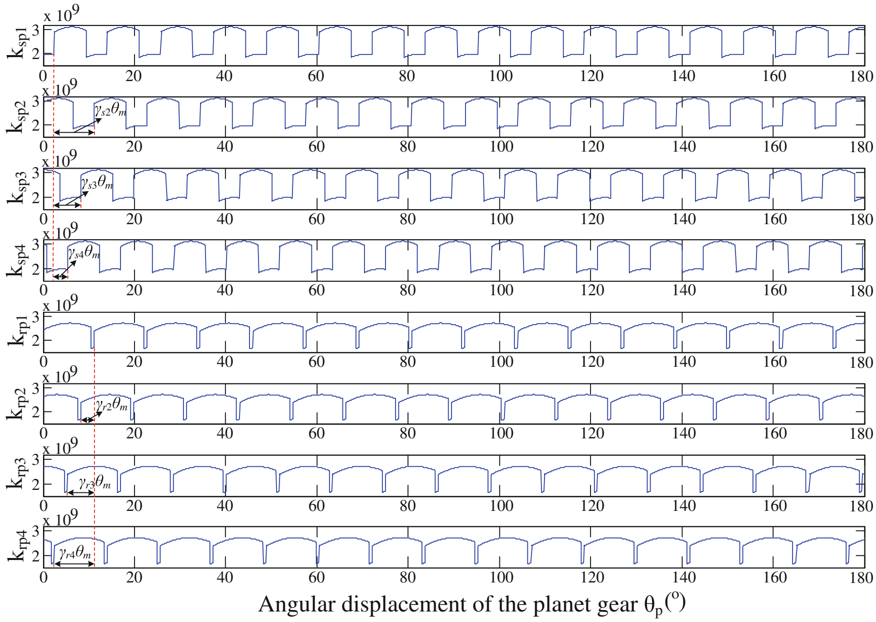


Fig. 33.4 Mesh stiffness of a planetary gear set when the carrier is fixed

33.4.1 Case 1: Fixed Carrier

For a planetary gear set, if the carrier is fixed, the sun–planet gear meshing is equivalent to a pair of fixed-shaft external–*external* gear meshing, while the ring–planet gear meshing is equivalent to a pair of fixed-shaft external–*internal* gear meshing. The time-varying mesh stiffness is calculated with the same parameters given in Table 33.1, and the results are shown in Fig. 33.4. The mesh stiffness shapes of different pairs of sun–planet (or ring–planet) meshing are the same except for the phase differences (The value of θ_m is the meshing period.). The number of meshes is 15.5 ($Z_p/2$) in half revolution of the planet gear.

After, in Fig. 33.4, the lines k_{sp1}, k_{sp2}, $k_{sp3,}$ and k_{sp4} represent the first, second, third, and fourth pairs of the sun–planet mesh stiffness, respectively. Similarly, the lines k_{rp1}, k_{rp2}, $k_{rp3,}$ and k_{rp4} denote the first, second, third, and fourth pairs of the ring–planet mesh stiffness, respectively.

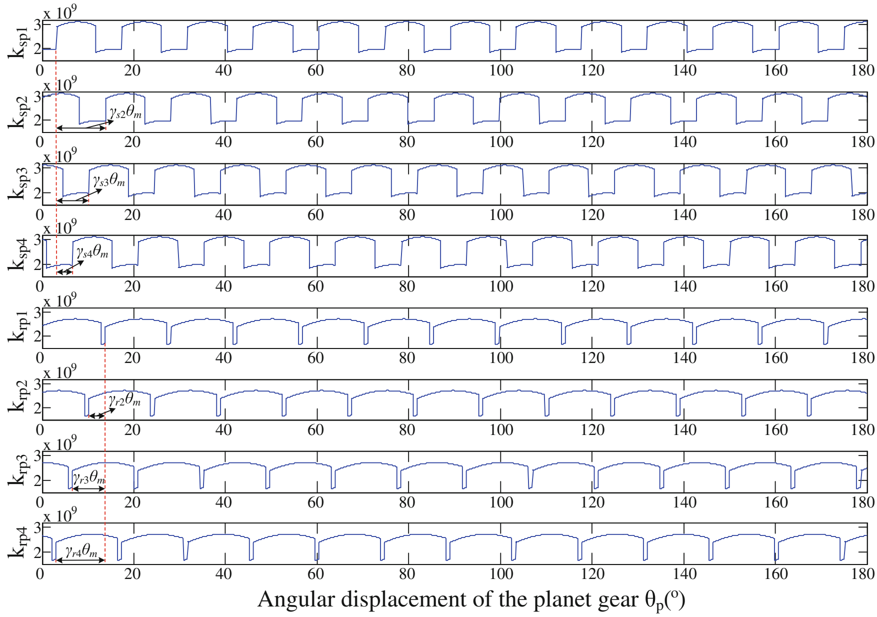


Fig. 33.5 Mesh stiffness of a planetary gear set when the ring gear is fixed

33.4.2 Case 2: Fixed Ring Gear

If the ring gear is fixed, the gear mesh frequency ω_m is smaller than the situation when the carrier is fixed with the same rotation speed of the sun gear. The mesh frequency ratio of a planetary gear set between the cases of ring gear fixed and of the carrier fixed can be expressed as: $\lambda_r = Z_r/(Z_s + Z_r)$ [8]. The mesh stiffness shapes are the same for the two situations except for the difference of the meshing periods. When the ring gear is fixed, the mesh stiffness can be obtained from an expansion of the shape of the mesh stiffness when the carrier is fixed:

$$k_{sp1} = k_t(\lambda_r \theta_p) \tag{33.8}$$

The time-varying mesh stiffness when the ring gear is fixed is depicted in Fig. 33.5. Here, we have also used the same parameter values shown in Table 33.1. We can see from Fig. 33.5 that 12.56 times ($\lambda_r Z_p/2$) of gear meshing are experienced in half revolution of the planet gear.

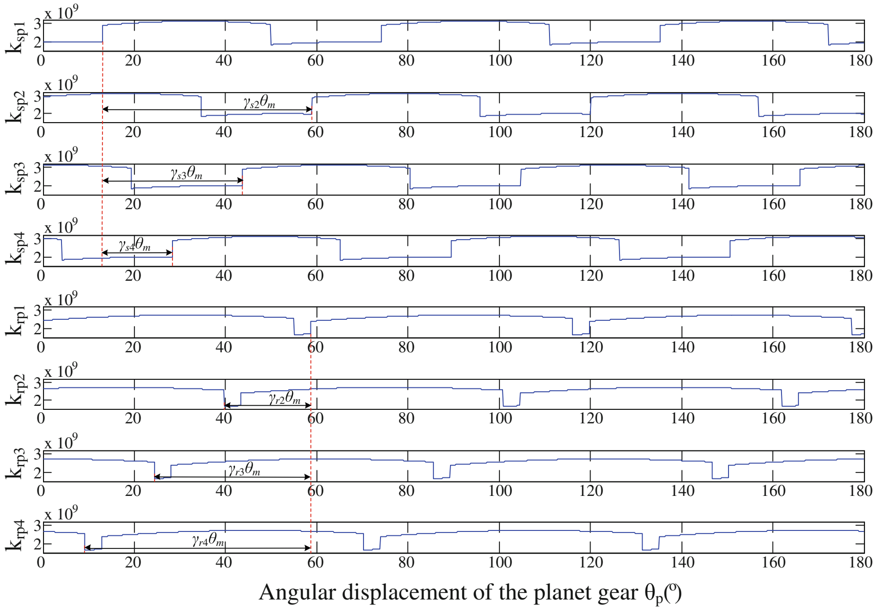


Fig. 33.6 Mesh stiffness of a planetary gear set when the sun gear is fixed

33.4.3 Case 3: Fixed Sun Gear

In this case, the mesh frequency ratio of a planetary gear set between the cases of fixed sun gear and of fixed carrier can be expressed as: $\lambda_s = Z_s/(Z_s + Z_r)$ [8]. The mesh stiffness with the fixed sun gear can be derived through an expansion of the shape of the mesh stiffness when the carrier is fixed:

$$k_{\mathrm{sp1}} = k_t(\lambda_s \theta_{\mathrm{p}}) \tag{33.9}$$

The time-varying mesh stiffness is illustrated in Fig. 33.6 using the same parameters in Table 33.1. The number of gear meshing in half revolution of the planet gear is 2.945 ($\lambda_s Z_p/2$).

33.5 Summary

In this study, the potential energy method is used to evaluate the time-varying mesh stiffness of a planetary gear set. The total potential energy is considered as a summation of Hertzian energy, bending energy, shear energy, and axial compressive energy in this approach. Three cases of fixed carrier, fixed ring gear, and fixed sun gear of a planetary gear set are considered. The results indicate that the obtained time-varying mesh stiffness can reflect the stiffness variation, because it is caused not only by the change of the number of the contact tooth pairs but also by the change of the contact position of the gear teeth. Furthermore, this approach can be further extended for stiffness evaluation when gears experience damages.

Acknowledgments This research is supported by the Natural Science and Engineering Research Council of Canada (NSERC), the China Scholarship Council (CSC), and the Reliability and Maintenance Research Laboratory.

References

1. Chaari F, Fakhfakh T, Haddar M (2006) Dynamic analysis of a planetary gear failure caused by tooth pitting and cracking. J Fail Anal Prev 6:73–78
2. Lin J (2000) Analytical investigation of planetary gear dynamics. Ph.D thesis, The Ohio State University, Columbus, Ohio, USA
3. Yang D, Lin J (1987) Hertzian damping, tooth friction and bending elasticity in gear impact dynamics. J Mech Transm-T ASME 109:189–196
4. Tian X (2004) Dynamic simulation for system response of gearbox including localized gear faults. Master's Thesis, University of Alberta, Edmonton, Alberta, Canada
5. Liang X, Zuo MJ, Patel T (2013) Evaluating the time-varying mesh stiffness of a planetary gear set using the potential energy method. Proc IMechE Part C: J Mech Eng Sci, Online available http://pic.sagepub.com/content/early/2013/04/24/0954406213486734.abstract
6. Yang D, Sun Z (1985) A rotary model for spur gear dynamics. J Mech Transm-T ASME 107:529–535
7. Parker R, Lin J (2004) Mesh phasing relationships in planetary and epicyclic gears. J Mech Des 126:365–370
8. Kahraman A (1994) Load sharing characteristics of planetary transmissions. Mech Mach Theory 29:1151–1165

Chapter 34
Signal Patterns of Piston Slap of a Four-Cylinder Diesel Engine

Tian Ran Lin, Andy C. C. Tan, Peter Crosby and Joseph Mathew

Abstract This paper presents an experimental study on the vibration signal patterns associated with a simulated piston slap test of a four-cylinder diesel engine. It is found that a simulated worn-off piston results in an increase in vibration root-mean-square (RMS) peak amplitudes associated with the major mechanical events of the corresponding cylinder (i.e., inlet and exhaust valve closing and combustion of Cylinder 1). This then led to an increase of overall vibration amplitude of the time-domain statistical features such as RMS, crest factor, skewness, and kurtosis in all loading conditions. The simulated worn-off piston not only increased the impact amplitude of piston slap during the engine combustion, but also produced a distinct impulse response during the air induction stroke of the cylinder attributing to an increase of lateral impact force as a result of piston reciprocating motion and the increased clearance between the worn-off piston and the cylinder. The unique signal patterns of piston slap disclosed in this paper can be utilized to assist in the development of condition monitoring tools for automated diagnosis of similar diesel engine faults in practical applications.

34.1 Introduction

As part of a broader project being undertaken by the CRC for Infrastructure and Engineering Asset Management, an online diesel engine monitoring system is being developed in collaboration with ASC Pty Ltd. The application of this system is initially focused on its application to a Collins class submarine test engine;

T. R. Lin (✉) · A. C. C. Tan · J. Mathew
Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: trlin@qut.edu.au

P. Crosby
ASC Pty Ltd, 694 Mersey Road North, Osborne, SA 5017, Australia

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_34,

however, such a system and its associated technologies have the potential for a much broader application. An experimental study on the vibration signal patterns associated with a simulated piston slap test of a four-cylinder diesel engine is one of several research elements undertaken by the CRC to support the development of the diagnostic and prognostic tools for this system.

Fuel efficiency, exhaust emission, and reliability are the main concerns in modern diesel engine operation. Unpredicted diesel engine faults, particularly combustion-related faults, not only produce unwanted radiated noise and reduce fuel efficiency and reliability, but can also increase exhaust emission of an engine which can pose an environment hazard to human health. The damage caused by particle emission of a diesel engine to human health can be even worse in enclosed spaces such as in an underground mine. It is thus essential to mitigate the impacts of unpredicted engine faults through early detection to ensure optimum engine performance and to lessen exhaust emission. To minimize or to prevent the occurrence of unpredicted engine faults, the health state of a diesel engine needs to be monitored continually so that a fault symptom and its root cause can be diagnosed and dealt with at the first instance before it deteriorates further and causes damage.

Piston slap is one of the most common combustion-related diesel engine faults, which is caused by excessive clearance between the piston and the cylinder wall or liner due to wear or inappropriate operation of an engine. Excessive piston slap can induce severe engine damage such as scuffing on the piston and cylinder wall/ liners. Vibration and noise radiation from diesel engines due to piston slap have been studied for several decades. Ungar and Ross [1] presented a theoretical model to estimate the vibration and noise power induced by piston slap of reciprocating machinery with a particular focus on diesel engines. Cuschieri and Richards [2] studied the noise radiated by combustion and piston slap of a diesel engine by considering single impacts. A method was also proposed in their study to reduce the noise due to piston slap. Nakashima et al. [3] presented a numerical simulation study to reduce piston slap noise of a diesel engine by optimizing the piston center of gravity and piston pin offset design. Piston slap noise was separated from the combustion noise of a four-stroke four-cylinder diesel engine by Badaoui et al. [4] by utilizing a cyclic Wiener filter. In their approach, the filtered pressure signal measured inside a cylinder of the engine was used to evaluate the combustion power component on accelerometers mounted on locations close to the cylinder. The result was used to estimate the contribution of combustion noise to the overall vibration signal, which was then subtracted from the vibration signal to obtain the specific vibration component due to piston slap. A blind source separation model was further developed in an accompanying work to separate the piston slap and combustion noise of the diesel engine by using the vibration signals detected by the accelerometers attached close to one of the cylinders [5]. Zheng et al. [6] investigated piston slap-induced ship hull vibration by using finite element and boundary element analysis. They found that the excitation due to piston slap can induce higher-level ship hull vibration and sound radiation than vibration induced by the vertical inertial force of the reciprocating masses of the engine.

Nevertheless, because excessive piston slap can be easily detected from the emitted audible noise, not much effort has been directed toward an automated detection of such faults in diesel engine monitoring. Furthermore, piston slap normally occurs only after a long life span of diesel engine operation since excessive wear of piston and cylinder wall/liner usually takes a long period to develop. Simulation of such faults in a controlled manner is thus vital to understand the signal patterns for automated engine monitoring and fault diagnosis. As part of a program to develop a comprehensive engine condition monitoring and fault diagnostic tool for online condition monitoring of the Hedemora diesel engine of ASC Pty Ltd [7], several common diesel engine faults such as injector fault, inlet and exhaust valve faults, piston blowby, and piston slap were simulated in a diesel engine test rig in the laboratory. Signal patterns associated with each fault were studied by using various condition monitoring techniques in the investigation.

The signal patterns from the baseline test of a diesel engine test rig at different loading conditions have been investigated and reported by Lin and Tan [8] using acoustic emission (AE) and vibration techniques. They showed that the major combustion-related events such as inlet valve opening and closing, combustion, exhaust valve opening, and closing produced specific AE root-mean-square (RMS) peaks. A near-period decaying piston rocking motion (essentially piston slaps) during the engine combustion process was also detected in the measured AE signals. It was found that the mechanical process and operation condition of the diesel engine were more stable under higher loads (e.g., full load) for a healthy engine. Lin et al. [9] further illustrated that the loading condition of a diesel engine can be estimated from the instantaneous angular speed (IAS) analysis. Moreover, it was found that the IAS technique can be employed to detect the leaking of an exhaust valve. Leakage of the exhaust valve will result in reduced combustion power output of a cylinder and thus lead to the reduced amplitude of the order component of the IAS waveform corresponding to the engine firing frequency. The amplitude reduction of the order component of the IAS waveform increased for higher engine loads attributing to the increased combustion power loss by the leaking exhaust valve. The result and finding for this part of work are summarized in a separate paper to be submitted for consideration of a journal publication. Signal patterns (the time- and frequency-domain features) of a simulated injector fault of the diesel engine were reported in another study [10] in which both AE and in-cylinder pressure techniques were employed.

Since piston slap can be detected in the audible frequency range, vibration data will be used in the analysis presented in this study. Typical signal patterns and features associated with piston slap of a four-cylinder diesel engine are studied. A description of the diesel engine test rig and the experimental setup is given in Sect. 34.2. Section 34.3 presents the analysis of vibration signals to extract the signal patterns and features induced by piston slap. A general discussion of the result is also included in the section. Section 34.4 provides a summary of the main findings from this study and is followed by a general conclusion.

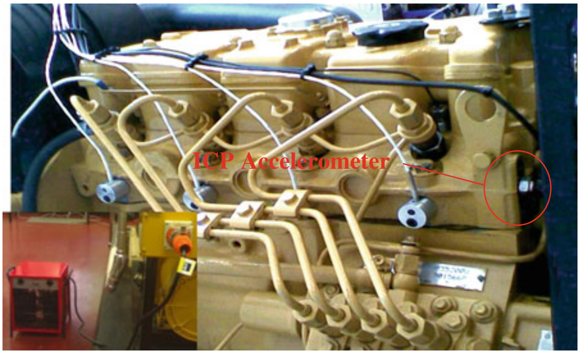


Fig. 34.1 A graphical illustration of the four-stroke, four-cylinder diesel engine test rig and the attached sensors

34.2 A Description of the Engine Test Rig and Experimental Setup

The diesel engine test rig as shown in Fig. 34.1 was employed in a range of engine fault simulations in the experimental study. The engine is an in-line four-stroke, four-cylinder diesel engine, which is commonly used for power generation in remote construction sites or an emergency power generator in hospitals. The engine generates about 15 kW of nominal power output at full-load condition. During the experiment, the power outlet of the diesel engine generator set was connected to a three-phase 15-kW industrial fan heater. The fan heater has three heat settings at 5, 10, and 15 kW, which can be adjusted during the experiment for different engine loadings. The engine specification is shown in Table 34.1.

A PCB ICP industrial-type piezoelectric accelerometer was stud-mounted on the non-drive end of the engine block close to Cylinder 1 as shown in Fig. 34.1 for vibration detection of piston slap and other engine operation noise. The accelerometer mounting location was chosen to minimize the effect of high-amplitude low-frequency engine rocking motion (i.e., swinging of the engine block on the flexible engine mounts) on the measured acceleration signal during engine operation. Other sensors installed on the test rig include four resonance-type AE sensors, one high-temperature pressure sensor, and a combined encoder and top dead centre (TDC) recorder unit. The signals acquired from these sensors are not used in this paper but used in the analysis presented in Refs. [7–10].

The setup of vibration measurement is schematically illustrated in Fig. 34.2. A National Instrument data acquisition card (NI DAQ 6062E), which has a maximum sampling frequency of 500 kHz and can acquire data synchronously for up to 16 channels, was used for the data acquisition. A sampling frequency of 30 kHz was used in the experiment with a low-pass filter set at 10 kHz for vibration measurement.

After completing the baseline measurement of the diesel engine at various loading conditions, a simulated faulty piston (see Fig. 34.3b) was used to replace the healthy piston (see Fig. 34.3a) in Cylinder 1 to simulate piston slap during the

Table 34.1 Specification of the diesel engine

Perkins 404C-22 engine data	
Number of cylinders	4
Arrangement	In-line
Running speed	1,500 rpm
Bore	84.0 mm
Stroke	100.0 mm
Displacement	2.216 l
Compression ratio	23.3:1
Firing order	1–3–4–2
Injection timing (estimated)	15° (±1°) before TDC
Exhaust valve open (measured)	143° after TDC
Exhaust valve close (measured)	370° after TDC
Inlet valve open (measured)	354° after TDC
Inlet valve close (measured)	584° after TDC

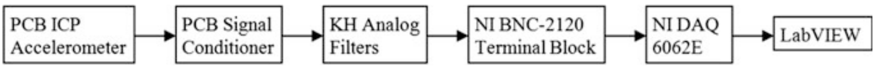


Fig. 34.2 A schematic illustration of the data acquisition setup in the experiment

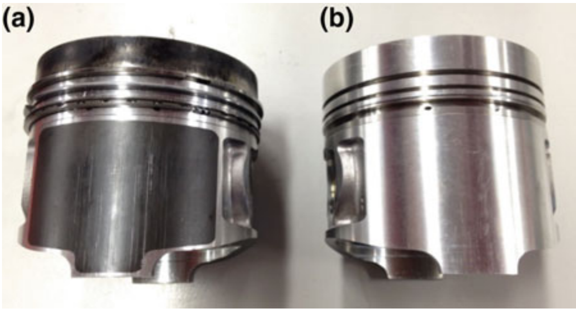


Fig. 34.3 A graphical illustration of the piston; **a** healthy piston and **b** faulty piston (the piston skirt was milled off by 0.1 mm)

engine operation. In this simulation, the skirt of the piston was milled off by 0.1 mm around the surface (i.e., the upper service limit according to the engine service manual) as shown in Fig. 34.3b. Signal characteristics of the simulated piston slap are analyzed in the next section to extract the useful features associated with piston slap to assist the development of automated condition monitoring applications for diesel engines.

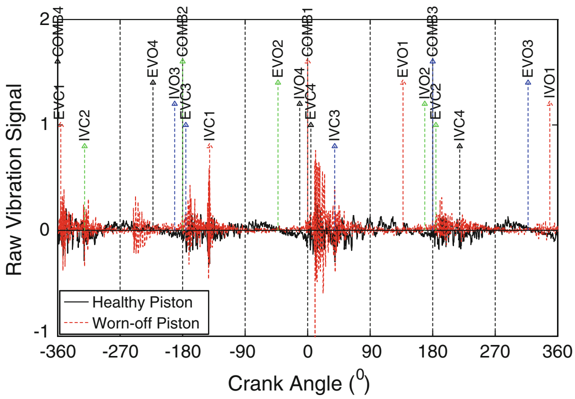


Fig. 34.4 Comparison of the measured raw vibration signals for the healthy and worn-off piston cases

34.3 Analysis of Vibration Signals from the Piston Slap Test

34.3.1 Vibration Signal Patterns

Figure 34.4 compares the measured raw vibration signals for the healthy and the worn-off piston cases of the engine at the unload condition. It is observed that there is a noticeable increase in vibration amplitudes during the combustion of Cylinder 1 when the simulated worn-off piston was used. In Fig. 34.4, COMB denotes the combustion, EVC stands for exhaust valve closing, EVO represents exhaust valve opening, IVC denotes inlet valve closing, and IVO indicates inlet valve opening. The number following these acronyms represents the corresponding cylinder number.

The averaged vibration RMS data (averaged over 192 engine cycles) calculated from the measured vibration signals for both health and faulty piston cases at the unload condition are shown in Fig. 34.5. It is shown that every valve closing event produces a corresponding vibration RMS peak. The simulated worn-off piston has led to a substantial amplitude increase for vibration RMS peaks associated with the major mechanical events of Cylinder 1 (i.e., COMB1, EVC1, IVC1) attributing to the increasing piston slap. The slight mismatch between some mechanical events and the corresponding vibration RMS peaks in the figure is caused by the engine speed variation and the error between the static measured valve timing and the actual valve timing, while the engine is in operation. A vibration RMS peak due to piston slap can be clearly identified during the air induction stroke of Cylinder 1 as shown in Fig. 34.5. The peak is present in the vibration signals for all loading conditions as shown in Fig. 34.6 but was absent in the vibration signals when the healthy piston was used in the experiment. The vibration signal produced by piston slap during engine combustion process overlapped with the engine combustion noise [4], which is difficult to differentiate from the low-frequency vibration

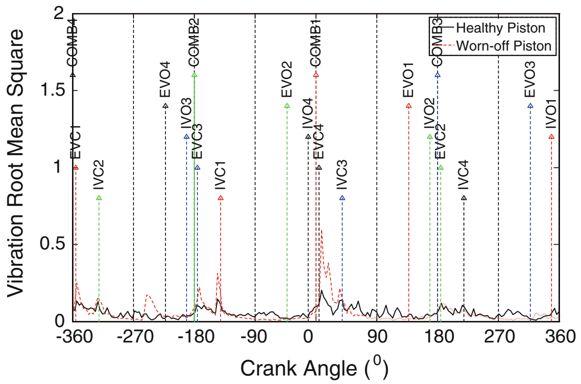


Fig. 34.5 Comparison of vibration RMS amplitude for the healthy and worn-off piston cases

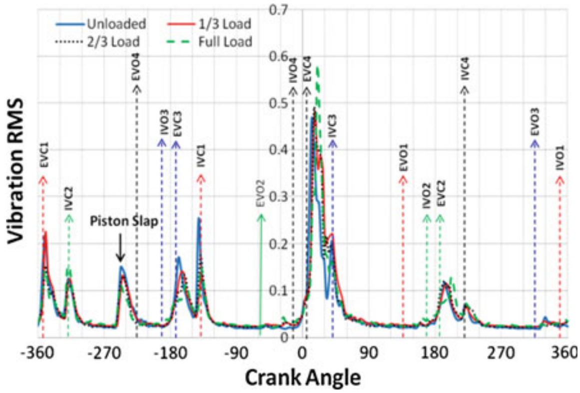


Fig. 34.6 Comparison of vibration RMS amplitude for the worn-off piston case at different loading conditions

signal. However, impacts produced by piston slap during engine combustion were clearly detected by using high-frequency AE technique in the experiment [8] where piston slap was termed as piston rocking motion. Piston slap induced by the direction change of the lateral force of the crank-rod-piston system for other piston strokes as disclosed by Geng and Chen [11] was not detected by the vibration sensor. This could be either due to the overlapping of piston slap-induced vibration and those generated by other mechanical events of the engine or due to the energy attenuation of the impact signal when it is propagated from the cylinder wall to the engine block.

In order to extract the useful frequency features of the signal, power spectra of the raw vibration data for the healthy and the worn-off piston cases at the unload condition (see Fig. 34.4) are presented in Figs. 34.7 and 34.8, respectively. It is shown that the spectra of the vibration signal in both cases are largely contaminated by the electric frequency (50 Hz) and its higher harmonics of the output power generated by the diesel engine operation. These electric frequency components can be clearly identified and isolated from the spectra using the order

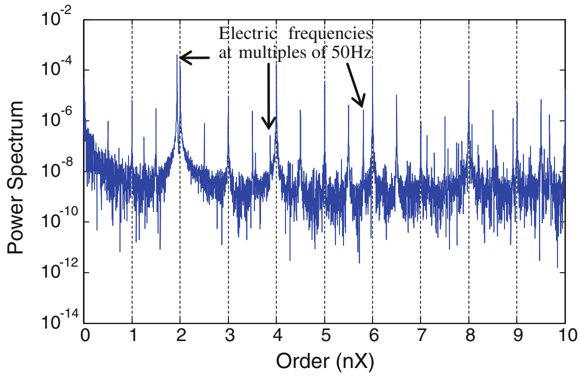


Fig. 34.7 Power spectrum of the vibration signal at the unloaded condition for the healthy piston case

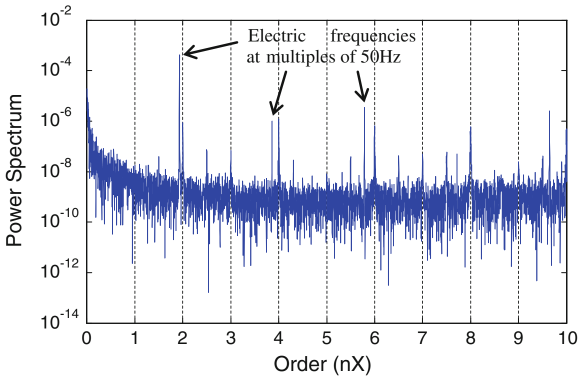


Fig. 34.8 Power spectrum of the vibration signal at the unloaded condition for the worn-off piston case

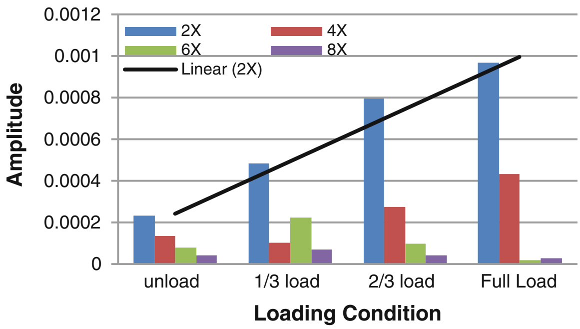


Fig. 34.9 The averaged amplitude of the four major order components at different engine loading conditions in the healthy piston case

analysis as shown in Figs. 34.7 and 34.8. However, care has to be taken when the frequency of the major order components is close to the electric frequency and its higher harmonics such as that in the full-load case, which is to be discussed in the subsequent analysis.

The major order components of the power spectra in both cases contain the orders of 2, 4, 6, and 8 times of the crankshaft rotating frequency. This is typical for a four-stroke, four-cylinder diesel engine. Other order components (the less dominant) such as 0.5, 1, 1.5, 2.5, and 3 can also be identified in the spectra although these order components have smaller amplitudes. For a better fault characterization, the major order components at different engine loading conditions were extracted from the spectra. The extracted order components at each loading condition were averaged across 10 data files to minimize the effect of amplitude variation of the diesel engine on the interpretation of the result.

34.3.2 Frequency-Domain Feature Extraction

Figure 34.9 presents the (averaged) major order components of the vibration power spectrum at different engine loading conditions for the healthy piston case. It is shown that the spectrum is dominated by the order component associated with the engine firing frequency (i.e., the order of 2). The amplitude of this order component increases proportionally with the increase in engine load, which indicates an increase in engine power output. Also presented in the figure is the linear trend line calculated based on the second-order component as a function of engine loading conditions. Figure 34.10 shows the amplitude of the four major order components (i.e., the orders of 2, 4, 6, and 8) at different engine loading conditions for the simulated worn-off piston case. Unlike that of the healthy piston case, the second-order component does not have a dominant amplitude when compared with the other order components in all loading cases except for the full-load condition (which is not shown in the figure due to the reason to be explained in the following paragraph). Since the impact time by piston slap is very short, the vibration response induced by piston slap is characterized by high-frequency contents. This helps to explain why the amplitude of the low-frequency-order components in the simulated piston slap case (as shown in Fig. 34.8) does not increase from that of the healthy piston case (as shown in Fig. 34.7) for all engine loading conditions. Instead, the amplitude of these order components decreases from that of the healthy piston case.

In the full-load condition (i.e., the mean engine speed of approximately 1,500 rpm), because the frequency of the second-order component overlaps with the electric power frequency, and the amplitude of this electric frequency component is much higher than that of the second-order component such as that shown in Fig. 34.8, the true amplitude of this order component is difficult to determine in this loading case. Although not uniformly, the sixth-order component appears to have a higher response amplitude than the other order components in this case. The amplitude of this order component has a general increasing trend as the engine loading increased as shown in Fig. 34.10. Further investigation is required to improve the understanding of the frequency-domain feature associated with the simulated piston slap.

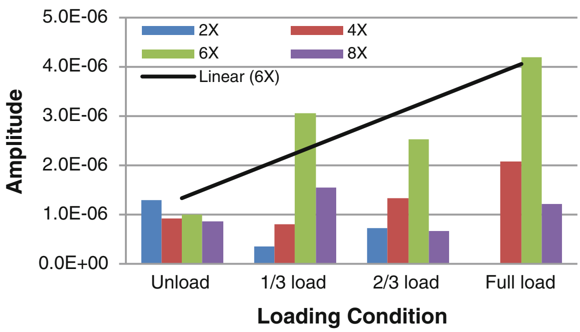


Fig. 34.10 The averaged amplitude of the four major order components at different engine loading conditions in the simulated worn-off piston case

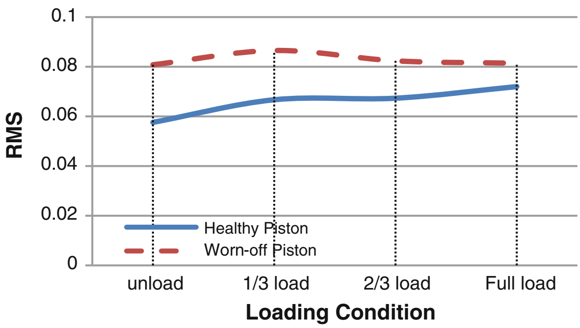


Fig. 34.11 The RMS value of vibration signals at different loading conditions

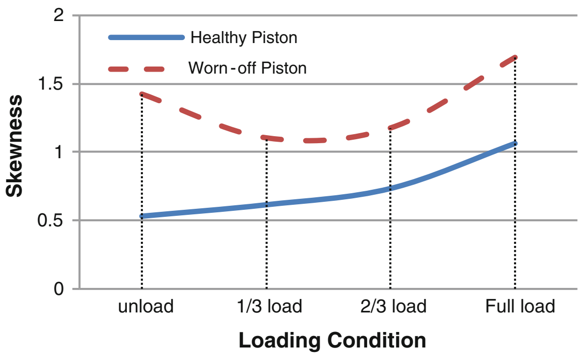


Fig. 34.12 The crest factor of vibration signals at different loading conditions

34.3.3 Time-Domain Feature Extraction

Four typical time-domain features, RMS, crest factor, skewness, and kurtosis were also calculated and are presented in Figs. 34.11, 34.12, 34.13 and 34.14, respectively. Each feature was calculated and averaged over ten data files. It is shown that all the time-domain features chosen in this study provide a better indication than the order domain components for the detection of the simulated worn-off

Fig. 34.13 The skewness feature of vibration signals at different loading conditions

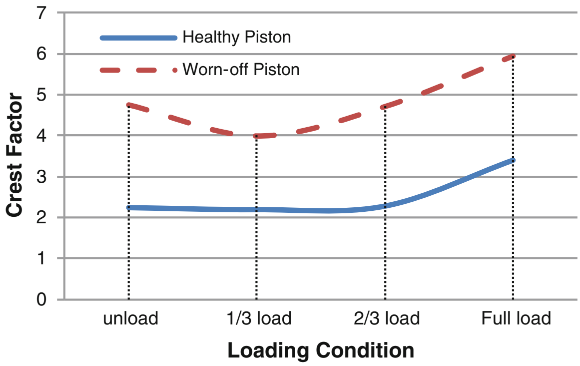
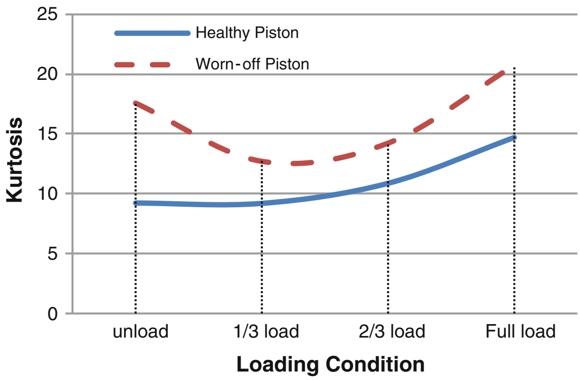


Fig. 34.14 The kurtosis value of vibration signals at different loading conditions



piston at all loading conditions. These time-domain features can be utilized for condition monitoring and automated fault diagnosis of diesel engines in practical applications.

34.4 Summary and Conclusion

In this paper, time- and frequency-domain features of vibration signals from a four-stroke, four-cylinder diesel engine due to a simulated piston fault were analyzed. Unique features associated with the simulated piston fault were extracted from the vibration signals for condition monitoring applications. It was found that the simulated worn-off piston resulted in increased vibration RMS peak amplitudes for major mechanical events of the corresponding cylinder (i.e., inlet and exhaust valve closing and combustion of Cylinder 1). This then led to the increase in overall amplitude of the time-domain statistical features, i.e., RMS, crest factor, skewness, and kurtosis in all loading conditions. The simulated worn-off piston not only increased the impact amplitude of piston slap during the combustion, but also produced a distinct impulse response soon after the compression stroke of the cylinder attributing to the increase in lateral impact force as a result of piston reciprocating motion and the increased clearance between the worn-off piston and the cylinder. These unique signal patterns of piston slap as disclosed in this paper can assist in the development of condition monitoring tools for automated diagnosis of similar diesel engine faults in practice.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC. The authors also acknowledge the financial support from ASC Pty Ltd, the CRC's industry sponsor for the purchase of the diesel engine test rig used in this research.

References

1. Ungar EE, Ross D (1965) Vibrations and noise due to piston-slap in reciprocating machinery. J Sound Vib 2:132–146
2. Cuschieri JM, Richards EJ (1985) On the prediction of impact noise, Part VIII: diesel engine noise. J Sound Vib 102:21–56
3. Nakashima K, Yajima Y, Suzuki K (1999) Approach to minimization of piston slap force for noise reduction—investigation of piston slap force by numerical simulation. JSAE Rev 20:211–216
4. El Badaoui M, Daniere J, Guillet F, Serviere C (2005) Separation of combustion noise and piston-slap in diesel engine—Part I: separation of combustion noise and piston-slap in diesel engine by cyclic Wiener filtering. Mech Sys Sign Proc 19:1209–1217
5. Serviere C, Lacoume JL, El Badaoui M (2005) Separation of combustion noise and piston-slap in diesel engine—Part II: Separation of combustion noise and piston-slap using blind source separation methods. Mech Sys Sig Proc 19:1218–1229
6. Zheng H, Liu GR, Tao JS, Lam KY (2001) FEM/BEM analysis of diesel piston-slap induced ship hull vibration and underwater noise. Appl Acoust 62:341–358
7. Lin TR et al (2011) Development of a diagnosis tool for condition monitoring of rotating machinery. In: Proceedings of ICOMS asset management conference, Gold Coast, Queensland Australia, 16–20 May 2011
8. Lin TR, Tan ACC (2011) Characterizing the signal pattern of a four-cylinder diesel engine using acoustic emission and vibration analysis. In: Shen G (ed) Proceedings of the world conference on acoustic emission, Beijing, P. R. China, 24–26 Aug 2011
9. Lin TR et al (2011) Estimating the loading condition of a diesel engine using instantaneous angular speed analysis. In: Lee J (ed) Proceedings of the 6th annual world congress on engineering asset management, Cincinnati, Ohio, USA, 2–4 Oct 2011
10. Lin TR et al (2011) Condition monitoring and diagnosis of injector faults in a diesel engine using in-cylinder pressure and acoustic emission technique. In: Law S (ed) Proceedings of the 14th Asia pacific vibration conference, Hong Kong, P. R. China, 5–8 Dec 2011
11. Geng Z, Chen J (2005) Investigation into piston-slap-induced vibration for engine condition simulation and monitoring. J Sound Vib 282:735–751

Chapter 35
Human—An Asset or a Liability: The Real Deal with Modern Humans in Intelligent Systems and Complex Operations

J. P. Liyanage

Abstract Popular understanding today characterizes human as an asset of an organization. On the contrary, various industrial cases often conclude human error as a major cause for unwanted events and incidents. Subsequently, much focus over the years has been paid on the human reliability, human error etc. Under this conventional focus lies the hidden presumption that performances and behaviors can be controlled or regulated by various technical, operational, and managerial means. These imply that humans take up a dual role in the use and operation of industrial systems; an asset on certain occasions and/or a liability on the others. This paper takes a critical look at this dual role with respect to the developments that has taken place over thousands of years and the nature of industrial systems that have been developed as well as those that are emerging. It strongly argues that the very nature of the humans are not understood or incorporated in the design, deployment, and operationalization of industrial systems. With respect to the complexity and dynamics of emerging systems, this has very sensitive implications in terms of safety and security.

35.1 Introduction

Over hundreds and thousands of years, humans have been living in various ages of *transforming systems*. It implies that the industrial systems have been changing their principal characteristics following the transitions and developments of the communities, technologies, political ideologies, civil systems, and so on. This process has resulted in a certain *culture of complexity* where the critical components of industrial systems on the one hand go through a physical miniaturization

J. P. Liyanage (✉)
Centre for Industrial Asset Management (CIAM), Faculty of Science and Technology, University of Stavanger, 4036 Stavanger, Norway
e-mail: j.p.liyanage@uis.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_35,

and on the other embrace advanced functions. This at the same time extends the boundaries of influence on humans with gradually increasing complexities. While the technical complexity of certain industrial systems is fairly explicit, some other types of *absorbed/embedded complexities* tend to remain implicit due to the very nature that the real impact of transforming systems has been difficult to interpret due to lack of knowledge or understanding of the hidden interfaces that those systems inherit/lead to, as the operational boundaries are continuously altered/extended.

Subsequently, humans have been exposed to a rapidly advancing technological environment, and the underlying assumption in this environment has been that the humans are able to cope with growing complexities. What was observed over the years at the same time was that such terms as human error and human reliability came to the forefront of the discussions, debates, and investigations as the very central terms to explain why certain incidents and events occur leading to serious losses. It implies that the human role in continuously advancing technological environments has embraced a notion of some form of liability. Interestingly, at a certain stage of the development process, the intellectual capabilities of the human received increasing attention. This perception is attributable to the growth of active human engagement in rapid technological innovation for commercial advantage and the strong focus on the abilities (knowledge, understanding, etc.), and the other values that can be transformed into some form of beneficial performance in a given context. In organizational management terms, thus, human role embraced the notion of an asset underlying that humans with certain attributes are critical to the value creating processes in industrial systems.

Thus in accordance with the above perceptions, humans have a dual role in use and operation of complex systems, either as an *asset* or as a *liability*. The underlying implicit assumptions here is that certain attributes can be used to classify humans into the two categories. In this regime, not much attention has been paid to the underlying growth of complexities and parallel development processes of different elements that shape human performance on a continuous basis under different circumstances and the dynamic/stochastic changes in human characteristics. This implies that the discussion of the human role as an asset or a liability embraces some form of a static character and apparently assumes an '*ability to cope*' with the nature in rapidly transforming environments. In contrast, what is clearly arguable is that the human role is inconsistent with nature in the engagement with various systems from time to time and human performances are very much situation and time dependent. Moreover, the added complexity due to rapidly transforming systems makes the human a more dynamic element of advanced systems, demanding frequent '*shifts*' of our engagement patterns and actions. Interestingly at the same time, the nature and scale of interaction of humans with the transforming systems have begun to contribute much to the change of fundamental characteristics of the human himself/herself in very many ways, indicating a mutual effect with the system development and usage/operational processes.

This paper, thus, takes a critical look at the dual role of human with respect to the developments that has taken place over thousands of years and the nature of industrial systems that have been developed, as well as those that are emerging. It strongly argues that the *true* nature of the humans is not effectively reflected and the dynamic complexities are not understood and incorporated in the design, deployment, and operationalization of industrial systems. It is more so in the assessment of usage/operational characteristics. This has a great potential toward performance shortfalls of humans in dealing with advanced systems and complex operations. With respect to the growth of complexity of emerging systems and the scale of influence on other systems, this will have very sensitive implications particularly in terms of safety and security.

35.2 The Technological Growth Process

The technological growth process has been quite continuous with the gradual development of human intellectual abilities from the very early ages. A closer look at the very beginning of this process reveals how the very primitive technologies were taken into a dedicated use for real needs, whereas in much later stages the motivation for designing and deploying technical systems have taken a dramatic turn. Since late, the nature and scale of this growth process has been relatively explosive and has largely been accelerated by the development both in industrial and in information technology sectors (Fig. 35.1).

The continuous growth process has come to a status today of *machine intelligence and virtual community*, where the structural configuration and dynamics of interaction have begun to show significantly different attributes and patterns in comparison to the previous stages. Subsequently, the process has added new dimensions (rather complexities) to the human expectations, values, characteristics, roles, functions, etc., providing new perspectives and adding new capabilities to the human growth process. The technological growth seems to continue toward an '*age of cyberspace*' (particularly due to the defining impact in information and communication technology) where significant changes can be anticipated in social, commercial, technical, and political terms.

The observation as of today is that as much as ongoing processes have contributed to narrow the *human-technology divide* in many different fronts, it has also induced other elements that have notable effects on '*who we are today.*' The sheer comparison indicates much not only about the diminishing divide but also about the characteristics of modern digital living that, for instance, tend to challenge the traditional social hierarchies and political authorities. Some of the landmarks of modern human living that, for instance, vary from by-pass surgeries, cloning, DNA research to ID theft and cyber terrorism, not only make a statement of intellectual capabilities of modern human civilizations but also make an expression of growth of complexity of humans and the world around.

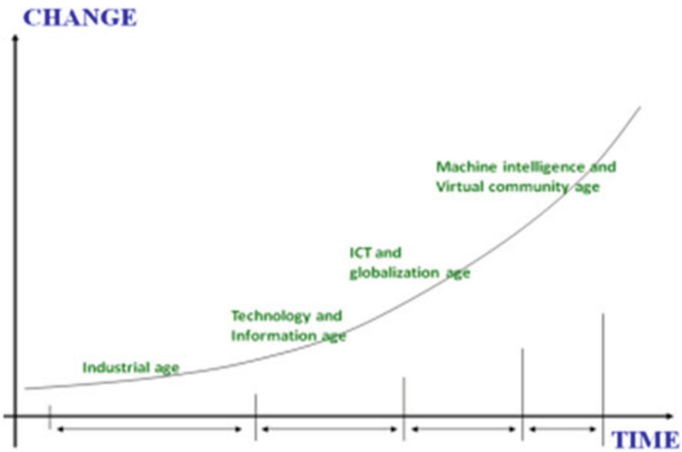


Fig. 35.1 Technological growth in a relative time domain

In terms of the development and deployment of industrial systems, many processes are underway to operationalize complex systems for human comfort, safety, and enhancing quality of life. The situation has matured to an extent that even the concepts related to *ultra* complex industrial systems are ready to be launched and the necessary know-how and the technologies are already at disposal. The next generation solution is in fact a matter of time and the market readiness to absorb the change. Many sectors today appear to go through this transformation process with systems solutions based on 24/7 *virtual community platforms* with varying degree of complexity that, for instance, include the following:

- Smart homes that obey our remote commands and execute functions to help support day-to-day household living.
- Smart cities with intelligent transportation networks where moving from A to B is optimized based on built systems intelligence that has capabilities to manage dynamic conditions
- Health monitoring and medical care delivery processes based on home-installed gadgets and devices connected to large-scale health networks
- Application services (based on near-field or far-field communication) that allows automated information exchange between digitized service offers and mobile devices carried in pockets.

Obviously, the development of such advanced systems is quite a demanding activity requiring a blend of innovation and creativity. Given the sensitivity of technological innovations in commercial terms in modern business context, it has begun to take the humans through a change process where the systems developers have been able to capitalize on *imaginations* to retain the competitivity in the product–service markets. It implies that in social and commercial scales there is a

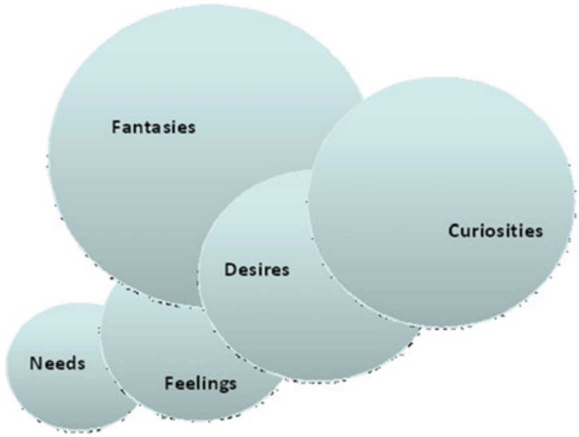


Fig. 35.2 Drivers of technological growth function with imagineering role

blend of *created expectations* (*and fantasizes*) to drive the complex system development process, where *imagineering* plays a very critical role, as opposed to simplistic needs (Fig. 35.2).

In this continuous process, humans apparently are exposed to dual forces; *push,* on the one hand, where the increasing needs for a safe and comfort life demand advanced products and services, and *pull* where product and service innovation has served the humans beyond their realistic expectations. It has introduced different impact levels more in sociological, behavioral, and psychological scales, and relatively less in physiological scale.

35.3 Core Elements of Complex Industrial Systems and the Impact Factor

No matter how technically advanced or functionally complex a system can be, every industrial system is designed on the basis of three principal elements, namely *operating concept/scenario, technological platform*, and *human workforce/user groups* (Fig. 35.3). The operating concept/scenario emphasizes on the structure, principles, and the functional configurations that are required to deploy the systems to ensure reliable and safe operations. The innovative technology on the other hand provides the necessary technical platform to keep the operational structure and the functions intact with principal role of ensuring the technical integrity of systems and its components. The workforce/use groups often engage with the systems based on the premises or the conditions set by the technical platform and the operating scenario. The interfaces between these core elements play meaningful roles in keeping the systems fully functional and intact with respect to the owners'/operators' expectations and acceptance criteria. The interface between the operating concept and the technical platform is primarily important for development and deployment

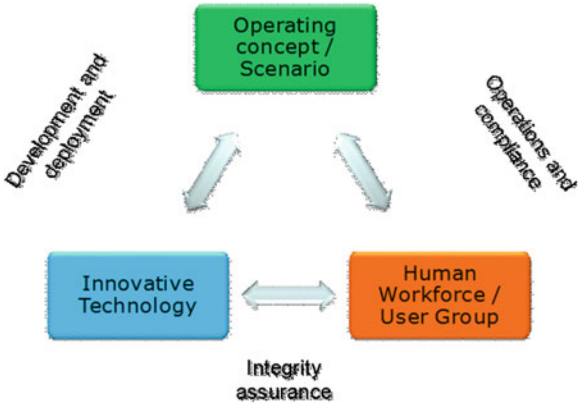


Fig. 35.3 Core elements of complex systems

processes that ensure appropriate initial conditions for work/use. The human interface with the other two elements has two unique roles; technical decision/tasks associated with assurance of integrity with the technical platform on the one hand and operational decisions/tasks associated with operational compliance on the other. In systems context, it is practically difficult to disintegrate the former from the later, since the work performance patterns on technical items are regulated by the operational attributes/demands. However, those interfaces have distinctive character with respect to the nature and content.

To a larger extent, the system development and deployment process so far have paid a relatively higher degree of attention on the operating concept and the technological platforms. The underlying presumption is that the workforce/user groups would be able to cope with the change processes, more or less, and that certain unavoidable gaps can be dealt with by means of pro-deployment measures that can vary from training modules, operational instructions, to job aids. However, in human domain the adaptive capability is not just limited to familiarization with new operational content and the technological components. It in fact brings three very dynamic issues into a complex perspective; *changing organizations, work forms*, and *human evolution*.

From a sheer *technology-induced organizational transition* perspective, the shift is apparently toward a virtual setting, where some of the key attributes include as follows:

- Overcome geographical barriers
- Transform into virtual organizations
- Capitalize on distributed teams based on real-time technologies
- Pursue B2B collaboration and network-based learning
- Relies heavily on active information platforms and distributed knowledge systems
- New governance processes and mechanisms to balance control versus trust
- Expand connectivity from point-to-point to many-to-many

This has unavoidable impacts on the work practices that seem to be heading toward more of a *hybrid work culture*. The emerging trend is toward a work setting where complexities are visible in many aspects due to the very nature of work configuration that gradually:

- Becomes more cognitively demanding
- Builds interactive information and knowledge systems
- Becomes greatly influenced by social intelligence and social competencies
- Promotes joints decision making and remote work support environment
- Becomes virtual and subsequently heavily multi-agent dependent
- Promotes leaning more toward 'tele-medicine'-type solutions

It is quite obvious that the innovative capabilities built into the modern complex systems have driving roles in shaping up both the organizational and work landscape. An important question in this fast growing *techno-organizational* environment is that *where are the humans in this change equation and what are the notable attributes that define human role in modern complex systems*.

35.4 Attributes and Role of Human in Emerging Systems

The history of humans runs back to approximately 5–7 millions of years. The migration theories, traces of settlements, and early civilizations show the signs of greatest qualities and intellectual capacities of the human ancestors. The gradual technological advancement that goes hand-in-hand with human development began with primitive technologies, for instance, in transportation, navigation, irrigation, and so on, that had a tight interface with the growth of civilizations. *Who the humans were yesterday, who they are today*, and signs of *who they will be in the future*, illustrate much about the evolving trends of the fast diminishing *digital divide* as well as the characteristics of the modern *digital living* that, for instance, tend to challenge the traditional social hierarchies and the political authorities. The modern landmarks of the human in the digital world, which, for instance, marked by advanced medical applications, communication systems, social dynamics, work configurations, etc., not only make a statement of the attributes of modern civilizations but also make an expression of growth of complexity of *who the humans actually have become* and *who they in fact are going to be*.

In this context, the impact of continuous human evolution on the dynamics and the characteristics of transforming systems can perhaps be considered as the most unknown variable of the global change processes in social, technical, political, and even commercial scale. The natural human evolution process has been influenced or rather accelerated to a great degree by the evolving and transforming technologies and societies having a complex impact on the human attributes. Perhaps the most critical factor is the *generation shift*, of which the real impact on the complex systems is largely uncertain, leaving two possible hypotheses; either toward a *seamless interface* or toward a *crash course*.

Table 35.1 The generation shift

Generation type	Key attributes
X	Born to post world war II with a different identity, relatively different sociopolitical expectations and engagement patterns, apparently reactive to events and life situations, open conflicts due to value shifts with pre-war generation
Y	The millennial generation, increased familiarity with digital technologies, neoliberal approach to political and economical exposures, notable conflict of interests with Y generation, relatively high work benefit driven at the cost of mere participation, tend to be mobile with a higher degree of changes in work styles, open socialization patterns
Z	Internet generation or the digital natives, born in the high-tech world, highly connected through digital technological platforms, extremely open socialization patterns and dynamic engagement, higher dependence on technology, branded as instantaneous and impatient

The *X–Y–Z shift* indicates itself of the patterns of human evolution as a continuous process. What is observed today is that it has taken a different turn due to various influence factors induced by the modern high-tech societies and advanced systems. As notable, cognitive capabilities are at the front end of the development process, implying that physiological changes in human are relatively negligible in comparison to *cognitive, behavioral*, and *sociological* transformations. This has already begun to give signs of significant implications on the values, expectations, attitudes, and other attributes of engagement of the new workforces with the current industrial and work systems (Table 35.1).

With respect to those ongoing change processes of the highly dynamic environment, subsequent changes in various other dimensions, and the observable human adaptive mechanisms, the key attributes of the future workforce can be explained in certain *transitional* terms. This, for instance, includes various transitions that can be highlighted from a multiple perspective, i.e.;

- From '*socio-technical*' systems to '*techno-social*' systems, where
  - the transfer is evident from the use of technology to support social functions, to re-configure the social fabric to cope with the technology
- From '*knowledge-based*' to '*hybrid intelligent*' performance, where
  - the transfer is evident from hands on experience and in-depth know-how to technology-dependent high-performance cultures (i.e., hybrid living)
- From 'risk *aversive*' to '*risk seeking*' orientation, where
  - the transfer is evident from risk as a negative attribute that should be avoided at all possible times, to risk as the means to advance oneself fast and aggressive to the next level.

- From 'smart *brains*' to '*cognitive coding*,' where
  - the transfer is evident from comprehension and ability to reproduce the knowledge, to the very ability to break the known codes through advanced cognitive functions/capabilities
- From '*patience*' to '*agile response*' behaviors, where
  - the transfer is evident from 'take your time' and 'one at a time' mind-set, to the pursuit of speed and multitasking in our daily interactions

These observations imply that as a new generation enters the workplace there is a gradual trend for major transformations in psychological, sociological, behavioral, and attitude terms. This to a large degree is further influenced by the digital technologies and the subsequent considerable changes in the lifestyle. It implies that as the new complex systems being developed, for instance, as observable in traffic, oil and gas, defense, health care, aviation, etc., the level of complexity associated with real usage also increases due to multiple factors both from *push* and *pull* ends. This requires a cautious and perhaps more novel approaches to the design, deployment, and operation of modern systems that can avoid all possible malfunction and failure scenarios emanating from various human–system interfaces.

In conventional terms, engineering systems are known to adapt principles and policies where a considerable portion of financial commitments are met by the time of deployment and commissioning process, leaving just a minor margin for sustenance. But in reality, the actual economical figures often give a reverse trend where the sustenance cost in fact bear a major portion of economical burden in comparison to initial design and deployment costs. A good portion of increased costs at the operational phase of engineering systems are attributable to priory uncalculated costs inclusive of liabilities, rectifying/adaptive solutions, systems modifications, etc.

A classical bottleneck in design and usage of complex industrial systems involves the very inability to integrate organizational and human issues/dimensions to create effective and safe interfaces from the very early stages as a means to ensure reliable as well as cost-effective performance. To a greater extent, this seems to lead to events, incidents, and situations where *human error* is often cited as the root cause. Perhaps the most important question in modern complex engineering systems context, with due consideration on both characteristics of human evolution and the growth of complexity levels of modern interconnected systems, is as follows:

> Is it the sheer human error that should be avoided at all times during design-to-operation phases, or should the modern systems be designed to defend against erroneous inputs taking into consideration the very nature that humans are irrational and unreliable: i.e. human characteristics are time and situation dependent, and that human attributes are dynamic?

The importance of designing and operating systems with due regard on '*true*' human characteristics lies on the fact that the design-to-operation phases take into account the very characteristics of the '*natural forces*' that can regulate/shape guide the systems' performance. There are a number of measures that can be taken to ensure the required *defensive* character of complex systems, which, for instance, include the following:

- Dynamic learning platforms
- Design of advanced barriers, both technical and operational
- Engineering, organizational, and behavioral resilience
- High-security defensive tools (such as for large-scale IT systems)
- Etc.

Perhaps the best option above all is to rethink the very basic engineering and technological processes themselves to identify various measures necessary to reduce the ever growing complexity of modern engineering/industrial systems so that the future systems will become much simpler with respect to a specific and targeted need. The very adaptive practice of adding additional technical features and functions to promote technological innovation to create and capitalize on the markets should be highly debated since the direct consequence of such a practice, as observable, is to contribute to the growth of complexity (and eventual *chaos*) leading the modern systems, societies, and humans to vulnerabilities. This will have a defining impact on the human role as an *asset* or a *liability*.

35.5 Conclusion

Modern practices associated with design and operations of complex systems contribute to characterize humans as assets or liabilities. The difference, in principle, places much weight on the human reliability, human error as well as other intellectual attributes of the organizational workforce. This paper paid attention to the growth of complexity of industrial/engineering systems as well as the parallel process of human evolution. While the former tend to follow a practice based on innovation (principally comparable to fantasies and imagination), the latter is influenced by the natural development forces as well as forces emerging from the rapidly growing digital technologies and associated capabilities. Thus, the paper sheds lights on the question that is it not more logical that the true human nature (i.e., irrational and unreliable) is seriously considered and subsequently modern systems are designed, deployed, and operated in a defensive manner, rather than counting more on human error and human reliability aspects. It also argues that perhaps the time is mature for an effective debate to rethink the engineering and technological processes that is expected to lead the way toward some form of a technological leadership, yet which in turn tend to promote the complexity. This has a greater potential to expose modern societies, systems, and humans to increased vulnerabilities.

Chapter 36
The Use of Expert Systems in Offshore Oil and Gas Assets: A Status Review with Respect to Emerging Needs for Innovative Solutions

Nii Nortey B. C. Lokko, Jawad Raza and Jayantha P. Liyanage

Abstract The ultimate goal of collaborative operating environments was value creation. All over the world, expert systems (ES) are being employed by various industrial sectors to foster this value creation process. Subsequently, under the umbrella of integrated operations (IO), this paper examines the current role and use of expert systems for value creation in the Norwegian offshore oil and gas (O&G) industry within the field of asset maintenance. Through comprehensive literature reviews, vendor surveys, and interactions with industry experts, the paper concludes that the Norwegian O&G industry closely mimics the global O&G industry in its adoption and use of expert systems technology. This study also reveals an apparent lack of widespread adoption of ES within maintenance is a contributory factor to the proliferation of preventive maintenance strategies on the Norwegian continental shelf.

36.1 Introduction and Background

Since the last quarter of 2004, the buzz word/phrase within the Norwegian continental shelf (NCS) has been 'integrated operations (IO).' This current dispensation, according to 'information managers (IM)' and O&M supervisors/engineers,

N. N. B. C. Lokko (✉) · J. P. Liyanage
Centre for Industrial Asset Management (CIAM), University of Stavanger,
4036 Stavanger, Norway
e-mail: nii.nortey.lokko@uis.no

J. P. Liyanage
e-mail: j.p.liyanage@uis.no

J. Raza
Operations and Engineering Services, Apply Sørco AS, Moseidsletta 122, Forus,
4033 Stavanger, Norway
e-mail: jawad.raza@applysorco.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_36,

translates into an increased complexity of maintenance management activities. Consequently, there is an increased need for more effective equipment fault diagnosis and prognosis capabilities and efficient decision support systems. IO have significantly influenced the development and use of tools/systems and processes that churn out and store millions of data in various forms and formats. Thus, the biggest challenge lies in employing powerful problem solving tools/systems that effectively use all of such data [1]. These systems should be able to quickly obtain, effectively transform, and efficiently analyze information from multiple databases, for more reliable decision-making, mimicking human behavior/responses/decision-making.

The Norwegian petroleum industry expects IO to enhance health, safety, and the environment (HSE) standards as well as its productivity. Since maintenance activities have a direct consequence on asset reliability and availability, it is considered as one of the major impact areas. An appropriate mix of data, information, expertise, and technology is thus essential in ensuring that HSE standards are not compromised. This intelligent combination of data, information, expertise, and technology for enhanced decision-making purposes in complex environments is the building blocks for computer-based tools known as expert systems (ES).

Assuming that the IO implementation is progressing as envisaged by the Norwegian Oil Industry Association (OLF), the NCS should thus be getting to the tail end of first generation (G1) and already entered the second generation (G2) [2]. The following are some of the main characteristics that can be driving forces for such a progression:

- Condition monitoring (CM) techniques interwoven with almost all aspects of operation and maintenance.
- Use of decision support systems (e.g., ES) to better deal with complex data to diagnose and prognose equipment fault/failure
- Moving from preventive to more predictive maintenance strategies as the most prevalent business strategy.

Most ES fall within the domain of artificial intelligence (AI). AI is in the area of computer science, which employs familiar principles such as programming language, programming technique, algorithm, and data and data structure [3]. With reference to several definitions from different literary works, ES may be defined as a computer-based tool (software/application), designed for problem solving (through the execution-specific tasks) by mimicking the thought processes of human experts, thereby equipping the less skilled with some of the abilities of experts [4–6]. This implies that ES are domain-specific decision support systems.

The purpose of this paper is thus to present a holistic view on the state of the art of existing ES in use by large oil and gas companies in the NCS, specifically in the area of asset maintenance.

Thus, to effectively embark on this venture, a comprehensive understanding and knowledge about ES technology is deemed essential. Three important aspects highlighted in this paper are as follows:

1. How to differentiate ES from other systems;
2. How to establish a general ES application area; and
3. How to determine where maintenance can best deploy ES.

36.2 Methodology

Published literature (books, journal articles, conference proceedings, etc.) was studied with the primary purpose of identifying global oil and gas (O&G) applications of ES over the past 40 years. This was to serve as a reference point in the status review of the NCS. Since the main focus area of this study is asset maintenance, more efforts were directed at identifying ES applications within this domain.

A vendor survey was then conducted with the intention of gaining a holistic overview of the utilization of ES on the NCS. Vendors/suppliers/providers of hardware, software, and information management products/services were sampled. Companies that provide inspection, maintenance, and repair products/services were also added to this sample group. The most important select criteria, therefore, was that each company needed to have a footprint on the NCS. Consequently, products from 32 vendors/suppliers/providers, who either has (1) its head office registered/located in Norway, (2) one of its branch offices in Norway, or (3) O&G clients in the NCS, were considered. Thus, a total of 132 software applications/systems (statistical analysis systems, database and management systems, information management systems, project management systems, and data analysis systems) were sampled and studied. Due to the fact that information about applications/systems developed in-house is not readily available to the general public, the study excludes such applications/systems.

36.3 Expert Systems: Some Reflections

36.3.1 Distinguishing Features of ES

There are several books and publications on understanding the concept of ES, knowing the history behind their development, and their past and current modes of application. The main theme running through all of them is that ES helps equip non-experts with some of the skills and abilities of experts [3–7]. An interesting observation from more recent literary works is an emergent trend where ES applications are being developed as components of larger software bundles rather than standalones [4]. Consequently, for the purposes of this study, knowing how to identify these systems is most relevant. To be able to effectively identify these software/applications on the NCS, it is essential to be familiar with the characteristics, the structure, and the critical components of any ES. Table 36.1 thus provides a summary of some differences between ES and conventional computer systems. This table is based on a review of several literary works.

Table 36.1 Summary: expert systems versus convention computer systems

Aspect	Expert system	Conventional computer system
Focus area	• Knowledge	• Data, information
User interface	• Very interactive – Responds to queries – Asks for clarification – Makes recommendations	• Not quite interactive
Programming language	• e.g., LIPS, PROLOG, CLIPS, and OPS	• e.g., C, C++, and Fortran
Primary function	• Learning • Problem solving • Adapting • Decision-making • Explanation/investigation	• Data storage and retrieval • Data manipulation and representation
Processing techniques	• Both symbolic and algorithmic – Fuzzy logic, ANN, if/then rules, GA	• Primarily algorithmic – Mathematical algorithms
Search techniques	• Heuristics and algorithms	• Algorithms
Logic reasoning capacity	• Capable of logic reasoning	• Incapable of logic reasoning
Uncertainty application	• Capable	• Not capable

Despite the fact that current versions of database management systems (DBMS) possess functionalities beyond simple organizing and storing of data [8], it would be erroneous to simply conclude that ES are a type of DBMS. Subsequently, each DBMS must be individually assessed to determine whether or not it possess the characteristics indicated in the table.

36.3.2 ES Application Area

Research suggests that ES are mainly deployed for problem solving purposes, and as such, being able to establish which problem categories are most suitable for ES applications will foster our cause. Expert opinion also indicates that categorizing problems according to (1) their frequency of occurrence (recurring or non-recurring) and (2) their mode of propagation (structured or unstructured), and subsequently combining them provides a 4×4 problem matrix as depicted in Fig. 36.1. Unstructured problems are generally thought to be the most difficult to diagnose because they almost never happen in the same manner. Subsequently, specific expertise may be required to be able to make any diagnosis that is close to being accurate. Non-recurring problems are generally regarded as the most difficult to solve because not much information on them has been captured. Here also, specific expertise as well as some level of intuition may be required to be able to design solutions from first principles.

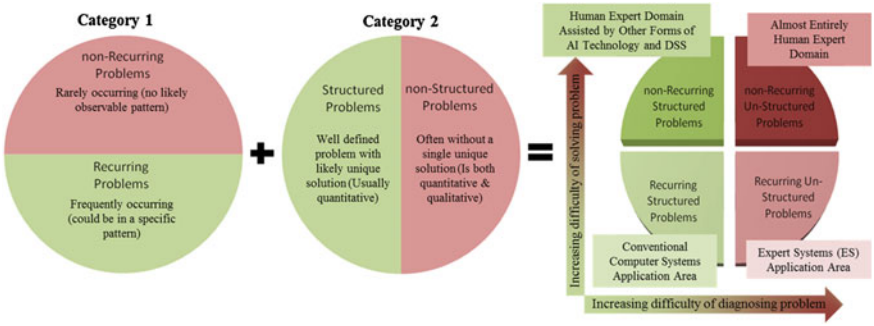


Fig. 36.1 Engineering problem categorization and ES application areas

Figure 36.1 suggests that recurring unstructured problems are perfect candidates for ES applications. Enough is known about them from their frequent occurrences. However, their unstructured nature makes them difficult to accurately detect and to decide which solution is most appropriate. Their frequency of occurrence also aids in the justification of the regularity of use of ES.

36.3.3 Deploying ES Within Asset Maintenance

Problem solving in asset maintenance (as in any other field of engineering), arrived at on the basis of the study, consists of four phases (being referred to here as the D^4 Process):

1. *Define*—identify/uncover the cause of the problem and describe it
2. *Design*—come up with possible solutions to the problem
3. *Determine*—choose the most appropriate solution based on the merits of the situation
4. *Deploy*—implement your chosen solution

Each of these phases requires expertise that may not always be immediately available. The unavailability of human experts increases the time frame within each phase of the D^4 process; subsequently, this may result in increasing asset downtime. Another factor that also adversely affects asset downtime (scarcity of expertise) is the limited capacity of human memory. It takes a longer time to process information, especially when this information is coming in large amounts and is not centrally located. An ES can thus be utilized for decision support in one or all phases of the D^4 process. Figure 36.2 depicts the possible time-saving potential of employing an ES in each phase of the problem solving process.

The design and use of expert systems for each phase make otherwise scarce expertise available for real-time decision-making without compromising consistency and quality of work output. An integrated ES can therefore be efficiently employed to help reduce time and cost of maintenance decision-making process.

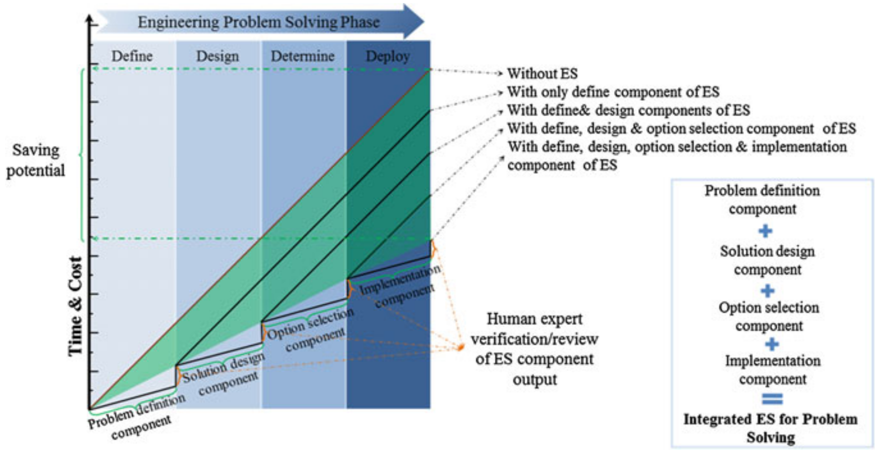


Fig. 36.2 ES potential in reducing time and cost of asset problem solving

36.4 Findings and Discussion

36.4.1 The Market Awareness

The global O&G industry has primarily seen the use of ES in geological applications—interpretation of seismic data, play analysis, and reservoir characterization (e.g., GEOPLAY [9]). Drilling of wells and well production have both experienced some notable applications as well (e.g., LSDO [10]). Some offshore design and construction applications are also known to exist (e.g., APDS). Several simulation systems (and to some extent, control systems, which has documented applications within the nuclear industry) are known to possess some ES components. Subsequently, we add operations to the list of O&G application areas.

Within the area of maintenance (and in a broader context, asset management), there is an apparent lack of ES specifically designed for O&G applications. Rail, automobile, aviation, and chemical process industries have several established ES's for their respective maintenance purposes. Most of which are employed for equipment fault diagnostics and the planning and scheduling of maintenance activities.

36.4.2 The Application Solutions/Software

The NCS is regarded as a leader in the development and implementation of groundbreaking technologies in the global O&G industry. Additionally, majority of the largest and most influential O&G companies have significant operations on the NCS. Consequently, the general picture of the utilization of ES on the NCS was expected to closely mimic that of the global O&G industry.

After analyzing all the identified software applications/systems and the opinions of some knowledgeable O&G industry professional, evidence was found to support our initial expectations for the NCS:

- Geology, drilling, and production seem to be the highest application areas.
- Operation has some applications; however, the numbers of software/systems that are unlikely to be considered ES are so great that we unable to conclusively say that ES applications are widespread.
- Environment/safety may have some applications but the investigation was not so conclusive.
- In the area of asset maintenance, ES had relatively very little application. The fault detection applications encountered seem to be more of other forms of AI applications than ES.

36.4.3 Developments in Maintenance Regimes

A further analysis of maintenance technologies led to an evaluation of past, prevailing, and expected future maintenance strategies on the NCS. The analogy here was that ES fosters the proliferation of predictive and proactive/dynamic maintenance strategies. The human–technology collaboration that utilizes all available data (design, performance, diagnostic, operator logs, and maintenance history) for timely and logical maintenance decisions is attainable via the application of much sophisticated technologies such as ES. Thus, knowing the level of predictive and proactive/dynamic maintenance strategies currently being practiced on the NCS would give us an indication of the level of sophisticated decision support systems (DSS) being employed. Figure 36.3 provides an illustration of the composition of maintenance strategies on the NCS.

The expectation under IO is for predictive and proactive/dynamic maintenance strategies to play a significant part in machine performance and uptime. However, the general feeling is that periodic/time-based maintenance (i.e., preventive maintenance) activities are still dominating the NCS (i.e., predictive maintenance and dynamic maintenance are playing only marginal roles). And as a result, it is likely that sophisticated decision support technologies (e.g., ES) are not being widely employed for maintenance purposes. Subsequently, it is our opinion that the NCS possess huge improvement potential for enhanced technology–human expert collaboration.

Within the purview of asset management, fast and efficient identification and description of maintenance problems are considered key to cultivating a predictive and dynamic maintenance culture. Subsequently, all the areas within the defined phase of the D^4 process (i.e., fault detection, diagnosis, prognosis, remaining useful life estimation, root cause analysis.) are suggested as immediate focus areas

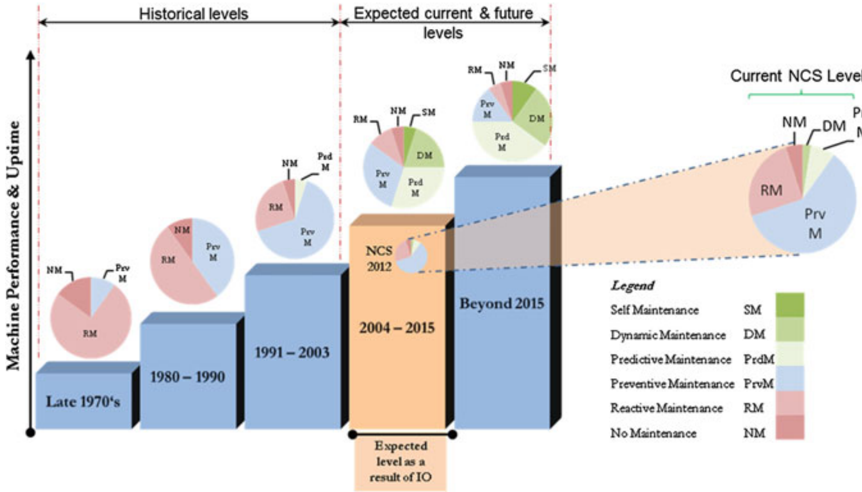


Fig. 36.3 Composition of maintenance strategies on the NCS

for ES application on the NCS. In principle, however, the NCS should be looking at value creation across all phases of the D^4 process. Reference is made to a follow-up research paper entitled, "*Value potential of Expert Systems for Offshore Oil & Gas assets from a Maintenance perspective.*" This follow-up paper elaborates on the utilization and value creation potential of ES within asset maintenance.

Inherent in IO is the challenge of efficiently utilizing data/information distributed in multiple databases, some of which may be geographically separated. Accordingly, industry professionals indicate that data analysts are constantly needed to assist in providing experts with the necessary information for decision-making, a scenario which they recognize as both expensive and slow. This study therefore suggests ES as a possible solution to this challenge within IO. ES (on the back of ICT) have the potential to enhance real-time decision-making by quickly and cost effectively utilizing available data/information. And it also has the potential to capture, store, and transfer scare expertise.

36.5 Conclusion

This paper has reflected on the role ES can play within the domain of offshore asset maintenance on the NCS. The study highlighted the status of ES in the field of O&G asset maintenance for which our preliminary findings seem to suggest that ES application areas on the NCS are similar to those of the global O&G industry. There also appears to be a lack of widespread adoption and use of ES in maintenance on the NCS which is perceived to be a direct consequence of limited knowledge and thorough understanding of the concept of ES by O&M engineers. Inappropriate branding (i.e., dissimilar nomenclatures) of ES products was identified in this study as a possible root cause of this challenge. Thus, on the basis of these initial findings, a more comprehensive study, which would also include in-house-developed software/applications, is recommended to verify the findings presented in the paper, thereby providing a sound basis for the design and development of new products/applications/services, which may well have significant economic and technical implications for different asset classes on the NCS.

Acknowledgments We would like to acknowledge all asset managers and maintenance engineers from the NCS who contributed to this study.

References

1. Liyanage J (2008) Integrated eOperations–eMaintenance: Applications in North Sea offshore assets. In: Murthy D, Kobbacy K (eds) Complex system maintenance handbook. Springer-Verlag London Limited, London, pp 585–610
2. Norwegian Oil Industry Association (OLF) (2005) Integrated work processes: future work processes on the Norwegian continental shelf. Stavanger
3. Krishnamoorthy C, Rajeev S (1996) Artificial intelligence and expert systems for engineers. CRC Press, Florida
4. Badiru A, Cheung J (2002) Fuzzy engineering expert systems with neural network applications. Wiley, New Jersey
5. Romem Y (2010) Darwin, culture and expert systems. In: Vizureanu P (ed) Expert systems, Intech, pp 1-16
6. Siler W, Buckley J (2005) Fuzzy expert systems and fuzzy reasoning. Wiley, New Jersy
7. Noran O (n.d.) The evolution of expert systems. Available at: http://www.ict.griffith.edu.au/noran/Docs/ES-Evolution.pdf. Accessed Apr 2012
8. Bassiliades N, Vlahavas I (2000) Active knowledge-based systems. In: Leondes CT (ed) Knowledged-based system: techniques and applications, Academic press, pp 1–34
9. Miller B (1995) GEOPLAY: a knowledge-based expert system—a model for exploration play analysis. U.S. Geological Survey Bulletin 2118, Washington
10. Ketata C, Satish M, Islam M (2005) Expert system knowledge management for laser drilling in the oil and gas industry, In: international conference on computational intelligence for modelling, control and automation and international conference on intelligent agents, web technologies and internet commerce, New York, vol 1, pp 930–938

Chapter 37
Value Potential of Expert Systems for Offshore Oil and Gas Assets from a Maintenance Perspective: A Case from Norway With Respect to Integrated Operations

Nii Nortey B. C. Lokko, Jawad Raza and Jayantha P. Liyanage

Abstract Expert systems (ES) broaden the possibilities for solving complex practical challenges, providing quick decision support, particularly when operations become complicated and collaborative. Such systems can contribute to improve overall technical integrity of offshore asset, subsequently creating value for oil and gas (O&G) companies. Consequently, this paper summarizes the results of the data collected through multiple case studies to investigate how sophisticated tools and technologies, such as ES, can contribute to improve the technical integrity of assets and add value to the Norwegian continental shelf (NCS) under the new operating environment known as integrated operation (IO). The paper highlights the potential of ES to be effectively deployed for real-time decision-making, enhancing predictive/dynamic maintenance capabilities, improving equipment reliability and availability, and optimizing work planning and resource allocation. Given the practical complexities of IO, the paper also identifies potential challenges, obstacles, and factors in the use of such advanced applications.

N. N. B. C. Lokko (✉) · J. P. Liyanage
Centre for Industrial Asset Management (CIAM), University of Stavanger, 4036 Stavanger, Norway
e-mail: nii.nortey.lokko@uis.no

J. P. Liyanage
e-mail: j.p.liyanage@uis.no

J. Raza
Operations and Engineering Services, Apply Sørco AS, Moseidsletta 122, Forus, 4033 Stavanger, Norway
e-mail: jawad.raza@applysorco.no

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_37,

37.1 Introduction and Background

The Norwegian continental shelf (NCS) is advancing more and more into deeper sea operations with significant subsea developments. It is faced with the challenge of making the most of marginal profitability fields as well as mature fields that are in the production decline phase. It also has to deal with managing the risks associated with unmanned installations coupled with an aging workforce whose expertise is in danger of being lost completely. And ultimately, it has to battle with ensuring compliance with health, safety, and environmental (HSE) regulations and standards that are continuously being stiffened. All these, according to industry professionals, translate into a need for value creation through the efficient use of real-time data for enhanced decision-making. This in turn would require the efficient preservation, development, and management of valuably scarce expertise. These needs thus contributed to the conception and initiation of the new collaborative environment concept known on the NCS as integrated operations (IO).

IO is commonly described as the, "integration of people, work processes, and technology to make smarter decisions and better execution [1]." This initiative, introduced by the Norwegian Oil Industry Association (OLF) about a decade ago, emphasizes the need to use "ubiquitous real-time data, collaborative techniques and multiple expertise across disciplines, organizations and geographical locations [2]." Ultimately, therefore, value creation is the expected outcome of this initiative. Value creation is in terms of enhanced HSE and optimized production. Whether it is through HSE or production, improving the functional and technical integrity of offshore assets plays a key role in creating value. Subsequently, we can relate certain aspects of value to asset maintenance under IO.

Within the domain of operations and maintenance (O&M), IO's practical application has been interpreted to encompass the following [3]:

- Testing out and implementing new technological solutions to especially enable predictive maintenance capabilities
- Implementing more robust technical platforms for effective O&M data management
- Establishing new organizational forms as compensation for the lack and/or shortage of experienced O&M workforce
- Standardizing the technical language used by different stakeholders for communication and cooperation enhancement purposes
- Providing fast access to technical expertise in challenging and urgent scenarios
- Building a lively competence network to enhance decision-making and the execution of activities
- Making quick and effective decisions based on data (data value creation).

The above list, according to industry professionals, highlights the need to employ sophisticated tools/systems that have the capacity to efficiently utilize the large volume of data that is being made available through the IO scenario. The concept of *expert systems* (ES) is an example of the tools/systems that, if efficiently employed, have the potential to significantly improve the real-time decision-making ability of O&M engineers and managers.

ES are computer-based tools designed for problem solving. They are designed to execute specific tasks (domain specific) by approaching problem solving as a human expert would. Therefore, it is considered as a tool that equips the less skilled with some of the reasoning and decision-making abilities of experts [4–7]. They are also commonly referred to as a type of decision support system (DSS) or knowledge-based systems (KBS) [8].

Our own preliminary studies reveal an apparent lack of widespread adoption and use of ES technology for asset maintenance on the NCS. Subsequently, the purpose of this paper is to investigate how sophisticated tools and technologies, such as ES, can contribute to improve the technical integrity of assets and create value in the NCS under IO.

As earlier indicated, the enhanced technical integrity of offshore assets can be related to effective asset maintenance. Technical integrity management, as we know it, is simply ensuring that facilities are in a sound condition (structurally and mechanically) such that they are able to perform and produce the outcomes they are designed for. These maintenance activities must therefore ensure that the assets are available and can be relied upon to deliver the expected outcome. Through the collaboration of people, technological systems, and processes/procedures, these maintenance activities can actually ensure asset availability and reliability, translating into enhanced technical integrity. It is the technological systems used to support maintenance decision-making and actions (e.g., sophisticated technology such as ES) that our attention is directed at here. IO is to realize a complex interactive environment of equipment, personnel, systems, processes, and organizations on the back of information and communication technology (ICT). Subsequently, the use of ES can only serve to foster the realization of a collaborative operating environment and improve the overall technical integrity of offshore assets. But for these systems to improve technical integrity, they must have a structure that is suitable for the task at hand. This means that they must possess (or have the capability to possess) functionalities relevant to the activities aiding the decision-making process.

37.2 Methodology

This study was conducted through multiple case studies using information obtained via two methods: interviews and questionnaires. The study was, however, limited to existing maintenance DSS within the field of asset maintenance. Four maintenance professionals from four O&G companies took part in the survey. Three of them were from large operating companies, and the other one was from a major maintenance service provider. The results and deductions are therefore limited by the data obtained from only these companies operating on the NCS.

Interview sessions were scheduled with highly experienced asset management professionals. Their expert insight was sought into "what," "why," and "how" ES/DSS are acquired in industry and also the identification of factors/challenges that affect their implementation. The availability/scarcity of expertise within maintenance and how this affected maintenance activities was also explored. The interviews also focused on IO's impact on the need for innovative technology and the acceptance of such technologies by employees within the O&G industry. The interviewees also expressed their professional views on the expertise requirement under IO, and the part ES/DSS plays (or can/will play) in the attainment of the O&M goals under IO.

The aim of the questionnaire administration was to assess the efficiency/effectiveness and impact of sophisticated technology (ES/DSS) for decision-making in maintenance. It covered three main system areas: *structure*, *functionality,* and *value impact*. Under system structure, the questionnaire respondents graded the domain specificity and knowledge base of a chosen maintenance system on a three-point scale. User friendliness, interoperability, reporting facility, large-volume data handling capacity, data uncertainty handling, response time, explanation facility, 24/7 online availability, knowledge acquisition capacity, symbolic processing capacity, and conflict resolution ability were graded on a five-point scale for system functionality. These functional areas were considered necessary for decision support within an IO environment. Finally, the respondents assessed the current and potential system impacts. Impact on productivity within maintenance, equipment availability and reliability, value-added gains, HSE activities, work planning and resource allocation, competence buildings, preventive/predictive/dynamic maintenance capacity, decision support, and expert task execution were graded on a five-point scale. These impact areas were considered relevant for value creation.

37.3 The Expert Systems Structure

Various researches (e.g., [4]) believe that the power of an ES lies in its ability to receive and combine factual and heuristic knowledge for complex problem solving. They therefore suggest that the knowledge setup within an ES be organized in a manner that fosters easy retrieval and in a format that can distinguish between data, control structures (parameters), and heuristics. Consequently, developers organize ES's around three main structures:

1. *Knowledge base*—this is the nucleus of all ES. It consists of a combination of the organized knowledge (a specific set of rules and procedures within the application domain for problem solving that have been captured by a knowledge engineer using knowledge representation techniques such as frames, semantic networks, and IF-Then rules) and the database (data and facts that may or may not be directly related to the application domain).

2. *Working memory*—this is where all the initial data about the problem are input/received, and the intermediate and final results/recommendations are displayed/retrieved.
3. *Inference engine*—this is the physical link between problem and possible solution. It is the control mechanism that organizes and matches knowledge in the knowledge base with the problem-specific data so that conclusions can be drawn and solutions can be found. It employs AI technologies such as ANN, GA, fuzzy logics, etc., that may be used singly or in combination.

37.4 Results and Discussion

In total, five maintenance tools/systems were studied and of these five systems (and based on an analysis of collected data), only one was considered to possess ES characteristics. To a certain degree, this is an indication of the alleged lack of widespread deployment of ES for asset maintenance on the NCS. This apparent lack of widespread ES adoption and use was perceived to be a direct consequence of a lack of knowledge and thorough understanding about the concept of ES by maintenance professionals. Proper product branding (or the lack of it for that matter) was highlighted as an underlying cause of the limited understanding of ES by maintenance professionals. A high-level approach involving the collaborative efforts of regulatory institutions and notable research and product development firms in ensuring that ES is appropriately branded was suggested as the first step in addressing this challenge. Other issues, such as the lack of confidence in unproven technology, embedded difficulties of the ES development process, lack of interest from service companies, and some perceived misconceptions about the implementation of IO, were also identified in this study as possible hindrances to the widespread adoption and application of ES and, in general, other sophisticated technologies.

The results of the study seem to suggest that, collectively, the maintenance systems may not be having their desire-valued impact on the NCS. Figure 37.1 below is an overall graphical representation of the estimated valued impact of maintenance systems on the NCS. It is derived from our five case studies by plotting the average rating given to each of the 9 impact areas. The impact assessment is lowest in the center with a rating of 0.0 (undeterminable impact/no expectation) and increases progressively toward the highest rating of 5.0 (significant impact/way above expectation). A rating of three is the minimum desirable and acceptable impact level. Thus, the further away a rating is from the center, the more desirable and valued the impact it has on maintenance activities on the NCS.

The consistency and quality of work output, together with productivity, seem to be highest valued impact areas. This we consider complementary of the systems. However, in an IO environment, lack of the desired impact on real-time decision-making, preventive/predictive/dynamic maintenance capabilities, and work

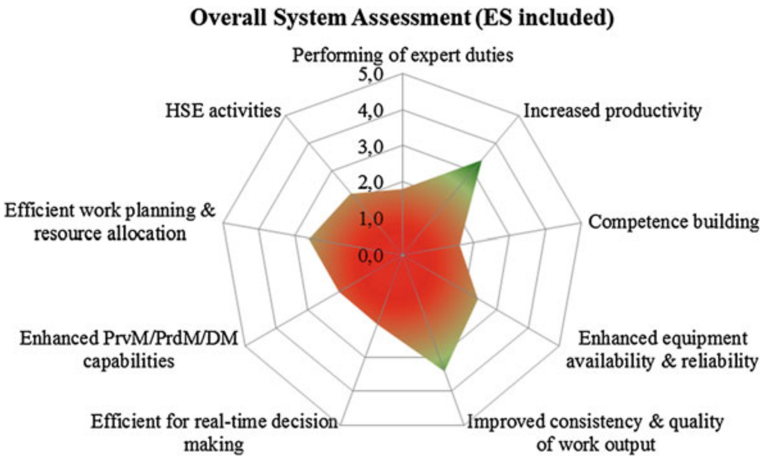


Fig. 37.1 Graphical representation of the overall impact of maintenance systems on the NCS (including ES)

planning and resource allocation is a conspicuous deficiency that would need attention. These deficiencies coupled with little or no impact on enhancing competence building and the performing of expert duties give the indication of lack "expertness"/sophistication in the maintenance systems on the NCS.

In this study, we employed an ES (sophisticated DSS technology) accept criteria of at least three. This is the least grade we expect any system considered to be an ES to have. Our argument for this acceptance criterion was that, since we rely on our human experts to assist us in enhancing our value creation process, we expect nothing less than quality work. Consequently, if we are to employ ES in our operations, the least-valued impact we can tolerate/expect is exactly what a human expert would have been expected to deliver (i.e., the system should meet and/or exceed or expectation). Any delivery/impact that is *below* our expectation is *not* considered value for money. Subsequently, we deduced from Fig. 37.1 that the overall impact assessment of the systems were below expectation (average of about 2.4). Hence, this adds to the perceived lack of widespread "expertness"/ sophistication in the maintenance systems on the NCS.

The results also seem to suggest a positive relationship between ES application and the value impact of maintenance systems/software on the NCS. When the ES's ratings were omitted from the analysis, the overall value impact assessment dropped by about 17 % (from 2.4 to 2.0), whereas the omission of the other non-ES resulted in less than a ±4 % difference in the overall assessment. The omissions were performed to investigate the impact each system had on the overall analysis. The impact of ES on the overall assessment can be seen in Fig. 37.2 where, in comparison with Fig. 37.1, the shaded area has diminished quite significantly.

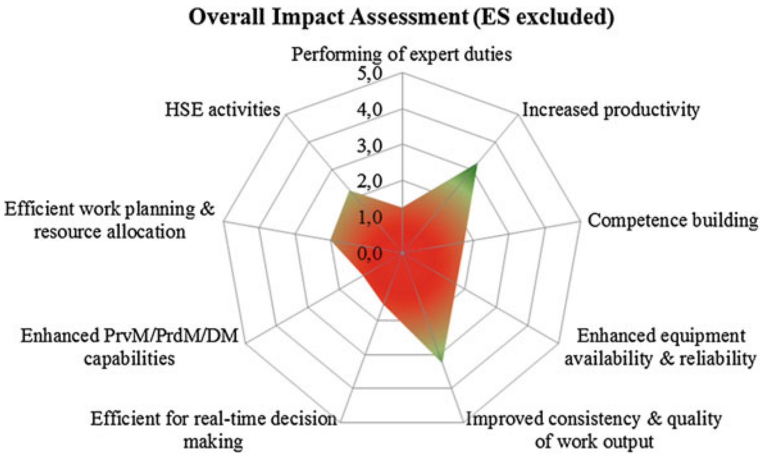


Fig. 37.2 Graphical representation of the overall impact of maintenance systems on the NCS (excluding ES)

The most affected value areas due to the omission of ES ratings were real-time decision-making, preventive/predictive/dynamic maintenance capabilities, equipment reliability and availability, performing of expert duties, and work planning and resource allocation. On the basis of this analysis, we could sufficiently infer that ES fosters the realization of the maintenance goals within an IO environment, i.e., value creation within maintenance as described in the introduction section of this paper.

Also, the study generally confirmed the notion that knowledge is more valuable than data and/or information. Subsequently, KBS's are most valuable to O&M managers and engineers. The system structural assessment part of this study indicated that for a system to effectively utilize a knowledge base, an interactive user interface, heuristics programming, and an inference engine should be present in the system structure. Anything short of this would require considerable human expertise to effectively link this knowledge to the problem at hand.

Figure 37.3 is a combination of 2 value pyramids: intelligence/wisdom is at the top of the decision (bigger) pyramid, and knowledge bases are at the top of the technology (smaller) pyramid. Since technological systems are currently not yet established in the intelligence/wisdom section, the height of the technology pyramid only reaches the knowledge section of the decision pyramid. Hence, Fig. 37.3 simply suggests that if the selection of maintenance tools/DSS is based on highest value creation potential, then ES should be one of the first to be selected. Since an ES is a KBS, it therefore possesses significant value creation potential.

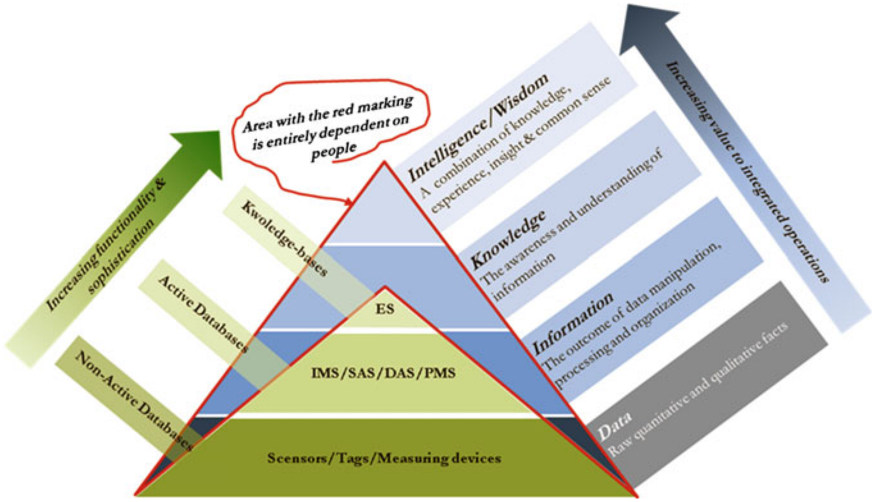


Fig. 37.3 Value assessment from data to intelligence (The initial figure was adopted from reference [9])

37.4.1 Potential for Expert Systems Utilization and Value Creation

With reference to the complementary research paper entitled, "*The Use of Expert Systems in Offshore Oil and Gas Assets: A Status Review With Respect to Emerging Needs for Innovative Solutions,*" it is therefore envisaged that the time and cost saving potential of ES within maintenance of offshore assets on the NCS is quite significant. Figure 37.2 from that paper was thus revised to illustrate the significant value creation potential being suggested in this study.

This (Fig. 37.4) thus indicates that the NCS seems to be utilizing ES applications (sophisticated technologies) for a comparatively small portion of its maintenance activities. Potentially large time and cost savings could be achieved in real-time decision-making if ES utilization can be extended to encompass the design, determine, and deploy phases of problems solving. There may also be the potential to develop an integrated ES (essentially domain specific) for maintenance problem solving that would incorporate all these phases in one application.

Essentially, the outcome of this study points to the need for more ES for value creation through enhanced asset maintenance:

1. IO is directed at transforming data/information into knowledge for decision-making → ES are technological systems that use knowledge bases efficiently/effectively.
2. IO is directed at dynamic operating regimes → ES fosters and enhances the quality of real-time decision-making, improves predictive and dynamic maintenance capabilities, and has the functional capacity to handle uncertainties.

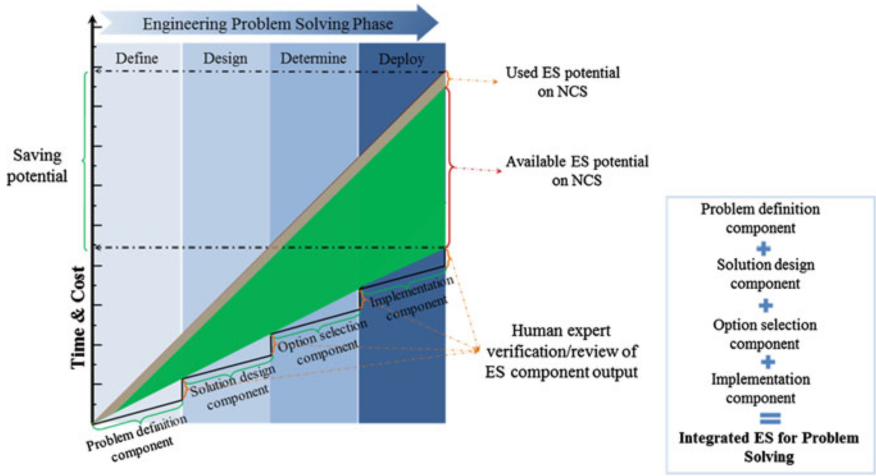


Fig. 37.4 Available asset problem-solving time and cost-reducing potential on the NCS

3. IO is directed at enhancing HSE → ES has functionalities that foster asset availability and reliability, which in turn influences the technical integrity of offshore assets.
4. Ultimately, IO is directed at enhancing value creation → ES is perceived to have the highest value creation potential.

37.5 Conclusion

The NCS has the potential to benefit immensely from more ES applications for asset maintenance. The study suggests that effectively deploying ES could assist in optimizing the technical integrity of O&G assets. This paper thus infers that significant maintenance time and cost savings stemming from widespread adoption and use of ES technologies are possible on the NCS. Areas such as real-time decision-making, enhancing predictive/dynamic maintenance capabilities, improving equipment reliability and availability, and optimizing work planning and resource allocation are prime ES impact areas that would be of worth in the IO environment. The study thus concludes that IO has created a conductive environment for the application of sophisticated technologies (such as ES) for value creation, especially within the maintenance discipline. Whether through enhanced HSE standards or through improved productivity of offshore installations, the optimization of the technical integrity of assets on the NCS can be achieved with the aid of ES.

Acknowledgments We would like to acknowledge all contributors to this study.

References

1. Norwegian University of Science and Technology (n.d.) Center for Integrated Operations in the Petroleum Industry. Available at http://www.ntnu.edu/iocenter. Accessed Apr 2012
2. Norwegian Oil Industry Association (2005) Integrated work processes: future work processes on the Norwegian continental shelf. Stavanger
3. Liyanage J (2008) Integrated eOperations—eMaintenance: applications in North Sea offshore assets. In: Murthy D, Kobbacy K (eds) Complex system maintenance handbook. Springer, London, pp 585–610
4. Badiru A, Cheung J (2002) Fuzzy engineering expert systems with neural network applications. Wiley, New York
5. Siler W, Buckley J (2005) Fuzzy expert systems and fuzzy reasoning. Wiley, New York
6. Noran O (n.d.) The evolution of expert systems. Available at http://www.ict.griffith.edu.au/noran/Docs/ES-Evolution.pdf. Accessed Apr 2012
7. Krishnamoorthy C, Rajeev S (1996) Artificial intelligence and expert systems for engineers. CRC Press, Boca Raton
8. Bassiliades N, Vlahavas I (2000) Active knowledge-based systems. In: Leondes CT (ed) Knowledged-based system: techniques and applications, Academic Press, USA, pp 1–34
9. Dwyer J, Chmela W, Lamborn R, Colomb C, Quinn T (2008) Real-time connectivity and potential benefits to Appalachian operations, in SPE eastern regional/AAPG eastern section joint meeting. Pittsburg, Pennsylvania

Chapter 38
Detection of Excessive Diesel Engine Piston Slap Using Acoustic Emission Signals

David P. Lowe, Weiliang Wu and Andy C. C. Tan

Abstract This paper presents a study whereby a series of tests was undertaken using a naturally aspirated, 4-cylinder, 2.216 L, Perkins diesel engine fitted with a piston having an undersized skirt. This experimental simulation resulted in engine running conditions that included abnormally high levels of piston slap occurring in one of the cylinders. The detectability of the resultant diesel engine piston slap was investigated using acoustic emission signals. Data corresponding to both normal and piston slap engine running conditions were captured using acoustic emission transducers along with both in-cylinder pressure and top-dead centre reference signals. Using these signals, it was possible to demonstrate that the increased piston slap running conditions were distinguishable by monitoring the piston slap events occurring near the piston mid-stroke positions. However, when monitoring the piston slap events occurring near the TDC/BDC piston stroke positions, the normal and excessive piston slap engine running conditions were not clearly distinguishable.

38.1 Introduction

Piston slap is a widely recognised mechanical impact phenomenon common to both reciprocating internal combustion engines and compressors. This phenomenon occurs when the piston undergoes transverse motion within the clearance between the piston and the cylinder bore causing it to impact on the cylinder wall. The transverse motion of the piston is the result of a direction change in the sideways acting force component of the overall piston force.

D. P. Lowe · W. Wu · A. C. C. Tan (✉)
Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: a.tan@qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_38,

The mechanical impacts generated by piston slap are generally acknowledged as being a major source of noise in diesel engines, and indeed, this has been one of the major motivating factors that have driven research into both the mechanics of piston slap and the methods by which the vibration and noise generated by piston slap may be reduced.

Earlier examples of research pertaining to piston slap as a source of noise in diesel engines include studies by Fielding and Skorecki [1] and Haddad and Pullen [2], while more recent studies by Okubo et al. [3], Haddad and Tjan [4], and Nakashima et al. [5] have also examined the relationship between piston slap and emitted engine noise.

All of these studies highlight the strong influence that the piston-cylinder bore clearance has on the severity of the resulting piston slap. In addition, the studies by Okubo et al. [3], Haddad and Tjan [4], and Nakashima et al. [5] have also shown that gudgeon pin eccentricity also has a major influence on the overall severity and nature of the piston slap. And indeed, the use of pistons having offset gudgeon pin positions is now common.

Perhaps more importantly from a reliability or condition-monitoring point of view, excessive piston slap has been identified as being a major cause of cavitation erosion damage of both diesel engine cylinder liners and cylinder blocks [6–8]. The cavitation erosion occurs on those cylinder liner or block surfaces that are in contact with engine coolant with the erosion being most obvious on the major thrust side of the cylinder bore [7]. Yonezawa and Kanda [8] showed that cylinder liner vibration caused by piston slap generated large cooling water pressure fluctuations. These cooling water pressure fluctuations in turn generate the rapid formation and destruction of air bubbles which lead to cavitation erosion.

38.2 Occurrence and Nature of Piston Slap

There are two sets of conditions that cause the piston to undergo transverse motion within the cylinder bore [2]; these conditions occur in the following conditions:

1. When the horizontal force component at the gudgeon pin changes direction due to the rotational motion of the crankshaft and resulting change in connecting rod geometry
2. When the connecting rod undergoes force reversals from compression to tension (and vice versa).

It has been found that the above-mentioned conditions can produce six piston slap events during one four-stroke engine cycle [1, 9].

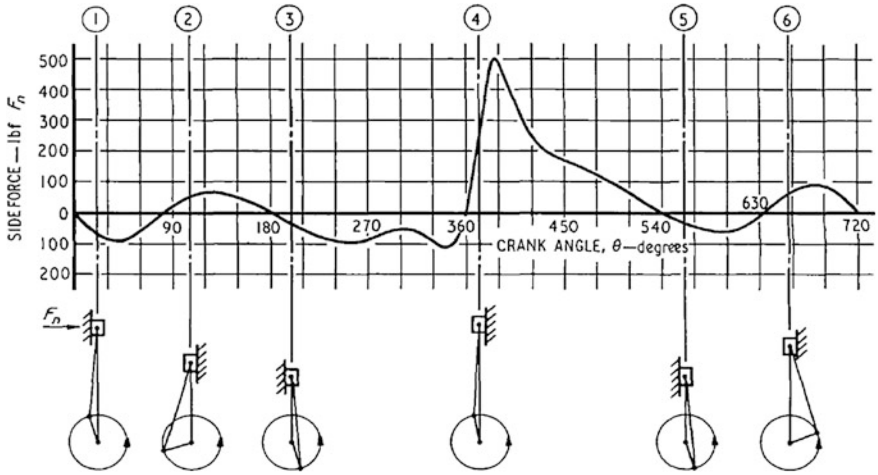


Fig. 38.1 A plot of the piston side force as well as the approximate piston slap and linkage positions in terms of crank angle [1]

38.2.1 Piston Slap Occurrence

The first condition mentioned above can generate four piston slap events during one complete (four-stroke) engine cycle. These four points occur at or near TDC and BDC, see Fig. 38.1 as events 1, 3, 4, and 5. The piston slap event at position 4 corresponds with the combustion event. The other two piston slap events labelled in Fig. 38.1 as events 2 and 6 occur due to the second condition listed above and occur at or near the mid-stroke positions.

Among other findings, Fielding and Skorecki [1] found that for increases in piston-cylinder bore clearance, the impact energy associated with non-combustion piston slap events (events 1, 2, 3, 5, and 6 in Fig. 38.1) increased in consistent manner, while the impact energy increase from the combustion-related piston slap event was approximately three times greater.

38.2.2 Description of Transverse Piston Motion

The motion of the pistons within reciprocating internal combustion engines and compressors is often referred to as either primary or secondary motions. The primary piston motion constitutes the displacement, velocity, and acceleration of the piston along the cylinder axis [10]. These are readily calculated as functions of crank angle [11].

Secondary piston motion refers to the lateral translational and rotational motions experienced by the piston within the piston-bore clearance due to the

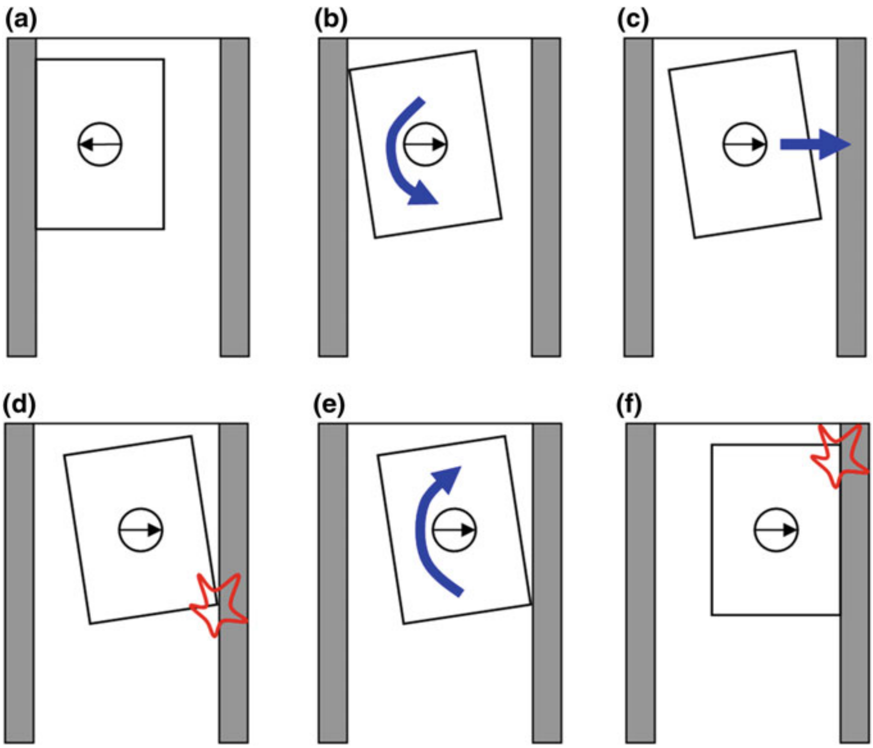


Fig. 38.2 The six stages of secondary piston motion as detailed by Fielding and Skorecki [1]

action of side forces [10]. Fielding and Skorecki [1] considered the transverse motion of the piston in six stages. These six stages are shown in Fig. 38.2.

As Fig. 38.2a shows, during the first stage, the piston travels within the cylinder with the entire piston skirt surface resting on the cylinder wall. As force changes direction, one end of the skirt lifts off the cylinder wall and the piston rotates, while the other end of the skirt remains in contact with the cylinder wall (Fig. 38.2b). The piston translates across the cylinder bore (Fig. 38.2c). The first impact occurs when one end of the skirt makes contact with the opposite side of the cylinder wall (Fig. 38.2d), after which the piston rotates about the end of the piston skirt in contact (Fig. 38.2e), until the other end of the skirt comes into contact with the cylinder wall resulting in the second impact (Fig. 38.2f). More recent studies undertaken by Haddad and Howard [12], Haddad and Tjan [4], and Haddad [13] describe piston motion using four modes of motion shown in Fig. 38.3. These modes of motion are as follows:

1. No horizontal piston movement
2. The piston rotates about the top of the piston skirt
3. The piston rotates about the bottom of the piston
4. The piston moves freely within the cylinder bore.

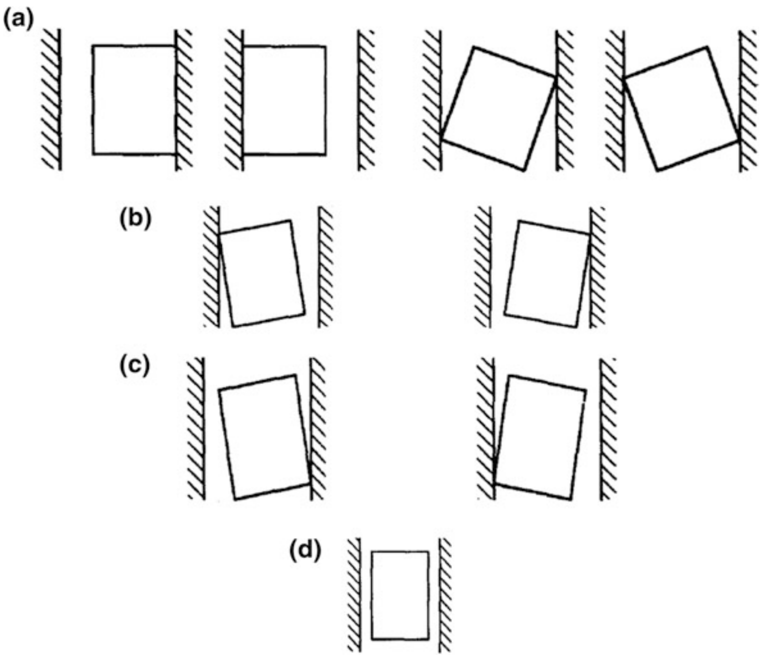


Fig. 38.3 The four modes of piston motion; **a** Vertical or diagonal contact, **b** Rotation about *top* edge, **c** Rotation about *bottom* edge **d** Free movement within bore [4, 12, 13]

Figure 38.3a shows the first mode of piston motion. As seen, both the top and bottom of the piston skirt in contact with the cylinder wall either vertically or diagonally. Figure 38.3b shows the second mode of piston motion, and as seen, the piston rotates about the top edge of the piston skirt with respect to either the major or the minor thrust sides of the cylinder bore. Similarly, Fig. 38.3c shows the third mode of piston motion whereby the piston rotates about the bottom edge of the piston skirt. Figure 38.3d shows the final mode of piston motion, and as seen, the piston is able to move freely within the bore according to the net horizontal force and moment acting on it.

38.3 The Experimental Campaign

The experimental campaign presented in this paper was undertaken at Queensland University of Technology using the diesel fault test rig (DFTR) shown in Fig. 38.4. The DFTR consists of a naturally aspirated, 4-cylinder, 2.216 L, Perkins diesel engine coupled to an Olympian 415 V, three-phase alternator. Table 38.1 lists all the relevant engine specifications.

Fig. 38.4 The diesel fault test rig

Table 38.1 Perkins 404C-22 diesel engine specifications

Engine make and model	Perkins 404D-22
Layout and number of cylinders	Inline, 4 cylinder
Engine cycle	Four stroke
Induction	Naturally aspirated
Engine speed	1,500 RPM (governed)
Gross engine power	Standby: 18 kW
	Prime: 16.2 kW
Firing sequence	1, 3, 4, 2
Bore	84.0 mm
Stroke	100 mm
Displacement	2216 c.c.
Compression ratio	23.3:1
Injection timing—start	15° before TDC

38.3.1 Experimental Setup

In order to induce an abnormal level of piston slap into the DFTR, one of the original pistons was removed and replaced with a piston having an undersized skirt diameter. The piston skirt diameter was machined from the initial diameter of 83.96 mm to 83.70 mm. This diameter is listed as being the piston skirt diameter service limit for the engine used in this study [13]. In qualitative terms, this experimental setup approach simulated a piston having skirt-cylinder wall clearance consistent with an engine reaching the end of its service life. Figure 38.5a shows the original piston, while Fig. 38.5b shows the undersized piston.

The DFTR was operated with both normal and undersized pistons using two engine load settings. These load settings were as follows: 0 and 15 kW. For the DFTR, these settings nominally represent no-load and full-load engine running conditions, respectively.

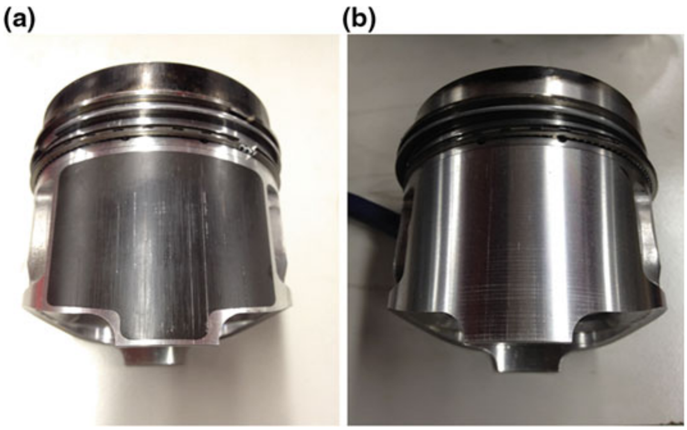
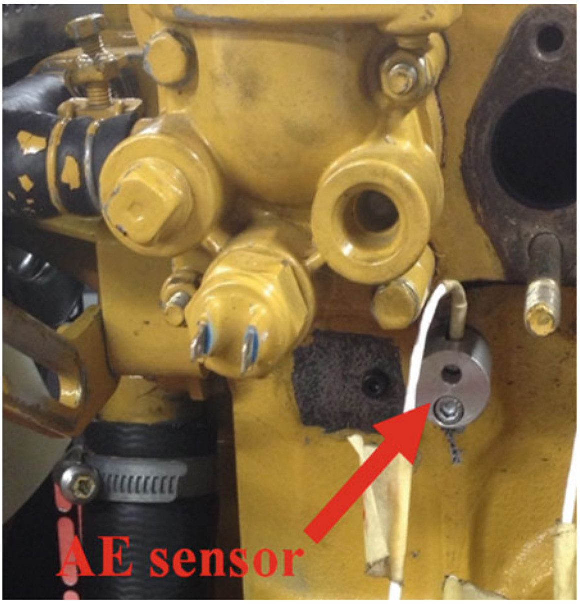


Fig. 38.5 Photographs showing **a** normal and **b** undersized pistons

Fig. 38.6 Photograph showing the AE sensor position



38.3.2 Data Acquisition

AE signals were acquired from the Perkins engine using a PAC MICRO-30D sensor mounted to the block of the engine in close proximity to cylinder 1. This sensor is shown in Fig. 38.6. AE signals were amplified using PAC 2/4/6 pre-amplifiers before being recorded and digitized using a PAC MicroDisp AE DAQ system and a Toshiba laptop computer.

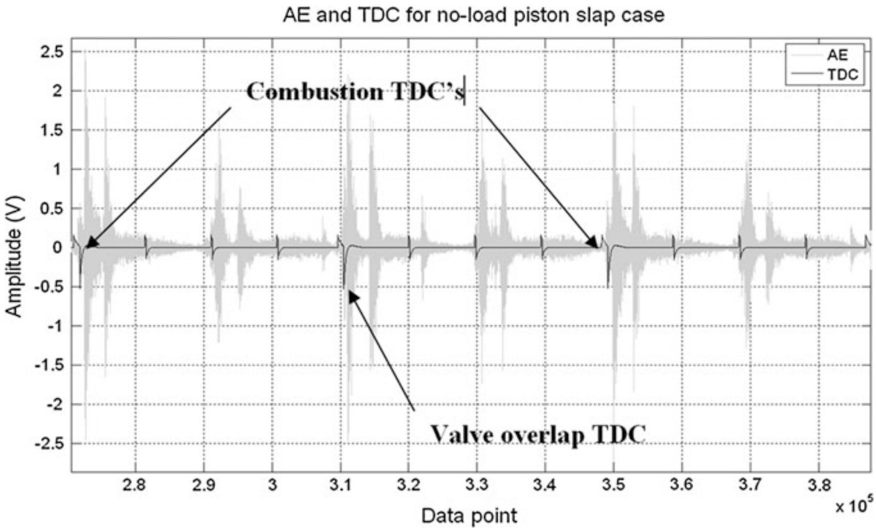


Fig. 38.7 Plot showing the simultaneously recorded AE and TDC signals

In addition to the AE signals, TDC and pressure reference signals were also recorded using the AE DAQ system. A plot showing the raw AE and TDC signals is shown in Fig. 38.7. As seen, the TDC and AE signals have been recorded simultaneously allowing the use of synchronous averaging techniques. In addition to the TDC and AE signals, AE and pressure signals were also simultaneously recorded in order to identify those portions of the signal associated with combustion TDC as opposed to the valve overlap TDC. Both of these are labelled in Fig. 38.7.

38.4 Results

The AE signals recorded during the various engine running conditions were synchronously averaged using the TDC signal and converted into AE rms. These averaged AE rms signals were then used to analyse the different running conditions in terms of piston slap-related events within windows determined using the approximate piston slap crank angle positions detailed in Fig. 38.1. Averaged AErms signals for the no-load engine-running case are shown in Fig. 38.8. Also shown are the six windows relating to the approximate piston slap crank angle positions mentioned above. As seen in Fig. 38.8, the four piston slap events associated with the TDC and BDC stroke positions are labelled as events 1, 3, 4, and 5. The other two mid-stroke-related piston slap events are labelled as events 2 and 6 on the plot.

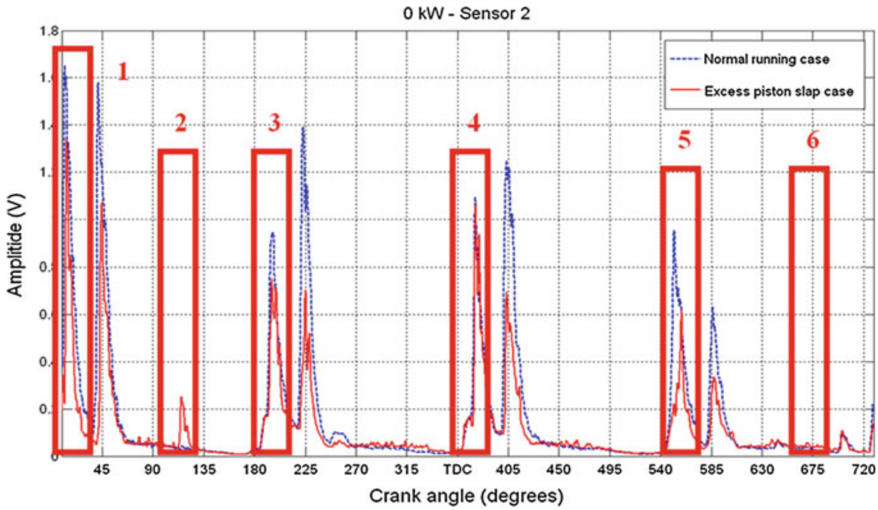


Fig. 38.8 Plot of averaged AErms for both normal and excess slap for the no-load case

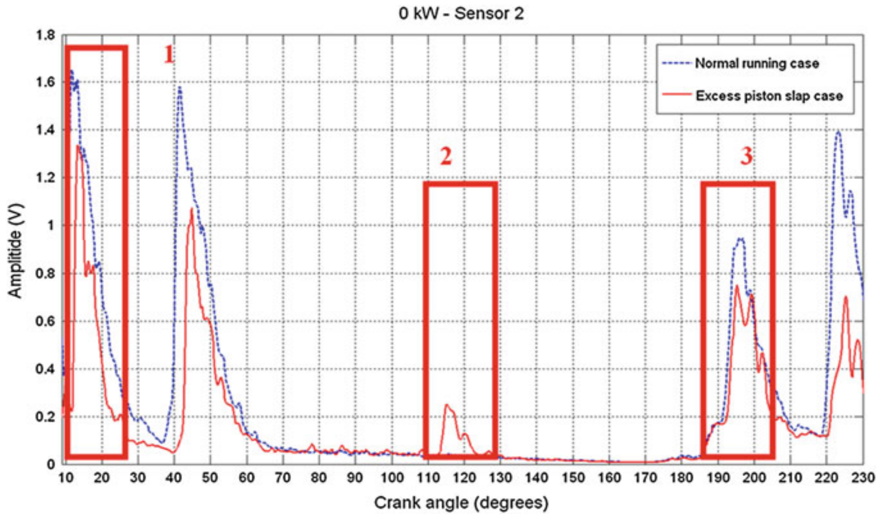


Fig. 38.9 Close-up view of windows *1*, *2*, and *3* for both normal and excess slap no-load cases

Figure 38.9 shows a close-up view of the first three piston slap position windows shown in Fig. 38.8. As seen, qualitative comparisons with respect to windows 1 and 3 show no indication of the excess piston slap. As a matter of fact, energy levels are seen to decrease for the excess slap case. The distinct peak occurring at approximately 115° within window 2 however is a clear indication of the increased piston slap.

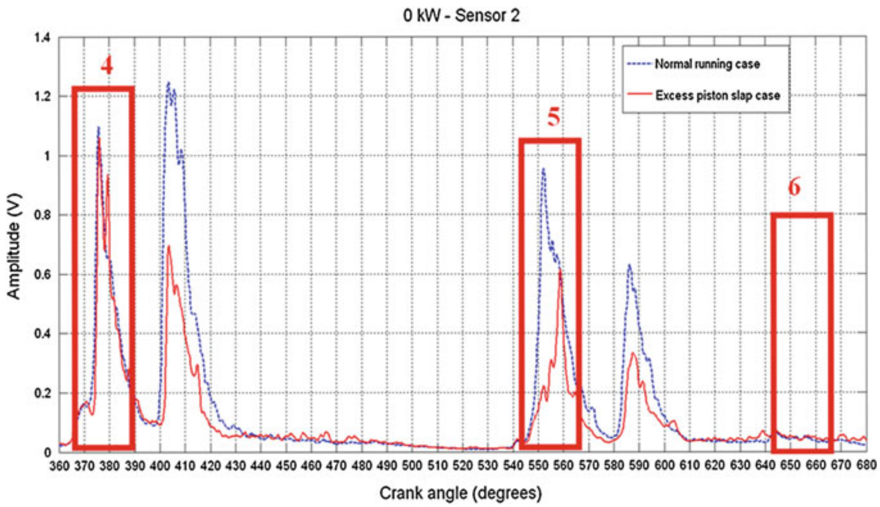


Fig. 38.10 Close-up view of windows *4*, *5*, and *6* for both normal and excess slap no-load cases

Table 38.2 Total window energy values for the two different engine load settings

Window	Running condition	Load setting	
		0 kW	15 kW
1	Normal	85.06	49.69
	Slap	51.08	20.49
2	Normal	3.66	2.65
	Slap	9.86	4.57
3	Normal	53.34	18.63
	Slap	41.45	14.34
4	Normal	47.61	10.52
	Slap	44.46	12.02
5	Normal	51.59	32.29
	Slap	25.74	15.61
6	Normal	4.58	3.25
	Slap	5.15	3.72

Figure 38.10 shows a close-up view of the last three piston slap position windows shown in Fig. 38.8. As seen, qualitative comparisons with respect to windows 4, 5, and 6 show no obvious indication of the excess piston slap, and indeed, window 5 shows a decreased energy level for the excess slap case.

In order to quantify these differences, comparisons were made between the normal and piston slap running conditions using total window energy. This total energy value was calculated by summing the AErms over the 20° window length. This procedure was performed for the six different windows in conjunction with the two different engine load settings mentioned in Sect. 38.3.1, namely 0 and 15 kW (no load and full load). Table 38.2 lists all of the calculated total window energy values.

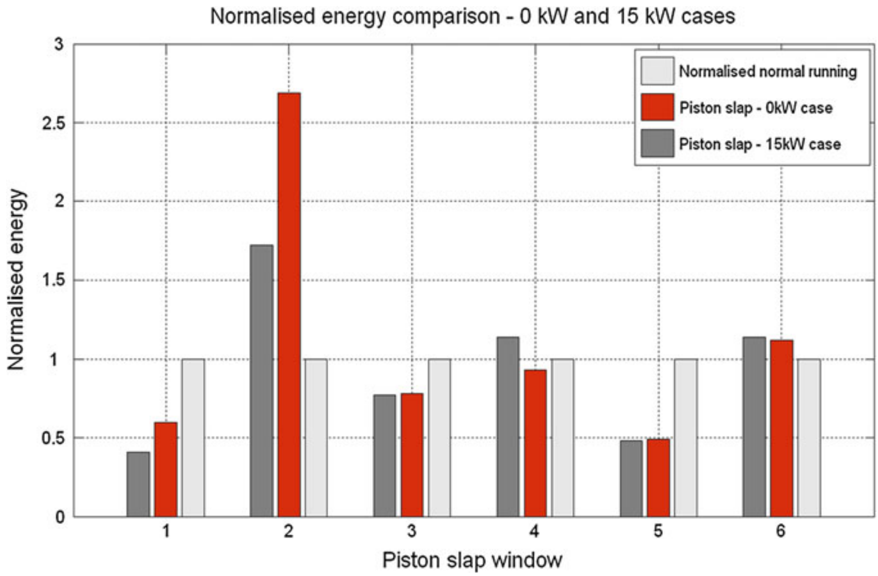


Fig. 38.11 Normalised total window energy values for the six piston slap windows

As seen, 3 of the 4 TDC/BDC-related windows (windows 1, 3, 4, and 5) show lower levels of total window energy between the normal and piston slap running conditions. The only increase seen in total window energy occurred for window 4 during the 15 kW load setting. This minor window 4 energy increase represented an increase of 14.26 % over the normal running case. It was also noticed that total window energy levels for both normal and excess piston slap running conditions decrease with load.

Figure 38.11 further highlights these findings by presenting the data listed in Table 38.2 in a normalised fashion whereby the normal running condition total window energy value was used to normalise the corresponding piston slap total window energy value.

As was seen in Table 38.2, the non-combustion TDC/BDC windows (windows 1, 3, and 5) shown in Fig. 38.11 all show reductions in total window energy for the piston slap running conditions for both engine loads. As was briefly mentioned, the only total window energy value inconsistent with the observed trend is the value corresponding to window 4 for the full-load case.

As window 4 relates specifically to the combustion TDC, it is postulated that this anomaly in the observed data is due to the increased combustion effects discussed in Sect. 38.2.1 for window 4 as compared to the other windows. The mid-stroke windows (windows 2 and 6), however, both show increases in total window energy.

While the increases in total window energy for window 6 were relatively minor, the piston slap event occurring in the mid-stroke position during the intake stroke (window 2) shows an increase in total window energy for the increased piston slap running conditions. As percentages of the normal running cases, these increases were 269 % for the 0 kW case and 172 % for the 15 kW case.

38.5 Conclusions and Future Work

The general trend regarding the TDC/BDC-related windows (windows 1, 3, 4, and 5) indicates that the increased level of piston slap was not clearly reflected in these portions of the signal. The mid-stroke-related windows (windows 2 and 6), however, both show increases in total window energy. Window 2 in particular showed a dramatic increase in total window energy for the increased piston slap running conditions.

With respect to the TDC/BDC-related windows, the only rise in total window energy corresponding to increased piston slap occurred in window 4 during the 15 kW load setting. It was postulated that the minor increase in total energy was due to the increased combustion effects. The specific cause of this trend, as well as the dramatic increase in piston slap-related total window 2 energy, could be established by the analysis of the experimental data in conjunction with a detailed dynamic piston slap model specific to the DFTR.

It is further postulated that the lack of indication concerning the presence of piston slap given by the TDC/BDC-related windows can be attributed to two major factors; the first being the substantial obstacle to AE propagation presented by the coolant-filled engine water jacket, while the second major factor is the large amount of other AE-generating events that occur near the TDC/BDC positions. These events include premixed (kinetic) combustion pressure fluctuations, exhaust and inlet valve impacts, gas blowby, and frictional regime changes. It is worth noting that in contrast to the numerous AE-generating events occurring at the TDC/BDC positions, the two mid-stroke positions contain few other sources of AE. This makes these positions very suitable from a monitoring point of view.

It is possible that these obstacles could be negated to a large degree by providing a more direct source-to-sensor path either by careful sensor position selection with respect to the engine block geometry, or with the use of a solid body of material to bridge the gap across the engine water jacket. In practical terms, this could involve the use of a threaded rod or bolt.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centres Program. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Fielding BJ, Skorecki J (1969) Identification of mechanical sources of noise in a diesel engine: sound originating from piston slap. Proc Inst Mech Eng 184(1):46
2. Haddad SD, Pullen HL (1974) Piston slap as a source of noise and vibration in diesel engines. J Sound Vib 34
3. Okubo M, Kanda H, Yonezawa T (1989) Analysis and reduction of piston slap noise in diesel engines. SAE technical paper 890127
4. Haddad SD, Tjan K (1995) An analytical study of offset piston and crankshaft designs and the effect of oil film on piston slap excitation in a diesel engine. Mech Mach Theory 30(2):271–284
5. Nakashima K, Yajima Y, Suzuki K (1999) Approach to minimization of piston slap force for noise reduction—Investigation of piston slap force by numerical simulation. JSAE Rev 20:211–216
6. Zhou Y, He J, Hammitt FG (1982) Cavitation erosion of diesel engine wet cylinder liners. Wear 76:321–328
7. Zhou Y, He J, Hammitt FG (1982) Cavitation erosion of cast iron diesel engine liners. Wear 76:321–328
8. Yonezawa T, Kanda H (1989) Analysis of cavitation erosion on cylinder liner and cylinder block. SAE technical paper 850401
9. Geng Z, Chen J (2005) Investigation into piston-slap induced vibration for engine condition simulation and monitoring. J Sound Vib 282:735–751
10. Mansouri SH, Wong VW (2005) Effects of piston design parameters on piston secondary motion and skirt-liner friction. Proceedings of the Institution of Mechanical Engineers, Part J: J Eng Tribol
11. Dong ZD, Cheng HS, Arai T, Hamai K (1992) A numerical analysis for piston skirts in mixed lubrication—Part 1: basic modelling. J Trbol 114:553–562
12. Haddad SD, Howard DA (1980). Analysis of piston slap-induced noise and assessment of some methods of control in diesel engines. SAE technical paper 800517
13. Perkins Service Manual 400C Series. Service manual publication number: RENR9825

Chapter 39
Towards an Integrated Maturity Model of Asset Management Capabilities

Muhammad Nateque Mahmood, Subas Prasad Dhakal, Anna Wiewiora, Robyn Keast and Kerry Brown

Abstract Asset service organizations often recognize asset management as a core competence to deliver benefits to their business. But, how do organizations know whether their asset management processes are adequate? Asset management maturity models, which combine best practices and competencies, provide a useful approach to test the capacity of organizations to manage their assets. Asset management frameworks are required to meet the dynamic challenges of managing assets in contemporary society. Although existing models are subject to wide variations in their implementation and sophistication, they also display a distinct weakness in that they tend to focus primarily on the operational and technical level and neglect the levels of strategy, policy, and governance as well as the social and human resources—the people elements. Moreover, asset management maturity models have to respond to the external environmental factors, including climate change and sustainability, stakeholders, and community demand management. Drawing on five dimensions of effective asset management—spatial, temporal, organizational, statistical, and evaluation—as identified by Amadi-Echendu et al. [1], this paper carries out a comprehensive comparative analysis of six existing maturity models to identify the gaps in key process areas. Results suggest incorporating these into an integrated approach to assess the maturity of asset-intensive organizations. It is contended that the adoption of an integrated asset management maturity model will enhance effective and efficient delivery of services.

M. N. Mahmood (✉)
University of Southern Queensland, West Street, Toowoomba, QLD 4350, Australia
e-mail: nateque.mahmood@usq.edu.au

S. P. Dhakal · K. Brown
Curtin University, Kent Street, Bentley, WA 6102, Australia

A. Wiewiora
Queensland University of Technology, George Street, Brisbane, QLD 4001, Australia

R. Keast
Southern Cross University, Southern Cross Drive, Bilinga, QLD 4225, Australia

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_39,

39.1 Introduction

Asset service organizations' face compounding challenges to ensure a sustainable balance between investment in new asset services and the need to maintain existing delivery of services at an optimal life cycle cost and quality while meeting community expectations. These challenges have placed significant pressures on organizations to improve the effectiveness of managing their infrastructure inventory through adopting more efficient, sustainable, and proactive engineering asset management (EAM) strategies. These challenges are due to the fact that these organizations have to deal with the growing concerns about resource scarcity, degrading environment, climate change, more stringent regulations, and a greater reliance on a multi-agency delivery model [2–4]. As a result, asset-intensive organizations are looking to improve their performance capabilities.

Assessing the maturity of an asset-intensive organization through the adoption of a capability maturity model is a practice that not only most effectively manages the resources; it supports the continual improvement in asset management performance. The capability maturity model is defined as an approach to assess the stages of development of business processes in organizations and a framework to improve processes through structuring a predefined set of levels. According to Hilson [5], the purpose of using a maturity model is to assess the current capability, strengths and weakness, and analyze gaps for improvement where it is required. In essence, these models are collections of best practices that help organizations to improve effectiveness, efficiency, and quality. A maturity model can thus be viewed as a set of structured levels that describe how well different processes of an organization are able to achieve staged outcomes in a reliable and sustainable way.

"EAM can be defined as the process of organizing, planning and controlling the acquisition, care, refurbishment, and disposal of infrastructure and engineering assets. It is a systematic, structured process covering the whole life of physical assets (p. 2)" [2]. Amadi-Echendu et al. [1] highlight five dimensions of EAM: spatial (consider all types of physical asset including interaction between the asset and stakeholders and clients, sustainability, industrial sector, and the government), temporal (consider short-term aspect such as operational management and long-term aspect such as strategic management of engineering assets), evaluation (consider financial measurement and capability measurement including social and physical capabilities), statistical (embedded in analysis process of risk), and organizational (including overall organizational management, the technology and information management, and the human factors management) [1, 6]. To assess the performance of an asset service organization's capabilities, this set of five dimensions is considered to be both comprehensive and crucial. There are growing numbers of maturity models being developed to assist with the assessment of how mature an organization is. However, the question is whether the existing capability maturity models for asset management met the five elements standard and consider all the dimensions in their processes. To answer this question, this paper undertakes a comparative study of existing asset management maturity models against the five dimensions of EAM to identify the gaps in terms of key process areas and to suggest incorporating all the key process areas for a more integrated approach to assess maturity of asset-intensive organizations. It is envisaged that the results of this analysis will better inform those practitioners in industry and academic researchers concerned with process improvement, intervention, and change management in organizations.

The paper proceeds as follows: Sect. 39.2 briefly describes the typical structure of a capability maturity model. This examination is followed by the methodology, which highlights the key process areas of EAM. Then, in Sect. 39.4, we compare some of the existing asset management maturity models against the different key process areas of EAM and identify the gaps for improvement. The paper concludes with a proposal for an integrated capability maturity model for asset management.

39.2 Capability Maturity Models

Capability maturity models (CMM) were first developed to objectively assess contractors' ability in undertaking software development projects [7]. Since then, capability maturity models have been widely used across a broad array of areas. In most cases, maturity is used in capability maturity models in the very technical sense to mean "the extent to which an organization has explicitly and consistently deployed processes that are documented, managed, measured, controlled, and continually improved" [8]. According to Paulk et al. [7], a capability maturity model is built upon five components: predetermined maturity levels, key process areas, goals, common features, and key practices. In this paper, we compare different asset management maturity models against the key process areas under five dimensions of asset management. As the key process areas provide the useful information for understanding different dimensions of EAM, we will use these to compare different capability maturity models. A key process area can be defined as the group of related activities that, when performed together, achieve a set of goals.

39.3 Methodology

39.3.1 Selection of the Asset Management Maturity Models

There have been a number of capability maturity model developed in the area of engineering asset management. From the list of potential models, we selected maturity models for comparative analysis that fulfilled the two criteria: relativity (consider the capability maturity models that are developed fully or partially in relation to the EAM) and publicly available (many maturity models are proprietary tools generated by consulting organizations. We consider only those maturity models that are in the public domain and available without cost).

In this paper, six maturity models are considered for a comprehensive comparative study. They are categorized as in the EAM field such as the PAS 55-BSI (publicly available specification 55-BSI), AMMM-OARSIK (Asset Management Maturity Model-OARSIK), PAMMM-OGC (Property Asset Management Maturity Matrix-OGC), AMM-IBM (Asset Management Maturity-IBM), AMM-SKF (Asset Maturity Management-SKF), and PAMCAM-OGC (Property Asset Management Capability Model-OGC).

39.3.2 Dimensions and Key Process Areas

The increasing complexity and sophistication of EAM processes has resulted in the creation of diverse areas of knowledge, expertise, and responsibilities within and across organizations. As a result, a state of process fragmentation has been created, and much inefficiency has subsequently arisen primarily due to the disintegration of process areas. An integrated approach to asset management can potentially eliminate many of the fragmentation inefficiencies by enabling the integration of processes. These dimensions and process areas of asset management, which are derived from the systematic review of extensive literature on asset management, have been concisely summarized in Table 39.1.

39.3.3 Evaluation of the Maturity Models

Content analysis of the related documents of the maturity models was carried out to identify the key process areas of the selected maturity models. According to Julien [34], a content analytic approach recognizes that text is open to subjective interpretation and reflects multiple meanings. For reliable evaluation of process of maturity models, in this paper, three authors carried out the contextual analysis individually and provided their ratings on the availability of the key process areas against the comprehensive list of key process areas. Later, these ratings were cross-checked and any discrepancies found in the evaluation were solved by consensus of the majority.

39.4 Findings

It is evident from Table 39.2 that all selected models incorporate process areas related to data management, asset register, and information systems. However, while all models focus on the asset data and knowledge standards related to data

Table 39.1 Dimensions and process areas of asset management

Dimensions	Elements	Process areas
Spatial	Community needs and expectations	Stakeholder management [9]; demand management [10]
	Environmental factors	Sustainability management [11]; climate change [12]
	Organizational governance	Interagency collaboration [13]
	Whole-of-government policy framework	Whole-of-government policy and whole-of-government model [14]
Temporal	Service delivery planning	Asset management policy [15]; asset management objectives [16]; asset management strategy [17]; acquisition plan [18]; operations plan [18]; maintenance plan [16]; disposal plan [18]
	Service delivery	Performance and condition monitoring [19]; incident management [19]; corrective and preventive actions [20]; procurement [18]
Organizational	Organizational governance	Corporate governance [21]; corporate policy [22]; corporate strategy [15]
	Knowledge management	Data management [20]; asset register [23]; information systems [24]; knowledge management [25]
	Organizational management	Leadership [26]; change management [27]; competence management [28]; organizational culture [29]
Statistical	Environmental factors	Risk management [30];
Evaluation	Evaluation	Asset performance measurement [31]; management reporting [32]; review [2]; audit [33]

quality and standards for collecting, categorizing, and providing asset information, but they do not deal with the more human-driven knowledge management aspects such as education and communication, trust-building, and team-enabling activities. Moreover, demand management under the spatial dimension, which is part of the forecasting long-term and short-term service demand, is not addressed in any of the models. The key process areas under the evaluation dimension are widely covered in almost all the maturity models described here with exceptions of AMM-IBM and AMM-SKF. The authors found that all the maturity models have the strong asset performance measurement process areas. However, apart from the risk management, the process areas related to the environmental factors in the form of sustainability management and climate change are overlooked in all the maturity models that are considered in this paper.

Compared to the other capability maturity models, the PAS55-BSI is the most comprehensive as it considers all the key process areas addressed under the temporal (in the form of service delivery planning and service delivery) and evaluation dimensions. Moreover, the model incorporates process areas related to the whole-of-life cycle asset management plans (acquisition, operations, maintenance and disposal plans). However, key process areas related to organizational governance (corporate governance, corporate policy and corporate strategy, and interagency collaboration) are not considered in this model. As part of the

Table 39.2 Comparison of six asset management maturity models

Elements (dimension)	Key process areas	PAS 55-BSI	AMMM-OARISK	PAMMM-OGC	AMM-IBM	AMM-SKF	PAMCAM-OGC
Service delivery planning (temporal)	Asset management policy	×	×	×			
	Asset management objectives	×	×	×			
	Asset management strategy	×	×	×	×	×	×
	Acquisition plan	×	×	×			×
	Operations plan	×	×	×	×	×	×
	Maintenance plan	×	×	×	×	×	×
	Disposal plan	×		×			
Service delivery (temporal)	Performance and condition monitoring	×	×	×	×	×	×
	Incident management	×				×	
	Corrective and preventive actions	×		×	×	×	×
	Procurement	×	×	×			×
Organizational governance (organizational)	Corporate governance			×			×
	Corporate policy			×			×
	Corporate strategy			×			×
	Interagency collaboration						
Knowledge management (organizational)	Data management	×	×	×	×	×	×
	Asset register	×	×	×	×	×	×
	Information systems	×	×	×	×	×	×
	Knowledge management						
Organizational management (organizational)	Leadership		×			×	
	Change management	×				×	
	Competence management	×	×		×	×	×
	Asset management culture					×	

(continued)

Table 39.2 (continued)

Elements (dimension)	Key process areas	PAS 55-BSI	AMMM-OARISK	PAMMM-OGC	AMM-IBM	AMM-SKF	PAMCAM-OGC
Environmental factors (statistical/spatial)	Risk management	×	×	×	×	×	
	Sustainability management						
	Climate change						
Community needs and expectations (spatial)	Stakeholder management	×					×
	Demand management						
Whole-of-government policy framework (spatial)	Whole-of-government policy	×				×	×
	Whole-of-government model	×				×	×
Evaluation (evaluation)	Asset performance measurement	×	×	×	×	×	×
	Management reporting	×	×	×		×	×
	Review	×	×	×	×		×
	Audit	×	×	×			×

organizational management process area, this model considers the change management and competence management but not the leadership and organizational culture, which means the model is lacking in two important process areas related to human aspects.

The next comprehensive maturity models are PAMMM-OGC and PAMCAM-OGC. In contrast to PAS55-BSI, these two models start from the organizational strategic governance (corporate governance, corporate policy, and corporate strategy) to asset management strategy. Similar to the PAS55-BSI, these two models have the whole-of-life cycle asset management plans. The process areas related to the organizational management and community needs and expectations are not considered in the PAMMM-OGC. As part of the process areas related to the organizational management, only competence management process area and stakeholder management under the community needs and expectations are taken into consideration in PAMCAM-OGC. It is clear that process areas related to human and social aspect are not fully covered in these two models.

The remaining maturity models (AMM-OARSIK, AMM-IBM, and AMM-SKF) provide mixed findings in relation to their suitability in attaining integrated process areas. None of these three models have process areas related to organizational governance and nor do they consider whole-of-life cycle management plans in their process areas. In terms of process areas related to organizational management, AMM-SKF is more comprehensive than other maturity models in the list as the model covers all the process areas in the form of leadership, change management, competence management, and asset management culture.

It is apparent from the comparative study that none of the maturity models have fully incorporated all the process areas against the five key dimensions. The desire to implement efficient and optimized asset management service delivery has created a strong demand for "bridging the gaps" through the adoption of integrated approaches. There is a need for a comprehensive asset management capability maturity model (AMCaMM) that considers all of the five dimensions of asset management mentioned earlier. To secure optimal benefit from a capability maturity model, it is obvious that the model that covers comprehensive process areas of asset management is more effective.

39.5 Summary and Conclusion

In summary, the strategic-level issues apart from the PAMMM-OGC and PAMCAM-OGC start only at the asset management level in the form of asset management policy, objective, and strategy but not at the strategic asset governance in the form of corporate governance and related corporate policies, objectives, and strategies. Furthermore, life cycle asset management plans are only considered in two maturity models, namely PAS55-BSI and PAMMM-OGC. Moreover, there are gaps associated with the process areas related to human and social issues in all the maturity models.

The integration of the process areas can be useful for achieving mature asset service organizations. The implications are that mature organizations are able to: manage all the projects undertaken by an organization effectively [35]; improve continually the performance of all projects undertaken by an organization and improve dialogue between the project management community and an organization's top management [36]. Therefore, the authors argue for an integrated AMCaMM based on a combination of all the process areas using all five dimensions of EAM. The paper concludes that a well-designed capability model for managing assets should include broader contextual elements and address higher-level organizational management levels by integrating corporate planning processes with their asset planning processes. Further research is required to establish the different graduations of maturity along a scale and how the different bundles of processes within each level of the scale should be assessed for maturity. The central unifying theme of the AMCaMM is the development of managerial and strategic solutions to the social and human issues along with the technical issues that potentially enhance an organization's ability to manage their engineering assets.

Acknowledgments This paper was developed within the CRC for infrastructure and engineering asset management, established and supported under the Australian Government's Cooperative Research Centres Programme. This paper is produced in collaboration with the Australian Asset Management Collaborative Group (AAMCoG).

References

1. Amadi-Echendu JE, Willett RJ, Brown K, Hope T, Lee J, Mathew J, Vyas N, Yang B-S (2010) What is engineering asset management? In: Amadi-Echendu JE, Brown K, Willett R, Mathew J (eds) Definitions, concepts and scope of engineering asset management. Springer, London, pp 3–16
2. AAMCoG (2011) Guide to integrated strategic asset management. Australian Asset Management Collaborative Group, Brisbane
3. Queensland Department of Environment and Resource Management (2010) Guidelines for the review and regular audit of strategic asset management plans. Queensland Department of Environment and Resource Management, Brisbane
4. Woodhouse J (2010) Asset management: the way forward. In: Lloyd C (ed) Asset management—whole life management of physical assets. Thomas Telford Press, London, pp 201–221
5. Hilson D (2003) Assessing organisational project management capability. J Facil Manage 2(3):298–311
6. Laue M, Brown K, Scherrer P, Keast R (2012) Integrated strategic asset management: frameworks and dimensions. In: Third international engineering systems symposium CESUN 2012. Delft University of Technology
7. Paulk MC, Curtis B, Chrissis MB, Weber CV (1993) Capability maturity model for software, version 1.1 (CMU/SEI-93-TR-24, ADA263403). Carnegie Mellon Software Engineering Institute, Pittsburgh
8. CMMI Product Team (2002) Capability maturity model integration (CMMI), Version 1.1. Carnegie Mellon Software Engineering Institute, Pittsburgh

9. Beach S, Brown K, Keast R (2010) How government agencies engage with stakeholders. Public Adm Today, 34–41
10. New South Wales Treasury (2004) Demand management guideline. New South Wales Treasury, Sydney
11. New South Wales Treasury (2004) Sustainable development guideline. New South Wales Treasury, Sydney
12. Rayner RF (2010) Incorporating climate change within asset management. In: Lloyd C (ed) Asset management—whole life management of physical assets. Thomas Telford Press, London, pp 161–180
13. Keast R, Mandell M, Brown K. Woolcock G (2004) Network structures: working differently and changing expectations. Public Adm Rev 64(3):353–361
14. New South Wales Government Asset Management Committee (2003) Total asset management manual. Available at http://www.gamc.nsw.gov.au/tam/default.asp?PageID=73
15. Viljoen J, Dann S (2000) Strategic management: planning and implementing successful corporate strategies. Pearson Education, Australia
16. Tasmania Department of Treasury and Finance (2004) Developing strategic asset management plans. Tasmania Department of Treasury and Finance, Hobart
17. New South Wales Treasury (2006) Asset strategic planning. New South Wales Treasury, Sydney
18. Australian National Audit Office (2010) Better practice guide on the strategic and operational management of assets by public sector entities—Delivering agreed outcomes through an efficient and optimal asset base. ACT
19. Western Australia Department of Local Government (2012) Asset management—managing service delivery. Western Australia Department of Local Government, Perth
20. The Institute of Asset Management (2011) Asset management—an anatomy. The Institute of Asset Management, Bristol
21. Cornish N, Morton K (2001) Asset governance—A radically new way to manage distribution networks in a competitive and deregulated market. In: 16th international conference and exhibition on electricity distribution (CIRED 2001). Amsterdam, Netherlands
22. Thompson AA, Strickland AJ (1999) Strategic management: concepts and cases. McGraw-Hill, UK
23. New South Wales Treasury (2004) Asset information guideline. New South Wales Treasury, Sydney
24. Cato WW, Mobley RK (2002) Computer-Managed maintenance systems. In: Baskarada S (ed) Information quality management capability maturity model. Vieweg + Teubner Research, Elsevier
25. James P (2005) Knowledge asset management—the strategic management and knowledge management nexus, DBA thesis, Southern Cross University, Lismore, NSW
26. House RJ, Hanges PJ, Javidan M, Dorfman PW, Gupta V (2004) Culture, leadership, and organizations: the GLOBE study of 62 societies. Sage Publications, Thousand Oaks
27. Osborne SP, Brown K (2005) Managing change and innovation in public service organizations. Routledge
28. Hyland P, Murphy G, Salomone S (2009) Final report—asset management skills and competencies, Queensland University of Technology
29. Brunetto Y, Xerri M (2011) Ensuring the implementation of engineering asset management: understanding organizational culture, WCAEM 2011. Cincinnati, USA
30. New South Wales Treasury (2004) Risk management guideline. New South Wales Treasury, Sydney
31. AAMCoG (2008) Public sector asset performance measurement and reporting. Australian Asset Management Collaborative Group (AAMCoG) and CRC for integrated engineering asset management (CIEAM)
32. New South Wales Treasury (2004) Post implementation review guideline. New South Wales Treasury, Sydney

33. Western Australia Department of Treasury and Finance (2010) Strategic asset management framework—High-level policy. Western Australia Department of Treasury and Finance, Perth
34. Julien H (2008) Content analysis. In: Given LM (ed) The Sage Encyclopedia of qual research met. SAGE Publications, London, pp 120–121
35. Suares I (1998) A real world look at achieving project management maturity. In: Proceedings of the 29th annual project management institute 1998 seminars and symposium, Long Beach, CA, (CD-ROM). Project Management Institute, Newtown Square
36. Peterson AS (2000) The impact of PM maturity on integrated PM processes. In: Proceedings of the 31st annual project management institute 2000 seminars and symposium, Houston, TX, (CD-ROM). Project Management Institute, Newtown Square

Chapter 40
Stagnancy in Indonesia's Reformed State Asset Management Policies and Practices: A Wicked Problem?

Diaswati Mardiasmo, Charles Sampford and Paul Barnes

Abstract The policies and regulations governing the practice of state asset management have emerged as an urgent question among many countries worldwide for there is heightened awareness of the complex and crucial role that state assets play in public service provision. Indonesia is an example of such country, introducing a 'big bang' reform in state asset management laws, policies, regulations, and technical guidelines. Indonesia exemplified its enthusiasm in reforming state asset management policies and practices through the establishment of the Directorate General of State Assets in 2006. The Directorate General of State Assets have stressed the new direction that it is taking state asset management laws and policies through the introduction of Republic of Indonesia Law Number 38 Year 2008, which is an amended regulation overruling Republic of Indonesia Law Number 6 Year 2006 on Central/Regional Government State Asset Management. Law number 38/2008 aims to further exemplify good governance principles and puts forward 'the highest and best use of assets' principle in state asset management. The purpose of this study is to explore and analyze specific contributing influences to state asset management practices, answering the question why innovative state asset management policy implementation is stagnant. The methodology of this study is that of qualitative case study approach, utilizing empirical data sample of four Indonesian regional governments. Through a thematic analytical approach, this study provides an in-depth analysis of each influencing factors to state asset management reform. Such analysis suggests the potential of an 'excuse rhetoric'; whereby the influencing factors identified are a smoke-screen, or are myths that public policy makers and implementers believe in, as a means to explain stagnant implementation of innovative state asset management practice. Thus, this study offers deeper insights into the intricate Web that influences state asset management innovative policies to state asset management policy makers; to be taken into consideration in future policy writing.

D. Mardiasmo (✉) · C. Sampford · P. Barnes
Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: d.mardiasmo@qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_40,

40.1 Introduction

Indonesia exemplified its enthusiasm in reforming state asset management policies and practices through the establishment of the Directorate General of State Assets in 2006. The Directorate General of State Assets have stressed the new direction that it is taking state asset management laws and policies through the introduction of Republic of Indonesia Law Number 38 Year 2008, which is an amended regulation overruling Republic of Indonesia Law Number 6 Year 2006 on Central/Regional Government State Asset Management [1]. Law number 38/2008 aims to further exemplify good governance principles and puts forward 'the highest and best use of assets' principle in state asset management [1].

Indonesia is chosen as a country case study for two reasons:

Firstly the re-introduction of good governance principles after the Asian Financial Crises in 1997 is an ongoing process in Indonesia and is a main objective of the current presidency (Susilo Bambang Yudhoyono) regime. This has resulted in a push for policy innovation, whereby conceptualizing good governance principles within all areas of government responsibilities (including state asset management) is emphasized. The Indonesian government has identified slow implementation of innovative policies and have acknowledged the 'lingering effect' of Soeharto's governing regime within the bureaucracy's 'mind-set'; however, yet to identify tangible impeding influences and/or construct relationships on how these impeding influences play a role in policy innovation implementation.

Secondly a review of state asset management practices (of various countries) and the literature on an integrated governance and asset management approach show that although Indonesia is acknowledged to have interesting complexities within its application of reformed state asset management practices [2–4], there is an absence of studies on Indonesia's state asset management practices—both prior to and after introduction of state asset management reform in 2006. The dearth of research in Indonesia's state asset management approach suggest the contribution of this study—that of theoretical and practical.

The purpose of this research is to create an understanding of the influences that play a role in Indonesia's conceptualization and implementation of reformed state asset management policies, ultimately providing an explanation of why there is stagnancy in said reform. To achieve this purpose, three research objectives are identified.

1. Identify and define influences that play a role in Indonesia's conceptualization and implementation of reformed state asset management policies
2. Evaluate and determine the validity of identified influences through statistical and qualitative methods.

40.2 State Asset Management in Indonesia: Pre- and Post-Reform

Pre-Reform. Up until 2003–2004, state asset management is minimum, in the sense that inventory of assets is done manually, there is incomplete public sector accounting standards, and incomplete financial reporting standards. It is acknowledged that the years of 1970s–1990s is known as the year of development, where there was euphoria of building and developing state assets and public infrastructure to support the role of government or government programs of the day. In the late 1990s, the central government started to acknowledge the importance of accountability and financial reporting, as well as compliance reporting. Hence, the central government introduced a simplistic inventory system for state asset management—manually executed and based on single entry bookkeeping.

The simplistic state asset management practice was evident during the Soeharto period, for approximately 32 years of his regime. However, during that time transparency, accountability, and maintenance state assets was not the main focus, be it by the government or the society itself. The state government that focused on building and purchasing, however, did not have the framework to manage acquired state assets. For example, during Soeharto's regime, there was no requirement for asset reporting or inventory keeping of state assets. This knowledge leads to several conclusions:

1. Confirms Wardhana's [5] statement that there has been an absence of 'caring' culture within the Indonesian government in terms of state asset management. It also details that an information system dedicated to state asset inventory and reporting have thus far been incomplete.
2. To a certain extent the absence of state asset maintenance is not 'only' the government's responsibility. The public/society plays a role, in terms of being swept up in the creating and building euphoria, taking for granted state assets that were built and utilizing it without consequence, and not insisting on financial and accountability reports from the government in regard to the assets [5].

Post-Reform. Hadiyanto [1] documented reform in state asset management dating back to 2004, with *Law No* 1 *Year* 2004 *on State Treasury* as a reform locomotive. The establishment of Law 1/2004 was deemed to be the highest decree, managing the function of state asset management as part of the treasury. At the time, state asset management was defined as budgeting, financing, and controlling [1]. It is observed that at the beginning of state asset management reform the definition and/or list of state asset management function is very different from the definition and list of asset management functions as detailed by Komonen et al. [6], Cornish and Morton [7], Kaganova [2], and others.

State asset management reform continued with the introduction of *Law No 6 Year 2006 on Central/Regional State Asset Management*, which along with the establishment of the Directorate General of State Assets, is deemed to be the establishment of state asset management reform in Indonesia [1]. The introduction of *Law No* 6 *Year* 2006 *on Central/Regional State Asset Management* exhibited a change of paradigm from the definition and perspective of state asset management as per *Law No* 1 *Year* 2004 *on State Treasury*. Changes include as follows:

1. Widening the scope of state asset management to budgeting, acquisition, utilization, maintenance and monitoring, valuation, disposal, change/hand over, administration/inventory, control mechanisms, and capacity building function.
2. Introduction of an asset manager role, whereupon the Directorate General of State Assets is appointed as the central government's state asset manager, in a bid to ensure professionalism in state asset management practices.
3. Integration of managerial and reporting aspects in state asset management in financial reporting as part of accountability.

State asset management reform in Indonesia is further marked by the establishment of The Directorate General of State Assets (DJKN) on December 7, 2006. The establishment of a department that specializes in state asset management shows a realization of the importance of state assets and is a positive start to state asset management reform; however, this also indicates a heavy task that involves addressing a myriad of complexities and issues. The Directorate General of State Asset introduced a **Road map to Strategic Asset Management**, which is designed to be viewed as the compass in an attempt to establish a state asset management practice that is sound and modern. The idea of strategic asset management is supported by Marlow and Burn [8] as they found that setting such a direction has assisted in establishing a more efficient and effective water waste management system in many countries such as the USA, UK, Australia, and New Zealand. Hadiyanto [1] provided the *definition of strategic state asset management* as the integration of functions such as planning, budgeting, maintenance, and accountability of state asset management that puts forward the principle of 'the highest and best use of assets.' This definition is within the corridor of asset management definitions, such as the definition offered by Komonen et al. [6], Lin et al. [9], and Jabiri et al. [10]. The concept of ensuring accountability and the highest and best use of asset is also in line with Cornish and Morton [7]; Woodhouse [11] and Marlow and Burn's [8] description of the integration of governance principles with asset management principles.

40.3 Stagnancy in Reform

Mardiasmo [12] warns that the implementation of a 'reform,' no matter in what area, will take time to eventuate, as Indonesia itself is still going through a thorough economic, political, and societal reform—thus any type of reform will be impacted by such dynamics. Hadiyanto [1] agrees with this as the possibility of revising state asset management practices due to internal and external factors is still very high. Ritonga [13] further explains that standardization will have to be done in a certain stages, which is considered to be a norm in any reform that is implemented in Indonesia. This is true based on the findings of Mardiasmo [14] within the context of good governance implementation in Indonesian regional governments, where there was a need to roll out the implementation plan in various stages during a long-term time frame.

Hadiyanto [1] has identified a stagnancy in state asset management reform, in particular in the implementation of reformed policies and practices; whereby regional government in Indonesia are at various levels of implementation. This statement is supported by Pardiman [15], identifying a time lag in regional government implementation (of the reform) which has impacted the Directorate General's time line to achieve strategic asset management. The Directorate General has put forward potential reasons for such time lag; however, empirical research is yet to be performed.

40.4 Methodology of Research

In order to achieve the purpose of this study, a qualitative orientation is taken. Qualitative research is a descriptive and interpretive method of gathering and analyzing data, which is particularly useful if the purpose of the research is to gain further in-depth understanding of a particular phenomenon or to explore relationships between identified influences [16–22]. Case studies are chosen in this methodology design as it provides a method for investigating complex social phenomena, such as decision making, and can holistically encompass many aspects and characteristics of real-life events [23].

To achieve the objectives of this study, it is essential to choose an organization (i.e., a case study) within Indonesia that has the responsibility of managing state assets. A main variable considered in the implementation of reformed state asset management is decentralization and regional autonomy regime, as it has the potential to add complexity in ensuring equal understanding and implementation of public policy reform [24]. Decentralization policy in Indonesia indicates regional independence in terms of policy making and economic autonomy [25–27]. Each region has the independence to enact policies based on the level of benefits reaped, where central government acts as an advisory as well as a control mechanism [28–31]. Geographical distance is found to dictate the level of central–regional government interaction, quantity and quality of face-to-face consultation, and the number of regional government officials attending workshops/training at central government offices [14]. Thus, *geographical distance* is considered a variable in the adaptation and implementation of policies by regional government.

Regional governments established after the introduction of decentralization and regional autonomy (in 2001) tend to embrace innovative ways of implementing government procedures, due to dissatisfaction to previous practice (i.e., during Soeharto's regime) in governing [32–34]. Mishra [34] concluded that newer

Fig. 40.1 The Johari window as a sampling method

	Near to Capital	Far from Capital
Old	Case study 1 (DKI Jakarta Region)	Case Study 2 (DIY Yogyakarta Region)
New	Case Study 3 (Directorate General of State Assets)	Case Study 4 (Gorontalo Region)

Table 40.1 Number of interviews

Regional government	High-level government official	Middle-level government official	Low-level government official
DIY Yogyakarta	5	9	8
DKI Jakarta	3	4	4
Gorontalo	7	9	11
Directorate General of State Assets	2	4	4

regional governments would be more open to implementing good governance principles in their public policy drafting and implementation as this is perceived as an innovative way of governing. Therefore, the 'age' of a regional government has arisen as a potential variable. The 'age' of a regional government refers to when the regional government is established, divided into two main 'age groups': prior to the introduction of decentralization and regional autonomy in 2001 (hence identified as an *'old' regional government*) and post-decentralization and regional autonomy (hence identified as a *'new' regional government*).

The identification of *geographical distance* and *age of regional government* captures the unique relationship between central and regional government of Indonesia. These two influences are utilized to contextualize the research issue, where a two-dimensional four cell matrix is used. The Johari Window compares two influences on each axis, in which each category is reflected within its quadrants [35–37]. The Johari Window is illustrated in Fig. 40.1.

The diesel engine test rig as shown in Fig. 40.1 was employed in a range of engine fault simulations in the experimental study. The engine is an in-line four stroke, four cylinder diesel engine, which is commonly used for power generation in remote construction sites or an emergency power generator in hospitals. The engine generates about 15 kW of nominal power output at full-load condition. During the experiment, the power outlet of the diesel engine generator set was connected to a three-phase 15 kW industrial fan heater. The fan heater has three heat settings at 5, 10, and 15 kW, which can be adjusted during the experiment for different engine loadings. The engine specification is shown in Table 40.1.

Figure 40.1 identifies the four case studies sampled for this research. It is acknowledged that the Directorate General of State Assets is not a province/ regional government; however, its inclusion is justified as it is a government department that acts as a state asset manager for the central government, whereby

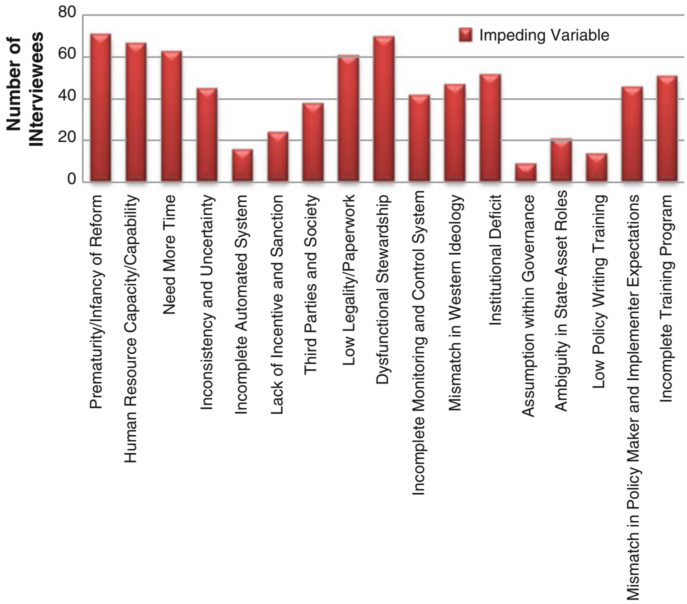


Fig. 40.2 Support for impeding influences as identified by interviewees

the Indonesian government consider the management and governing of central government as equal to a province and as an individual entity, due to its size and complexity [1, 38].

The number of interviews involved within each case study is outlined in Table 40.1, categorized by echelon level.

As evident in Table 40.1, there is a slight discrepancy in the number of government officials interviewed; due to time restrictions, availability of government officials, and willingness of government officials to participate. It is also noted that the number of participants from DKI Jakarta and Directorate General of State Assets may seem low (in comparison to DIY Yogyakarta and Gorontalo). Lower numbers of interview participants can be explained by the centralized government structure and state asset management practice of both case studies, in contrast to the decentralized government structure and state asset management practice of DYI Yogyakarta and Gorontalo.

40.5 Why Stagnant? A Wicked Problem or a Rhetoric Excuse?

Throughout the data collection process of this study, interviewees identified numerous influences that play a role in the conceptualization and implementation of state asset management reform in Indonesia. Interviewees had a tendency of viewing these influences as ‘impeding influences,’ in the sense that these influences contribute to the challenges of achieving a state asset management reform that fully incorporates good governance principles and is equally adopted/implemented by all regional governments within Indonesia. These ‘impeding influences’ are presented in Fig. 40.2, along with the number of interviewees who has identified it as a contributing challenge to state asset management reform.

40.5.1 Premature/Infancy of Reform

Indonesia’s state asset management reform was introduced in 2006 through Law 6/2006, with regional state asset management conceptualized and ‘enforced’ in 2007 through Law 17/2007. The reform was revised in 2008, with the introduction of Law 38/3008. As data collection/interviews were performed in July 2010, interviewees reasoned that the state asset management reform is ‘new’ or ‘premature’ for it was introduced a mere two years (at most, from its latest revision in 2008) prior to the study. Interviewees have associated the new state asset management practice with high uncertainty, for due to its perceived radical and ‘big bang’ reform nature, there is much confusion surrounding answers to questions such as what is state asset management, who are the state asset managers, what is best practice within it, how to implement it on a day-to-day basis, etc. Thus, interviewees believe that the reform is still premature, due to the myriad of questions unanswered, and that there is a need for policy makers to continuously revise laws and policies of the reform. This is evident in Table 40.2.

40.5.2 Human Resource Capacity and Capability

The ‘people,’ or state asset management-related actors, are identified as an influencing factors to state asset management reform by 89 % of interviewees. Within this influencing factors (human resource), characteristics such as commitment, level of knowledge, willingness, personality, and capability/competence, and sense of stewardship are included. In fact, dysfunctional sense of stewardship as an influencing factors is supported by 93 % of interviewees and incomplete training in state asset management laws, policies, and implementation as an influencing factors are supported by 68 % of interviewees. As dysfunctional sense of stewardship and incomplete training in state asset management matters are considered by interviewees to be part of human resources (state asset management-related actors) characteristics, these figures adds to the validity of the ‘people’ as an influencing factors to state asset management reform.

Interviewees of this study provided several reasoning as to why ‘the people’ is a valid influencing factors. One of them is the disparity in state asset management knowledge between government officials, where it is argued that the disparity in

Table 40.2 Prematurity of state asset management reform as a factor

'…that the rules and regulations that govern the management of state assets can still be classified as something that is relatively new. This includes everything that governs state asset management—the constitution, the umbrella laws, the regional level regulations, and even the technical guidelines. All of this is relatively new and can still be classified as "pre-mature", where there are many changes and updates to laws, rules and regulations, and technical guidelines…'
'…asset management is still a new topic in Indonesia and to be honest there is a need for higher level of knowledge specifically on the matter (instead of it being done by accountant and auditors). It can also be a hindrance, because people may say 'we don't have the necessary laws and guidelines yet' and use it as a potential 'excuse' or reason as to why a particular policy is not being implemented…'

Table 40.3 Support for the 'People' as an influencing factors

'…here identifies the kind of personality or capability that is needed in government official personnel—one that is committed, competent, and performs the activities in a conducive way. It is the combination of these three main personalities that I find to be difficult—to find it in one person, or each person within a team. Hence perhaps this is what is hindering the whole thing, the inability, or the rareness in finding these three personalities in one person—or finding all three at the highest level is probably even a better description!
'…The key reason here is human resources, or perhaps not human resources, but the human nature of our government officials. I think from the supporting infrastructure etc. we already have everything that we need, however when it comes to the humans/government officials who are supposed to implement the policies, they start to look right and left to see where can they make extra money…'

knowledge have resulted in a discrepancy in expectation and outcome. Interestingly, the mismatch between policy maker and implementer gained 61 % of interviewee support (Fig. 40.2), which suggests that there is agreement surrounding the mismatch as a product of disparity in state asset management knowledge. This is evident in Table 40.3.

40.5.3 The Notion of 'Needing More Time'

Time is considered to be an impediment to state asset management reform by government officials, with 83 % support. The concept of 'I need more time' is divided into three senses:

1. The time that has passed between the procurement and/or acquirement of a new state asset and the introduction of state asset management reform,
2. The time that has passed between the establishment of a region (both that is considered to be new after 2001 and old prior to the decentralization regime) and the introduction of state asset management reform,

Table 40.4 Opposing statement to 'I need more time' as an influencing factors

'…its not that we need more time to implement this state asset management reform, its more about whether we want to implement the new laws and policies and way of doing things. I mean yes it does take some time for it to be fully understood and implemented, but not decades, saying that I need more time is just an excuse because that person doesn't want to do it, or the new way of doing things does not support their own personal views…'
'…thats just an excuse. We already have an automated system (SIMDA—red) from the internal audit body, all we have to do is follow that and implement it, is not like we don't have to make everything by ourselves. The internal audit body have already provided us with the system, all we need to do is implement. Why do we need more time?…'

3. That ultimately there is need for 'more time' to ensure uniform understanding and implementation of state asset management.

It is argued by government officials that the time that has passed between the procurement and/or acquirement of new state asset and the introduction of state asset management reform is an impediment as the 'older' the state asset is (in age) the more likely that it is subjected to the 'old' practices of state asset management. Interviewees suggest that the longer the time lapsed the higher the chance that the condition of a state asset is unknown, its utilization rate and by whom is unknown, certificate ownership lost or not yet completed, and maintenance of state asset to be behind schedule.

The time that has passed between the establishment of a region and the introduction of state asset management is considered to be a variable in the implementation of state asset management reform, as it is perceived that newer (i.e., post 2001) regional governments (such as Gorontalo) would have less state assets (in terms of amount and history) and thus have a more manageable state asset size than regional governments established prior to decentralization regime (such as DKI Jakarta and DIY Yogyakarta). This is perceived to be an advantage in terms of recording, inventory, recalling information, tracking down utilization, and creating ownership certificates.

The third sense of 'time' as an influencing factors to state asset management implementation is, to a certain extent, related to the prematurity or newness of state asset management reform—in the sense that not enough of time has passed since the introduction of the reform and the present study. Interviewees believe that 'more time' and a 'step-by-step' approach are needed for activities such as drafting specific laws, increasing knowledge and understanding, complete socialization, and implementation workshops. This view of 'I need more time' is supported by 84 % of interviewees. Interestingly, 16 % of interviewees contradict this view, claiming that 'I need more time' is 'an excuse' for non-implementation. These views are captured in Table 40.4.

40.5.4 Dysfunctional Stewardship

Low stewardship and/or dysfunctional perception of custodianship have been identified by interviewees as a potential influencing factors to state asset management. This is supported by 93 % of interviewees, which leads to the belief that it is indeed a valid influencing factors. Here, low level of stewardship or dysfunctional perception of custodianship refers to state asset management actors viewing the state assets in their care as a free good, not a good that needs to be managed in a best practice manner to ensure optimization of utilization or wealth creation. State asset management actors do not feel a sense of belonging toward the state assets in their care, viewing that it belongs to another party (though who is unclear), and thus, it is not within their interest to ensure best practices in managing the state asset. How this hinders state asset management reform is captured in Table 40.5.

Acknowledging the negative impact that low level of stewardship or dysfunctional custodianship may have on regional state asset management, few government officials attempted to provide a justification as to why there is dysfunctional custodianship of regional state assets and have put forward suggestions in order to increase stewardship. One of the reasons put forward is the fact that there is no 'ownership giving /transition' ritual or ceremony, where there is no formal process in which regional state assets are delegated or bequeathed to regional heads or government official—thus no symbolized transfer of responsibility so to speak. This has resulted in regional heads or government officials not feeling the responsibility of being 'trusted' with a state asset. It is even identified that at times, due to the lack of this symbolic process, regional heads are unaware of the new state assets that exists in his/her jurisdiction.

40.6 Discussion and Conclusion

The findings section has highlighted and discussed main influencing factors to state asset management reform that are identified by interviewees, which include as follows: newness and prematurity of state asset management concept, the perceived impact of 'time,' the people, and dysfunction custodianship/low stewardship. It is interesting to observe inconsistencies and contradictions between views surrounding each influencing factor, which suggests that there could be an underlying, more basic and simple, explanation to why state asset management reform is not yet implemented fully.

For example, in discussing dysfunctional custodianship or low stewardship, lower level echelons argue that this is evident in state asset management practices due to the challenges in changing a culture and 'way of doing things' as Indonesia has a history of neglect in relation to its state assets. Therefore, they believe that 'that is how things are' and that changing a culture would take decades as opposed

Table 40.5 Support for low stewardship /dysfunction custodianship as influencing factors

'…In regards to stewardship, I strongly agree that we are lacking this particular feeling. For example a government vehicle. If something happens to the vehicle while a government official is using it, then the government official should replace and ensure that the vehicle is no longer broken—this is of course other than filing paperwork for insurance to look after the damages. So people (government officials) can no longer just use state assets to their heart's content—perhaps the government vehicle is used by his nanny or a distant family cousin or his neighbour for example. People at the moment do not want to be responsible, the do not want to pay if some sort of damage has happened…'
'…I think the main problem here is that people take assets for granted. They think its a free good, comes out of the government's pocket/money and not theirs, and therefore they don't have a high level of responsibility toward it. They think if something is broken, oh well, just buy a new one using the government's budget. They don't think about preserving it for best use or that it is wealth creation, they just use it however they want. So thats a big problem, if people don't feel like they are responsible for assets then how are we going to get them to change their mind? And follow the new rules and regulations?'

to years, and would also need to be devised in a step-by-step manner to ensure full implementation. High-level echelon government officials, however, disagree to a certain extent. They agree that culture change is a difficult and challenging process; however, this is not the main reason behind low stewardship or a dysfunctional sense of custodianship, rather it is caused by the differing levels of commitment and willingness in taking responsibility of the state assets within one's jurisdiction. High-level echelon government officials argue that despite the potential challenge within change, if the policy implementers are willing and committed to take drastic action and admit responsibility, stewardship and custodianship will eventuate.

Another example is an inconsistency of opinion regarding human resource capabilities. Lower level echelon government officials provided the reasoning of low training, socialization program, and information dissemination; and the unexplained overlap in state asset management legal products. They believe that, thus far, there is not enough information regarding state asset management to appease their anxieties and uncertainties concerning the change in state asset management-related activities. It is also believed that there are too many state asset management legal products (both from central and regional government) and that there is confusion on which one to follow. High-level government officials, however, partially rejects this argument, as although they believe there is a need for continuous training and socialization of state asset management reform; there are scheduled training and socialization programs from central government on a bi-monthly basis (which is open to any regional government), regional governments provide their own seminars concerning state asset management implementation to members within the office, and there is an assistance team from the internal audit body to oversee/accompany the process of understanding and implementing state asset management reform.

Lastly, an inconsistency is found within the interviewee's opinion themselves throughout the interview, in particular in their indecisiveness of what is considered to be a main impediment to state asset management reform, and the cause of each influencing factors. It is found, more often than not, that interviewees were able to champion for a particular angle of argument depending upon the 'bait' that the interviewer provided during the interview process. It is also found that interviewees were able to provide supportive or disagreeing comments for a particular impediment variable, which shows that there is a lack of certainty among interviewees—which has the effect of unconvincing arguments regarding the reasons as to why state asset management reform is facing challenges.

The inconsistencies and contradictions found (as per above) throughout the interview process resulted in further questions of whether the impediment variables provided are, in fact, the reality of why state asset management reform is not yet implemented (or is facing challenges). It is questioned whether the impediment put forward are in fact the myths that government officials believe in, in order to justify the lack of implementation of state asset management reform policies. The notion of 'the myths that we believe in' was recently discussed by Bessant [39] in her work concerning the reasons for lack of implementation of public policy in both developed and developing countries.

Bessant [39] further elaborates that due to the myths that government officials live or believe in, more often than not government officials are caught up in the fantasies of policy making—referring to the ideal world of their policy objectives but not considering the reality in which they are living in. Hence, other government officials who are relegated the job of implementing these policies are not able to make a connection between the policy and the reality. Furthermore, Bessant [39] spoke of a multiplier effect of the myths that government officials believe in, whereby between the policy makers and the policy implementers there is an 'imaginary institution' who is going to translate policies into reality, provide needed support, and play the role of monitor and controlling of said policies. More often than not however such an institution is 'imaginary' as the existence of such an institution within a government is questioned—that is, the number of government institutions with these specific functions are questioned [39].

The contradictions of interviewee opinions regarding influencing factors to stagnant reformed state asset management, coupled with the ideas of Bessant [39], provokes the thought that perhaps influencing factors are part of the myths that government officials live/believe in, where they are in fact an 'excuse rhetoric' that is provided by government officials to justify the lack of reformed public policy implementation. Here, 'excuse rhetoric' is likened to the term 'rhetorical question'—whereby more often than not the answer to the question is implicitly found in the question asked. Hence, 'excuse rhetoric' refers to the real 'excuse' being implicitly found within the influencing factors provided.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established, and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Hadiyanto (2009) Strategic asset management: Kontribusi Pengelolaan Aset Negara dalam Mewujudkan APBN yang Efektif dan Optimal. In: Abimanyu A, Megantara A (eds) Era Baru Kebijakan Fiskal: Pemikiran, Konsep, dan Implementasi. PT Kompas Media, Jakarta
2. Kaganova O (2006) A need for guidance in countries with emerging markets. In: Kaganova O, McKellar J (eds) Managing government property assets: international experiences. The Urban Institute Press, Washington DC
3. MacAndrews C (1998) Improving the management of Indonesia's national parks: lessons from two case studies. Bull Indonesian Econ 34:121–137
4. MacAndrews C, Saunders L (1997) Conservation and national park financing in Indonesia, natural resources management project occasional paper no 6. Natural Resources Management Project, Jakarta
5. Wardhana A (2009) 30 Tahun Reformasi Ekonomi di Indonesia: Transisi dari Kebergantungan akan Sumber Daya menuju Kemampuan Bersaing Secara Internasional. In: Abimanyu A, Megantara A (eds) Era Baru Kebijakan Fiskal: Pemikiran, Konsep, dan Implementasi. Jakarta, PT Kompas Media
6. Komonen K, Kortelainen H, Räikkönen M (2006) An asset management framework to improve longer term returns on investments in the capital intensive industries. WCEAM, Queensland
7. Cornish N, Morton K (2001) Asset governance—A radical new way to manage distribution networks in a competitive and deregulated market. London Electricity's Dist Bus 482:18–21
8. Marlow DR, Burn S (2008) Effective use of condition assessment within asset management. J Am Water Works Assoc 100:54
9. Lin S, Gao J, Koronios A, Chanana V (2007) Developing a data quality framework for asset management in engineering organisations. Int J Inf Qual 1:100–126
10. Jabiri NZ, Jaafari A, Platfoot P, Gunaratram D (2005) Promoting asset management policies by considering OEE in products' TLCC estimation. In: IEEE international proceedings of engineering management conference
11. Woodhouse J (2004) PAS 55: specification for the optimized management of physical infrastructure assets. http://www.iam-uk.org/downloads/PASworkshop.pdf. Accessed 28 Nov 2012
12. Mardiasmo (2009) Kebijakan Desentralisasi Fiskal di Era Reformasi: 2005–2008. In: Abimanyu A, Megantara A (eds) Era Baru Kebijakan Fiskal: Pemikiran, Konsep, dan Implementasi. PT Kompas Media, Jakarta
13. Ritonga AA (2009) Upaya-upaya Efisiensi Belanja Negara. In: Abimanyu A, Megantara A (eds) Era Baru Kebijakan Fiskal: Pemikiran, Konsep, dan Implementasi. PT Kompas Media, Jakarta
14. Mardiasmo D (2007) Good governance implementation and international alignment: the case of regional governments in Indonesia. School of management, faculty of business, Queensland University of Technology, Brisbane, Queensland
15. Pardiman (2009) Prinsip-Prinsip Barang Milik Negara/Daerah (BMN/D). In: Workshop Training of Trainer Latihan Keuangan Daerah dan Kursus Keuangan Daerah. Direktorat Jenderal Kekayaan Negara, Jakarta, Indonesia
16. Denzin NK, Lincoln YS (1994) Handbook of qualitative research. Sage, Thousand Oaks
17. Ezzy D (2002) Qualitative analysis: practice and innovation. New Fetter Lane, Routledge
18. Golden-Biddle K, Locke K (1997) Composing qualitative research. Sage, Thousands Oaks
19. Padgett D (2004) The qualitative research experience. Wadsworth Publishing, New York
20. Silverman D (1997) The logics of qualitative research. In: Miller G, Dingwall R (eds) Context and method in qualitative research. Sage, London
21. Sofaer S (2002) Qualitative research methods. Int J for Qual in Health Care 14:329
22. Willis JW (2007) Foundations of qualitative research: interpretive and critical approaches. Sage, Thousand Oaks

23. Yin RK (1994) Case study research: design and methods. Sage, Thousand Oaks
24. Siddik M (2009) Kebijakan Awal Desentralisasi Fiskal 1999–2004. In: Abimanyu A, Megantara A (eds) Era Baru Kebijakan Fiskal: Pemikiran, Konsep, dan Implementasi. PT Kompas Media, Jakarta
25. Devas N (1997) Indonesia: what do we mean by decentralization? Public Adm Dev 17:351
26. Federspiel HM (2005) Regionalism in post-Suharto Indonesia. Cont Southeast Asia 27:529
27. Schroeder L (2003) Fiscal decentralization in Southeast Asia. J Public Budg, Acc Fin Manage 15:385
28. Crane R (1995) The practice of regional development: resolving central-local coordination issues in planning and finance. Public Adm Dev 15:139
29. Kristiansen S, Trijono L (2005) Authority and law enforcement: local government reforms and security systems in Indonesia. Cont Southeast Asia 27:236
30. Silver C (2003) Do the donors have it right? Decentralization and changing local governance in Indonesia. Ann Reg Sci 37(3):421
31. Tambunan M (2000) Indonesia's new challenges and opportunities: blueprint for reform after the economic crisis. East Asia An Int Q 18(2):50
32. Indrawati SM (2005) Efektivitas Kebijakan Fiskal 2006 (pp 6). Departemen Keuangan Republik Indonesia, Jakarta, Indonesia
33. Kivimaki T, Thorning R (2002) Democratization and regional power sharing in Papua/Irian Jaya. Asian Surv 42:651
34. Mishra SC (2005) Pembuatan Kebijakan Demokratis dalam Konteks yang Berubah; Bahan Latar Belakang untuk Laporan Studi Mengenai Proses Pembuatan Kebijakan di Indonesia Discussion Paper Series No. 05/7-IND (pp 94). United Nations Support Facility for Indonesian Recovery, Jakarta
35. Davis BJ (2005) PREPARE: seeking systemic solutions for technological crisis management. Knowl Process Manage 12:123
36. Horine L (1990) The Johari window-solving sport management communication problems. J Phys Ed, Rec Dance 61:49
37. Sole D (1997) Johari's window for generating questions. J Adolesc Adult Lit 40:481
38. Mardiasmo (2004) Otonomi dan Manajemen Keuangan Daerah. Andi Offset, Yogyakarta
39. Bessant J (2010) El Dorado adventures: a policy making assignment in South America. Paper presented at the The School of Justice Seminar Series Brisbane, Australia

Chapter 41
Policy Delphi: Contribution to Infrastructure and Engineering Asset Management Organizations

Mary M. McGeoch, Kerry Brown and Yvonne Brunetto

Abstract Policy Delphi was first introduced in 1969, and it is a process that seeks to generate the strongest possible opposing views on the potential resolutions of a major policy issue, such as strategy and policy for infrastructure and engineering asset management within organizations. The objectives of the Policy Delphi are to: ensure that all possible options have been tabled for consideration; estimate the impact and consequences of any particular option; and examine and estimate the acceptability of all the individual options. There are six phases that include formulation of the issues, identifying options, determining initial positions on the issues, exploring and obtaining the reasons for disagreements, evaluation, and finally, reevaluation of options. The methodology relies on a two-round limit by utilizing the following procedures: (a) The monitor team devoting a considerable amount of time to carefully pre-formulate the obvious issues; (b) Seeding the list with an initial range of options but allowing for the respondents to add to the lists; and (c) Asking for positions on an item and establishing underlying assumptions in the first round. The research investigates what kind of asset management experts is needed to determine the scope of policy requirements for this process? The paper provides a case study of a Policy Delphi to elaborate the efficacy of this method for asset-intensive organizations to develop strategic policy.

M. M. McGeoch (✉)
Southern Cross Business School, Southern Cross University, Lismore, Australia
e-mail: mary.mcgeoch@scu.edu.au

K. Brown
School of Tourism and Hospitality Management, Southern Cross University, Lismore, Australia

Y. Brunetto
Cooperative Research Centre for Infrastructure and Engineering Asset Management (CIEAM), Queensland University of Technology, Brisbane, QLD 4001, Australia

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_41,

41.1 Introduction

A committee may be simply defined as a particular group of people put together to work out an issue [1]. Committees are commonly used to address a wide range of issues within organizations at management and operational levels. Indeed, it is fair to say that committees dominate governance and management within nearly all types of organizations. Committees are an institutional norm of corporations, the governance of which typically, and increasingly, is addressed through board committees encompassing a range policy issues pertaining to management support, monitoring and oversight, public policy and social responsibility [2, 3]. Committees are commonly used in the workings of public sector agencies (see for example [1]). Committees represent formal sites for interaction between members, aiding information exchange and 'sense-making' [1]. They provide a formal structure through which a group of people can explore an issue, form recommendations, and implement a particular initiative [1]. The committee system initially emerged as a way to promote the advocacy process associated with policy analyses, and in most organizations, the examination and exploration of policy issues occur through a committee process. The committee structure connects people across organizational lines at similar organizational levels, enabling representation of various interests and presentation of differing views, hence encouraging the distillation of a consensus through processes of expression and advocacy. Without effective application, the committee approach arguably may not function effectively in the realm of policy formulation [4].

The Delphi technique is seen as an alternative to the conventional committee approach that can be used to produce information and to stimulate creative exploration of ideas for decision-making. It provides a method for structuring a group communication process in such a way as to enable a group to deal with a complex problem [4]. The Delphi technique bypasses the typical problems of group dynamics in committees. Communication is structured through feedback of individual contributions, assessment of the group judgment or view, and opportunities for revision of individual views [5]. The key characteristics of the Delphi technique are the structuring of information flow; the provision of feedback to the participants; and participant anonymity [6] or anonymity, iteration, controlled feedback, and the statistical aggregation of group response [5]. The Delphi technique was originally developed as a systematic, interactive forecasting method utilizing a panel (group) of geographically separated experts to analyze information about a, usually technical, issue in two or more rounds. After each round, a facilitator summarized the experts' forecasts from the previous round, along with the rationale for the experts' earlier judgments. This information was then provided to the expert panelists, who could then revise their earlier answers in light of the replies of other members of the panel. This iterative process would continue so as to reduce the range of the answers, the aim being that the group would ultimately converge toward the 'correct' answer. The process concluded upon the attainment of a pre-defined stop criterion (e.g. number of rounds, achievement of consensus or stability of results), at which point the mean or median scores of the final round determined the results [5]. Delphi remains a popular technological forecasting technique [7]. It is a widely used and accepted method for achieving convergence of opinion concerning real-world knowledge solicited from experts within certain topic areas [6].

The Policy Delphi method was introduced in 1969 and was first reported upon in 1970. It represented a significant departure from the understanding and application of the Delphi technique up to that point in time. Until then, the use of the Delphi technique had been limited mainly to technical topics and had typically focussed upon the attainment of consensus among homogeneous groups of experts [4]. In stark contrast, the Policy Delphi purposely seeks to generate widely diametrically opposing on potential alternative resolutions of a given policy issue. The Policy Delphi technique concerns systematic attempts to identify widely divergent views and facilitate consensus among stakeholders with opposing views [8]. An expert may contribute a quantifiable estimation of some effect resulting from a particular resolution of a policy issue, but this would be unlikely to result in a consensual resolution of the policy issue [5].

It has been suggested (e.g. [8]) that it is unfortunate that the Policy Delphi technique has not been more widely used, as it is suitable method to address a multiplicity of policy issues in complex arenas. While their observation is made within the context of health policy, the same line of reasoning beckons in the similarly complex environment of asset organizations.

41.2 The Complexity of Engineering Asset Organizations

Asset management is an area of research that concerns the management of engineering assets including plant, equipment, and infrastructure. Asset management has received increasing attention within both academic and industry circles as an organizational undertaking of strategic importance [9–14]. Many large organizations own substantial plant, property, and infrastructure assets, prompting the need for a strategic approach to asset management (see for instance: [13]). The challenge of managing assets to achieve organizational objectives is most acutely obvious within these asset organizations. Asset management is increasingly viewed as being central to organizational efforts to manage risks and maximize the performance of physical assets. The purpose of asset management arises within the organizational context. For private corporations, the performance of physical assets is important because it ultimately impacts upon business competitiveness, profitability, and hence survival. For public organizations, the performance of physical assets is important because it ultimately affects social or economic functions that are important to society [15]. Service delivery is fundamental to assets are important because they support service delivery so the overarching goal of asset management concerns optimizing the service delivery potential of assets [9].

The asset management undertaking embodies significant challenges. Foremost among these are that asset-intensive businesses typically confront escalating pressure to deliver improved financial returns on the investment in their physical assets, while simultaneously facing diminution in the resources available to achieve this outcome. This basically represents the challenge of organizational efficiency or 'doing more with less'. This challenge of organizational efficiency is hardly new, and in fact, organizations have grappled with the problems of obtaining greater efficiencies from infrastructure assets for a considerable time [16–18]. However, rising customer expectations also mean that there is a need to achieve greater service delivery returns from physical assets as well. Hence, there is an increasingly pressing need to just 'do more' too. Customers have progressively come to expect higher standards of service delivery, and asset management has become increasingly important as the means through which such demands can be addressed. Transfield's power generation services is a business that operates and maintains power stations, some of which are part company-owned. Transfield's customers now want more than base-level operating and maintenance services, and the market is demanding a more comprehensive suite of value-added services; this requires that service providers such as Transfield engage more closely with their business customers and become more intimately involved in efforts to create long-term value for customers' businesses [19].

The twin challenges of rising customer expectations ('doing more') and escalating efficiency pressures ('doing more with less') necessitate the delivery of lasting and sustainable performance improvements and so fundamentally drive the need for continual improvement in asset management [20]. Enterprises face the challenge of needing to maintain or improve operational effectiveness, revenue, and customer satisfaction, while simultaneously reducing capital, operating and support costs. This impacts competiveness; hence, asset management ultimately plays an increasingly important role in optimizing business profitability [21].

The durability of physical assets necessitates a whole-of-life perspective if the full benefits of asset management are to be achieved. To gain even greater value, the asset management process should extend from design, procurement, and installation through operation, maintenance, and retirement, i.e. over the complete life cycle [21]. Accordingly, owners of asset-intensive enterprises are increasingly seeking robust asset management and reliability solutions in both the design and operational aspects of their businesses. So fundamental is this life-cycle perspective that asset management is defined in these very terms, namely as the life-cycle management of physical assets to achieve the stated outputs of the enterprise [19]. Similarly, the British standard PAS 55 defines asset management as systematic and coordinated activities and practices through which an organization optimally manages its physical assets and their associated performance, risks, and expenditures over their life cycles for the purpose of achieving its organizational strategic plan [22].

Simultaneous fulfillment of organizational efficiency and service delivery requirements, through the management of enduring and valuable physical assets presents a complex challenge. Power generators such as Transfield provide a good example of this complex asset management challenge so much so that the power generation sector is considered to be where the discipline of asset management actually originated. In particular, power stations comprise discrete, capital-intensive, and durable assets; with the bulk of investment expense arising at the beginning of a long asset life, the owners of power stations rely upon assured performance from their assets; asset management is pivotal in the attainment of this outcome as it enables an asset to perform at the required level throughout its entire life cycle, in a safe and cost-effective manner [19].

41.3 Key Issues in Managing Complex Physical Assets

An asset-intensive operation's life cycle is a series process comprising: feasibility, detailed design, procurement, installation, commissioning, and operation. Any mistake, miscalculation or misjudgement made anywhere over the life cycle will inevitably find its way into operating costs. This is particularly important in the initial feasibility and design phases because as equipment is selected, so too are the operating costs that come with them. In fact, most operating costs are the result of the choices made during feasibility and design. According to a study [17], 65 % of a facility's life-cycle costs (LCCs) are fixed during the design phase [21]. Low operating costs and high profits are therefore largely established during the initial project phase, and vice versa. High costs and operational failures are nearly always due to project phase choices: Latent problems introduced during feasibility and design phases will eventually cause troubles and higher costs during operation [19].

Asset management implementation and maintenance strategies can be applied to address life-cycle considerations and their associated risks and cost implications. For example, equipment reliability is malleable by choice of policy and quality of practice; reliably removing the causes of equipment failure will result in equipment that works trouble-free for decades. Reliability growth cause analysis (RGCA) can be applied to identify and prevent failures of equipment parts from errors and defects created during the life cycle by engineering, manufacturing, installation, operations, maintenance, and procurement [17]. RGCA allows the identification of operational risk in all stages of a part's life cycle so that the necessary improvements to policy, processes, and practices can be made to reduce operational risk and meet availability goals. However, although such logic is well understood in terms of asset and service availability, scant consideration has been given to safety as a dimension of risk and cost in asset management. Notably, safety concerns risk and the impact of asset management as a probability driver in respect to asset failures that have a safety consequence [23]. Further, as an oft-cited concern in the mission statements and core values of Australian asset organizations, one would expect safety to enjoy an equal weighting to availability in asset management strategies of asset operators. Nevertheless, there are in fact few asset management implementation and maintenance strategy types presently available and utilized, where safety is as much a part of defining a level of service as asset performance [19]. This is an important issue in need of research attention in asset management.

Asset operators confront an array of, often related, risks that need to be considered when managing physical assets. For example, a major risk facing electricity distributors, such as Powercor, is the risk that their vast networks of electricity distribution assets might cause a major bushfire. For instance, Powercor's electricity distribution assets include some 530,000 poles, 85,000 km of lines (circa 9,000 km underground and 76,000 km of overhead lines), 81,553 transformers, and 137 zone substation transformers [24], any of which can potentially ignite a fire. Furthermore, because these assets are often situated in bushland areas susceptible to bushfires, which are a common occurrence in Australia, electricity distributors such as Powercor confront a risk of asset damage from bushfires otherwise ignited. So in considering the potential consequences of any fire, electricity distributors such as Powercor must also recognize that their assets are exposed to risk of damage and destruction from fire as well, irrespective of the fire's cause. Accordingly, asset management risk strategies need to consider service delivery and safety. Asset management is an interactive process that requires attention to engineering, business planning, and community implications in the determination of desired and affordable service levels [16]. Railway organizations recognize this, and nowadays, most railway organizations adopt a similar emphasis, whereby safety is as much a part of defining a level of service as on-time running: The goal of a railway system is now typically expressed as achieving a defined level of rail traffic in a given time and safely. For instance, RailCorp stresses safety and service delivery as core values and states its mission in these terms, namely to deliver safe, clean, and reliable services that are efficient, sustainable and to the satisfaction of all customers [25].

Similar views are reflected in the mission statements and core values of other Australian asset operators. Risk evaluation is a multidimensional task, requiring attention to commercial, technical, safety or customer/public perception risks. The management of complex physical assets is even further complicated by factors affecting the implementation of asset management strategies. Particularly, when managing physical assets in complex organizations, the asset management regime may encounter significant barriers to implementation. Such as silo thinking, short-termism, communication barriers, and risk evaluation complexity may all variously operate to undermine effective implementation of asset management within asset organizations [26].

Silo thinking occurs when functional, departmental or geographic barriers between groups operate to prevent collaboration and shared solutions being generated. This concept draws upon the metaphor of independent grain storage silos to depict disconnectedness between different parts of an organization, typically describing situations of organizational fragmentation and dysfunction; silos give rise to a 'them and us' mentality, impeding communication and cooperation between groups within an organization. This often arises as a result of previous poor experience of organizational change, strong local management personalities, and/or badly structured performance/reward mechanisms. Silo thinking may, for instance, confront asset organizations when those responsible for financial management come to perceive those making asset maintenance expenditures as people who complicate or impede established budgeting goals [26, 27].

Common language is essential to communication and discourse; this is because common language allows the sharing of information with a mutual understanding of its terms. Communication barriers may impede implementation of asset management simply because the lexicon of different groups leads to miscommunication, misunderstanding, and mistakes. For example, the language used by those in engineering and facilities management roles often differs from that used by those in finance and business management roles. Such language differences may also reinforce cultural divides, contributing to the formation of, or reinforcing organizational silos [28].

Short-termism describes the tendency to focus attention on short-term outcomes without due consideration to, the often crucial, longer-term consequences, an issue of particular concern in light of the inherent long-term nature of physical assets. Short-termism is defined as a systematic characteristic of an organization that overvalues short-term rewards and undervalues long-term consequences. It may be an inherent organizational characteristic arising from culture, processes, or routines or from managerial characteristics that tend to lead to myopic decisions overvaluing short-term rewards and undervaluing long-term consequences. An example of short-termism within asset organizations can be found in outsourced project work, where success is typically measured in terms of 'on-time' and 'on budget' project delivery, irrespective of longer-term asset performance and value. As mentioned previously, poor decisions in the project phase will have a substantial impact on subsequent operating costs [29].

Conflict relates to incompatible activities, goals or preferences, and it is an interactive process manifested in incompatibility, disagreement or dissonance within or between social entities (i.e. individual, group, organization, [30]). Conflicting performance measures, wherein one group can only succeed at the expense of another, can foment competition and conflict. Even balanced scorecards can reinforce competing priorities and foment conflict. For example, the financial controller whose performance is evaluated according to cost minimization will likely come into conflict with the asset manager whose performance is judged on asset availability without regard to cost.

Risk analysis entails the identification of potential sources of harm, the likelihood of their occurrence, and the severity of consequences arising. Risk management addresses what might go wrong and also what should be done about it [12]. Complex asset organizations typically face a wide range of risks. Risk management requires the rational and consistent identification, quantification, and management of commercial, technical, safety, or customer/public perception risks [26]. Poor risk management impedes the implementation of effective asset management, the objective of asset management is to optimize the service delivery potential of assets, and this requires the intelligent deployment of business systems to minimize risks and costs [9].

The managers of asset organizations face a complex task; they need to undertake appropriate consideration of risks, financial and non-financial business objectives, short- and long-term consequences, and the inevitable trade-offs that occur therein. The common thread underlying many of the problems in asset management implementation is arguably found in the lack of structured, fact-based decision methods. Indeed, some asset managers take the view that clear and auditable processes are needed to identify what type of data is needed and how it should be used in supporting a structured, fact-based decision-making approach to asset management. Overreliance on data and models potentially reduces strategic thinking; so when asset management is more narrowly focused on feeding data into models, organizations run the risk of failing to be strategic in planning, development, and operations. Furthermore, while models seemingly offer a reliable and verifiable basis for decision-making, this is not necessarily so [31]. In particular, when asset managers attempt to extrapolate data from the past to inform strategies and actions for the future, they necessarily rely on assumptions, which guide interpretation and decision [31, 32]. These assumptions are often implicit and unrecorded; hence, defying external examination or analysis and precluding verification and validation, even when underlying assumptions are explicit, in practice, they are rarely subjected to any examination. Underlying assumptions in asset liability modeling are often neither widely understood nor necessarily consistent with each other and so can be difficult to change [33].

41.4 Policy Delphi Applied to Complex Asset Organizations

Asset organizations have become bigger and more complex; they typically deliver a wide range of functions and span a wide range of complex interacting processes. Escalating pressures of organizational efficiency and rising customer service delivery expectations compound the challenges for complex asset organizations. The complex and dynamic context of asset organizations presents a significant challenge to the management of asset organizations and particularly to traditional tools such as committees. Within this context, the committee structure becomes an unwieldy, and arguably unworkable, method for including the full range of interests pertinent to addressing a given issue and within an acceptable time frame. Specifically, committees that truly represent all interests on an issue are often quite large and unwieldy; by the time one has reached the point of twenty or more people, a complete and free exchange of views among all concerned is often impossible, at least within the constraints of the time and resources allocated for the job [5, 16, 34–37].

With the increasing size and complexity of asset organizations, the ratio of the number of people at the top echelons to those in the lower echelons has decreased over time, particularly in government agencies. This trend implies that those at the top must now spend more time devoted to day-to-day management functions and have less time for committee participation on the longer-range issues associated with policy [12]. Consequently, the responsibility for committee participation falls increasingly on lower-level staff within organizations. Individuals at these lower levels have constrained decision-making authority and are often unable to be advocates of anything until they have had ample time to clear it with their supervisors. This need for checking often forces the committee into added, sometimes lengthy, delays whenever any new point is made, and this usually results in the early or premature termination of new considerations that might result from the advocacy process [5]. The difficulties are compounded by the fact that the rising complexity of issues facing modern asset organizations necessitates a great deal of additional staff resources to properly support the committee process. At best, this requirement only further convolutes the committee process; however, more often than not, this support, or the time needed for its delivery, is simply not available to committee participants. In an atmosphere of ruthless efficiency, budget cuts and competition for limited funds may well prompt risk-averse behavior by committee members; it may well appear advantageous not to advocate, not to be noticed, and especially not to be held accountable for views, promises, or positions which require effort (resources) to document or substantiate [5].

Different cognitive approaches also influence decision-making processes and determine the way that people rely upon and use data and decision aids. As some people analyze information and then make decisions using their intuitive feeling, others make emotional decisions and then justify the hunch with rational analysis [32]. In most organizations, however, individuals are not familiar with many of the new decision aids coming out of operations research and systems analyses, instead relying on their intuitive ‘feel’ for the complexities of the particular business or function the organization is involved in. Others, on the other hand, may be familiar with modern management techniques but may suffer from dogmatism, sometimes being overly confident that these approaches can be applied to every problem. The lack of effective communication between these two groups has brought about the ineffectiveness of many committees [5].

While it is the above contextual factors, or any combinations of thereof, that have generally motivated attempts to seek substitutes for the committee process, other, different factors also warrant mention. Notably, the earlier writings on Delphi have usually presented a separate but canonical set of problems associated with committees that largely reflect psychological, as opposed to contextual, characteristics of committee processes, namely:

- The domineering personality or outspoken individual that takes over the committee process;
- The unwillingness of individuals to take a position on an issue before all the facts are in or before it is known which way the majority is headed;
- The difficulty of publicly contradicting individuals in higher positions;
- The unwillingness to abandon a position once it is publicly taken; and
- The fear of bringing up an uncertain idea that might result in embarrassment [5].

41.5 Application of the Policy Delphi to Complex Asset Organizations

The Policy Delphi method can be utilized to revise the effectiveness of the predominant committee approach. A Policy Delphi can be administered to anywhere from ten to fifty people as a precursor to a committee activity. Its goal in this application is to expose all the differing positions advocated and the principal arguments ‘for’ and ‘against’ positions on a specific policy issue under. Once the Delphi has been accomplished, a smaller, and workable, committee can then utilize the results to formulate the required policy. As a precursor to the committee approach, the Policy Delphi method presents as a mechanism for addressing the contextual and psychological factors variously undermining the effectiveness of the committee process, thereby improving the effectiveness of lateral policy formulating committees and reviving the advocacy process within complex asset organizations [5]. Rather than supplanting the dominant committee approach, the Policy Delphi assists it to be more effective. This combination offers ‘the best of both worlds.’ Particularly, the Policy Delphi provides an organized method for correlating views and information pertaining to a specific policy area and for allowing the respondents representing such views and information the opportunity to react to and assess differing viewpoints. Because the respondents are anonymous, fears of potential repercussions and embarrassment are removed and no single individual needs to commit themselves publicly to a particular view until after the alternatives have been identified. Even in those cases where the Policy Delphi uses only the committee or sponsoring body as the respondent group, it retains the advantage of eliminating the principal bottleneck in the committee procedure by providing a clear delineation of specific differing views, thereby providing an opportunity for the committee members to prepare their respective cases adequately. A Policy Delphi can be utilized to serve any one or any combination of the following objectives:

- To ensure that all possible options have been presented for consideration;
- To estimate the impact of, and consequences arising from, any particular option;
- To examine and estimate the acceptability of any particular option [5].

To be effective, however, the Policy Delphi needs to be properly implemented. Implementation of the Policy Delphi method is a very demanding exercise, both for the design team and for the respondents. There are six phases in the communication process that takes place, each giving rise to key questions that must be carefully considered. These are:

1. Formulation of the issues: What is the issue that really should be under consideration? How should it be stated?
2. Exposing the options: Given the issue, what are the policy options available?

3. Determining initial positions on the issues: Which are the ones everyone already agrees upon and which are the unimportant ones to be discarded? Which are the ones exhibiting disagreement among the respondents?
4. Exploring and obtaining the reasons for disagreements: What underlying assumptions, views, or facts are being used by the individuals to support their respective positions?
5. Evaluating the underlying reasons: How does the group view the separate arguments used to defend various positions and how do they compare to one another on a relative basis?
6. Reevaluating the options: Reevaluation is based upon the views of the underlying 'evidence' and the assessment of its relevance to each position taken [5].

41.6 Summary and Conclusion

The Policy Delphi method provides a useful means to inform asset management implementation within the context of complex asset organizations. The Policy Delphi can be applied to assist the adoption of scientifically supported continuous improvement processes in asset management, in order to identify and remove operating risk, and achieve maximum equipment reliability, operating availability, and asset utilization. Ideally, this objective would be linked to the overarching asset management strategy in order to holistically address the complete life cycle of the operational assets that are critical to successful performance of asset organizations, from business concept through to decommission [36].

The formation of a committee is the predominant method for dealing with policy within asset-intensive organizations. It is demonstrated that committees face limitations, particularly in the increasingly complex and challenging environment of asset management within complex engineering asset organizations.

Asset management is a complex and challenging environment, compounding asset management policy formation challenges. Some critical areas, such as safety, have not been well integrated into asset management strategy. A holistic approach is needed in asset management, and this approach is vital to service delivery and community needs.

Delphi is an alternative approach to the committee process. A derivative of the process, the Policy Delphi is suited to complex arenas such as asset management organizations. This process can be used to reform and reinvigorate advocacy processes within asset management. This could be done via supplementing, rather than supplanting, the committee approach, as a precursor that delivers the best of both worlds.

This would ideally be achieved by integrating Policy Delphi to aid adoption of scientifically supported continuous improvement process in asset organizations to optimize service delivery outcomes taking cost and risks into account. This should then be integrated into the holistic asset management strategy necessary to meet the demands of the increasingly complex challenge of asset management that confronts modern asset managers.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established, and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Hoon C (2007) Committees as strategic practice: the role of strategic conversation in a public administration. Hum Relat 60(6):921–952
2. Carson E (2002) Factors associated with the development of board sub-committees. Corp Governance: Int Rev 10(1):4
3. Harrison JR (1987) Strategic use of corporate board committees. Calif Manag Rev 30(1):109
4. Linstone HA, Turoff M (1975) Delphi method: techniques and applications. Addison-Wesley Publishing, Boston
5. Rowe G, Wright G (1999) The Delphi technique as a forecasting tool: issues and analysis. Int J Forecast 15:353–375
6. Hsu CC, Sandford BA (2007) The Delphi technique: making sense of consensus. Pract Assess Res Eval 12(10):1–8
7. Martino JP (2003) A review of selected recent advances in technological forecasting. Technol Forecast Soc Change 70(8):719–733. doi:http://dx.doi.org/10.1016/S0040-1625(02)00375-X
8. Rayens MK, Hahn EJ (2000) Building consensus using the policy Delphi method. Policy Polit Nurs Pract 1(4):308–315. doi:10.1177/152715440000100409
9. AAMCoG (2011) Guide to integrated strategic asset management (version 1.0 ed.). Australian Asset Management Collaborative Group (AAMCoG)
10. Brown KA, Keast R, Wiewiora A, Laue M, Mahmood MN (2012) ISAM for XPERTS: integrated strategic asset management (Advanced Copy). CIEAM (Cooperative Research Centre for Infrastructure and Engineering Asset Management) & AAMCoG (Australian Asset Management Collaborative Group), Brisbane, Australia
11. Gruis V, Nieboer N, Thomas A (2004) Strategic asset management in the social rented sector: approaches of Dutch and English Housing Associations. Urban Stud 41(7):1229–1248. doi: http://usj.sagepub.com/archive/
12. Jolicoeur PW, Barrett JT (2004) Coming of age: strategic asset management in the municipal sector. J Facil Manage 3(1):41–52
13. Office of the Water Supply Regulator (2010) Guidelines for preparing strategic asset management plans. Queensland Government retrieved from http://www.derm.qld.gov.au/water/regulation/pdf/guidelines/guidelines_03.pdf
14. Woodhouse J (2009) Making the most of your assets. Process Eng 90(1):8–9
15. Amadi-Echendu J, Willett RJ, Brown KA, Lee J, Mathew J, Vyas N, Yang BS (2007) What is engineering asset management? Paper presented at the world congress on engineering asset management and the 4th international conference on condition monitoring, Harrogate, United Kingdom
16. AMQ International (2008) Strategic asset management
17. Barringer HP (1998) Life cycle cost and good practices. Paper presented at the National Petrochemical & Refiners Association reliability and maintenance conference, San Antonio, TX
18. Childs J (1998) The military use of land: a history of the defence estate. Peter Lang Publishing Incorporated, New York

19. The Asset Management Council (2007) The Asset Management Council perspectives: major ASSET owners back council objectives. Asset J (online) 2
20. Canders M (1999) Strategy and performance in the Defense Industry: an empirical investigation. PhD thesis, Walden University
21. Schuman CA, Brent AC (2005) Asset life cycle management: towards improving physical asset performance in the process industry. Int J Oper Prod Manage 25(6):566–579. doi:10.1108/01443570510599728
22. Woodhouse J (2007) Asset management: joining up the jigsaw puzzle-PAS 55 standards for the integrated management of assets. Maint Eng (Sept/Oct)
23. Mardiasmo D, Tywoniak S, Brown KA, Burgess K (2008) Asset management and governance: an analysis of fleet management process issues in an asset-intensive organisation. In: International conference on infrastructure systems: building networks for a brighter future, 10–12 November, Rotterdam, Netherlands
24. CitiPower & Powercor Australia (2012) Powercor network statistics. Retrieved 29 December 2012, from http://www.powercor.com.au/Powercor_Network_Statistics/
25. RailCorp (2012) Corporate plan 2012–2013. Sydney, Australia. NSW Government Retrieved from http://www.railcorp.info/__data/assets/pdf_file/0010/7021/RailCorp_Corporate_Plan.pdf
26. Woodhouse J (2001) Asset management decision making. Strateg Asset Manage 173–175
27. Cilliers F, Greyvenstein H (2012) The impact of silo mentality on team identity: an organisational case study. SA J Ind Psychol 38(2)
28. Lum F (2005) Editorial: why is a terminology important? Ophthamology 112(2):173–174. doi:10.1016/j.ophtha.2004.11.024
29. Laverty KJ (2004) Managerial myopia or systemic short-termism? The importance of managerial systems in valuing the long term. Manag Decis 42(7/8):949–962
30. Rahim MA (2002) Toward a theory of managing organizational conflict. Int J Confl Manage 13(3):206
31. Burns P (2005) Assumptions and 'gut feel'. Strateg Asset Manage
32. Azavedo D (2005) Gut feel. Strateg Asset Manage 1010–1011
33. Dempster MAH, Germano M, Medova EA, Villaverde M (2002) Global asset liability management. Presented to the Institute of Actuaries, 25 November 2002
34. Weijnen MPC, Herder PM, Bouwmans I (2009) Designing complex systems, a contradiction in terms. In: Eekhout M, Visser R, Tomiyama T (eds) Delft science in design 2. IOS Press, Amsterdam
35. Bauer JM, Herder PM (2009) Designing socio-technical systems. In: Gabbay DM, Meijers A, Thagard P, Woods J (eds) Handbook of the philosophy of science: handbook philosophy of technology and engineering sciences. Elsevier, North-Holland, pp 601–631 (ISBN-13: 978-0-444-51667-1, ISBN-10: 0-444-51667-0)
36. Harpur PD (2008) Clothing manufacturing supply chains, contractual layers and hold harmless clauses: how OHS duties can be imposed over retailers. Aust J Labour Law 21(3):316–346
37. US Office of Personnel Management (1999) Strategic human resources management: aligning with the mission. US Government, USA

Chapter 42
A Single Cavitation Indicator Based on Statistical Parameters for a Centrifugal Pump

Kristoffer K. McKee, Gareth Forbes, Ilyas Mazhar, Rodney Entwistle, Melinda Hodkiewicz and Ian Howard

Abstract Cavitation is one of the major problems associated with the operation of centrifugal pumps. Cavitation occurs when vapor bubbles that are formed due to a drop in pressure in the pipes upstream of the centrifugal pump implode under the added pressure within the volute of the pump. These implosions wear away the impeller, and sometimes the volute itself, which if left unchecked, would render the pump inoperable. Much research has been done in the detection of cavitation through: indicators in certain audible frequencies, drop in the net positive suction head (NPSHa), visual inspection using a transparent casing and a stroboscopic light, paint erosion inside the volute, and on the impeller, changes in pressure within the flow or volute, and vibration within certain frequency ranges. Vibration detection is deemed as one of the more difficult methods due to other structural and environmental factors that may influence which frequencies may be present during the onset of cavitation. Vibration measurement, however, is most easily measured and deployable in an automated condition monitoring scenario. It is proposed that an increasing trend in a set of statistical parameters, rather than a firm threshold of a single parameter, would provide a robust indication for the onset of cavitation. Trends in these statistical parameters were obtained from data collected on a pump forced to cavitate under several different operating conditions. A single cavitation indicator is outlined utilizing these statistical parameters that can quantify the level of cavitation in a centrifugal pump.

K. K. McKee (✉) · G. Forbes · I. Mazhar · R. Entwistle · I. Howard
Department of Mechanical Engineering, Perth, Australia
e-mail: k.mckee@curtin.edu.au

M. Hodkiewicz
School of Mechanical and Chemical Engineering, The University of Western Australia, Perth, Australia

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_42,

42.1 Introduction

Centrifugal pumps are used in a wide variety of applications in industry, ranging from pumping clean water into cities to pumping waste water or sludge into treatment plants. They are composed of two main parts, the impeller and the volute. The impeller, either open or closed, imparts kinetic energy via rotation to the incoming fluid as it moves through the impeller vanes. The volute redirects the fluid from the impeller and channels it in a single direction. Through this action, the fluid's kinetic energy changes from kinetic energy to pressure. Consequently, the main goal of the pump is to add pressure to the incoming fluid and channel it in a desired direction.

Cavitation is one of many fault modes associated with the centrifugal pump and is also one of the most common ones. Cavitation is the formation and subsequent implosion of vapor bubbles in the moving fluid. These vapor bubbles are formed when the pressure upstream of the pump decreases below the vapor pressure of the fluid. The additional pressure added to the fluid in the pump causes the vapor bubbles to implode, causing a release in kinetic energy, which produces damage to the impeller and volute. Despite all attempts by pump designers to avoid this fault mode, it inevitably happens to most pumps since they cannot always run at or close to their best efficiency point. In terms of quantifiable amounts, cavitation occurs when the actual net positive suction head (NPSHa) within a system falls below that which is required by the pump manufacturer (NPSHr).

Symptoms of cavitation are not usually apparent until the severity of the damage is high. When this happens, a low-frequency vibration, reduction in pumping efficiency, erosion, and a crackling noise usually described as "pumping gravel" accompanies the phenomena. Extensive research has been conducted attempting to determine early onsets of cavitation via experimental or mathematical means, mostly to no avail. Various methods have been explored including data-driven models (statistical methods and neural networks), knowledge-based models (expert systems and fuzzy logic systems), and physical models [1].

42.2 Background Research

Čudina used a set of microphones near a centrifugal pump setup to detect the beginning of cavitation. Experimentally, it was determined that a discrete frequency tone of half the blade pass frequency, or 147 Hz in this case, was separate from the noise spectra of the pump and was a clear indication of the onset of cavitation and its development [2, 3]. In a later study, using three different pumps to verify this hypothesis, Čudina et al. determined that there was a distinct characteristic frequency based on pump design, such as the pump's geometry and material used, that increased with the onset of cavitation. This frequency was separate from the noise spectra and can be determined from the first positive slope on the noise-level curve [4].

Neil et al. performed industrial scale tests to determine whether acoustic emission sensors are able to determine the occurrence of cavitation in a centrifugal pump. Experimental conclusions showed that acoustic emissions RMS values spiked at the onset of cavitation, and then decreased after [5]. Alfayez et al. performed a case study to verify the correlation between the increases in the acoustic emissions RMS levels on a centrifugal pump to the point of cavitation inception, followed by a decrease in the acoustic emissions RMS levels afterward. They concluded that AE RMS levels could detect the onset, not the development, of cavitation [6].

Athavale et al. created a full cavitation model, utilizing reduced Rayleigh-Plesset equations to account for bubble formation and to determine the time-mean phase-change rates utilizing the local pressures and characteristic velocities. The models were incorporated into commercial CFD code. Simulations of cavitating flows were conducted on a water jet propulsion axial pump, a centrifugal water pump and an inducer from a LOX turbo pump [7]. Hofman et al. [8] created 3D numerical models of centrifugal pumps and simulated cavitation. Uchiyama [9] created a numerical simulation utilizing a finite element method to attempt to predict cavitation.

Franz et al. attempted to determine the rotor dynamic forces on an impeller during cavitation. They concluded that cavitation corresponding to a head loss of three percent did not have a significant effect upon the unsteady force. However, a lesser degree of cavitation had caused an increase in the destabilizing forces for a given set of whirl ratios [10]. Rapposelli et al. [11] created a cavitating pump rotordynamic test facility that was able to measure the instantaneous fluid loads and thermal effects on an impeller during cavitation. Jeremy et al. [12] determined that dynamic pressure monitoring is a direct indication of cavitation and a viable method of detection, while monitoring NPSHa is an indirect indicator and not as reliable.

McKee et al. performed research on the advantages of using an adapted version of the octave band spectral analysis for vibration data on a centrifugal pump. Three types of limits were applied to the data, each altering the limits for the velocity RMS values provided by ISO 10816 by a scaling factor. Limits for the RMS of velocity vibrations given by NASA for newly commissioned machines were shown to be the most useful. This combination of limits provided better information about the state of the machine and a more detailed insight as to fault modes appearing in different octave bands [13].

42.3 Theory

Vibration monitoring has been found to be a proven method of determining a variety of fault modes within pumps. As a result, an increasing number of pump manufacturers include on-board sensors to aid the end user in diagnosing pump problems. The International Organization of Standards (ISO) published standard 10816 that addresses acceptable levels of RMS velocity vibration for centrifugal pumps. Four different zones (A, B, C, and D) are used to categorize the machine as being either a newly commissioned machine, an acceptable machine for long-term continuous operation, a machine suitable for short term but not for long-term use, or a machine that is considered inoperable. Limits for each zone are dependent on characteristics of the machine, such as the pump's power rating, location of the motor, height of the motor shaft, and the pump's orientation [14, 15]. Although very robust in determining the health of the machine, the single RMS velocity value lacks the ability to differentiate between fault modes.

Octave band spectral analysis was then used to expand on the information contained within the frequency spectrum that is hidden by the single RMS velocity value. Octave band spectral analysis has been used in the field of acoustics, where standards divide the frequency spectrum into several bands. ISO 532 defines the standard octave center frequencies of 31.5, 63, 125, 250, 500, 1,000, 2,000, 4,000, 8,000, and 16,000 Hz. Lower limits of a band are found by dividing the central frequency by the square root of 2, and the upper limits are found by multiplying the central frequency by the square root of 2. Anything found outside of these bands are considered not important since the audible range of frequencies for humans is from 20 Hz to 20 kHz [16]. McKee et al. adapted the octave band spectral analysis for RMS velocity values by altering the central frequency of the second band to be the running speed of the machine, and altering the first band to contain all frequencies below the lower limit of the second band. By doing so, each octave band contains information about vibrations surrounding specific harmonics of the machine [13]. This approach was utilized to create octave bands in the frequency spectrum based from the accelerometer readings on the centrifugal pump.

Since the number of octave bands can become numerous, depending on the sampling frequency of the accelerometers, principal component analysis (PCA) was then utilized to determine which variables are greatly affected by the onset of cavitation. The purpose of PCA is to reduce a set of observed variables into a smaller set of artificial variables that are a weighted sum of the observed variables. These artificial variables are called principal components. The first principal component accounts for the maximum amount of variance in the observed variables, thus producing an artificial variable that could possibly show the maximum change in the set of observed variables due to cavitation. The subsequent principal components account for the maximum amount of variance under the constraint that the component must be orthogonal to the preceding components [17].

42.4 Results

Data were acquired from pump tests performed by Melinda Hodkiewicz on a Gould's 3700, 90 kW, four vane impeller, centrifugal pump which ran under at a nominal speed of 2,990 rpm. Hodkiewicz recorded signals from four accelerometers

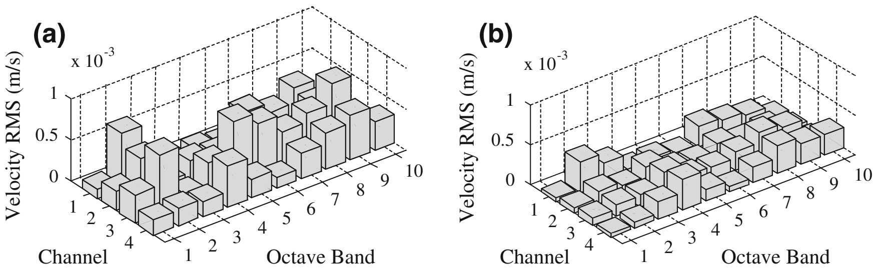


Fig. 42.1 Velocity RMS before and after cavitation (set 1) **a** Overall RMS Velocity Values (*baseline*): Ch1: 0.61 mm/s, Ch2: 0.44 mm/s, Ch3: 0.70 mm/s, Ch4: 0.71 mm/s, and **b** Overall RMS Velocity Values (*cavitation*): Ch1: 0.96 mm/s, Ch2: 0.88 mm/s, Ch3: 1.7 mm/s, and Ch4: 1.1 mm/s

and dynamic pressure transducers while modifying the pump to cavitate using a suction vacuum under a range of operating conditions on a purpose built test facility. The data were then analyzed to determine whether any indicators exist in the processed data as it approached a cavitating state. The measurements were divided into five datasets, where each set contained from 2 to 5 entries showing the progression from a non-cavitating state to an incipient through to a fully cavitating state. Each dataset has four channels—channels 1 and 2 were taken at perpendicular positions on the bearing closest to the pump's volute, channel 3 was located at the suction flange, and channel 4 was located at the discharge flange [18]. Statistics obtained from the data included RMS of the velocity signal, peak of the acceleration signal divided by RMS of the acceleration signal, RMS of the acceleration signal, kurtosis of the velocity and acceleration signal, severity levels of the RMS velocity vibration according to ISO 10816, and the severity levels of the velocity vibration in octave bands while applying limits given by ISO 10816, as performed by McKee et al. [13]. Baseline data were provided prior to any experiments being performed. Figure 42.1 shows the velocity RMS for each octave band for four channels for one of the sets.

Comparing statistical parameters between baseline data and data containing cavitation, several parameters were chosen that increased as the pump's state approached cavitation. The parameters chosen to be investigated were velocity RMS and the velocity RMS values found in octave bands. Since McKee et al. [13] performed a study on the advantage of velocity RMS values in octave bands over the overall velocity RMS values, the former was investigated to see whether signs of cavitation appeared. To do so, the difference between the values of the velocity RMS in the octave bands at cavitation and those of baseline data was obtained. Figure 42.2 shows the results for data obtained in Set 1 in terms of a gray-scale color scheme. The color scheme ranged from white, signifying small differences in value, to black, signifying large differences in value. Figure 42.2 shows that channel 3, the suction flange, has the largest difference between cavitation and the baseline data. In addition, octave bands 2, 5, 6, and 9 also have large differences

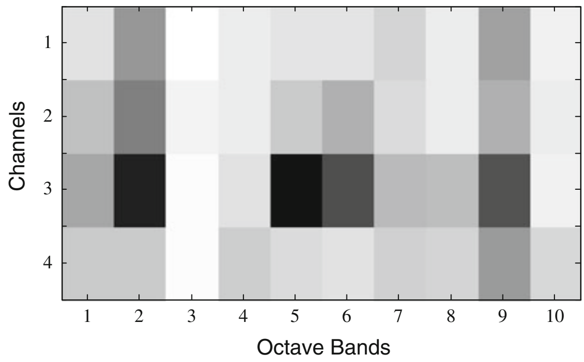


Fig. 42.2 Difference between baseline and cavitation data

Table 42.1 Values for principal component in all octave bands

	Channel 1	Channel 2	Channel 3	Channel 4
Band 1	0.0532	*0.3870*	0.1563	0.1211
Band 2	*0.1847*	*0.8520*	*0.6701*	*0.3118*
Band 3	0.1113	0.0136	*0.2235*	*0.4671*
Band 4	*0.1491*	0.0622	*0.2182*	*0.6578*
Band 5	*0.1334*	0.0215	0.1503	0.0228
Band 6	0.0804	0.0389	*0.1761*	0.0510
Band 7	0.0745	0.0232	0.1033	0.2127
Band 8	*0.6641*	*0.0676*	*0.6016*	*0.2767*
Band 9	*0.6710*	*0.3292*	0.0404	*0.2810*
Band 10	0.0856	*0.0688*	0.0328	0.1827

between cavitation and baseline data. Through consideration of all given datasets, it was found that octave bands 2, 4, 5, 6, 8, and 9 have been shown to increase as cavitation sets in.

To provide a quantifiable justification for choosing these octave bands to track, principal component analysis (PCA) was performed on the cavitation data in each set after the baseline data was subtracted. Hence, by looking at the weights given in the first principal component, the observed variables that produce the largest variance, and thus most greatly affected by cavitation, can be determined. Each data file used in the cavitation analysis contained vibration measurements sampled at 25 kHz. The results are shown in Table 42.1.

Entries in the table were highlighted whether they fell into the upper half of the values for the column. Then, looking across the octave bands, and using the criteria of taking the three bands that have the most highlights, it was seen that bands 2, 8, and 9 are to be used. Band 9 was chosen over band 4 since the average of the absolute value of band 9 was greater than that of band 4, thus representing a greater influence in the first principal component. This coincides with the literature, where cavitation is shown to appear in the vibrations surrounding the running speed, as well as in the high-frequency components.

Table 42.2 Average coefficients for bands from PCA

Band 2	Band 8	Band 9
0.5047	0.4025	0.3304

Table 42.3 CSP values for three runs with cavitation (mm/s)

Channel 1	Channel 2	Channel 3	Channel 4
0.5686	0.4376	0.8181	0.4580
0.4721	0.2213	0.4472	0.5452
0.5703	0.2259	0.6242	0.3863

To determine a cavitation sensitivity indicator from Table 42.1, an average value of the principal component coefficients were taken for each band. Table 42.2 contains these values for the chosen bands.

These values are then used as coefficients in an equation to create a cavitation sensitivity parameter (CSP), as shown in Eq. 42.1 below.

$$\begin{aligned} \text{CSP} = (0.50 \times \text{RMS Vel in Band } 2) + (0.40 \times \text{RMS Vel in Band } 8) \\ + (0.33 \times \text{RMS Vel in Band } 9) \end{aligned} \tag{42.1}$$

The next step is to set a threshold value for the CSP. This is done by applying the equation to datasets containing cavitation, and looking for an average CSP value. Below in Table 42.3 shows the results of applying the CSP to the data used in Table 42.1, in terms of mm/s. More datasets containing accelerometer readings from cavitation will be needed in order to confirm these threshold values.

42.5 Conclusion

Cavitation is one of the major problems resulting in degraded health of a centrifugal pump. Its appearance is a result of the pump not being operated at the ideal conditions, which are at or around the best efficiency point. Due to its damaging effect on a pump, it is imperative that pumps are diagnosed with this problem upon its onset. ISO standard 10816 sets limits on broadband RMS velocity values that can be used to assess the health of a pump when monitored over a long period of time. However, these severity limits are unable to discriminate between the various fault modes that are associated with the centrifugal pump.

Octave band spectral analysis has been shown to be able to help discriminate between various fault modes within the pump since narrow spectral bands show information surrounding specific multiples of the running speed. Due to the possible large number of Octave bands that could result based on the sampling frequency of the accelerometers, it is not ideal to follow all Octave bands. What is needed is a more focused view of the information in the Octave bands. PCA provides the ability to determine which Octave bands produce the most variance under cavitation as well as a set of weights corresponding to each Octave band which are used to produce the principal components. Taking the weights of the three Octave bands that contain the highest variance due to cavitation in the first principal component, a cavitation sensitivity parameter was created to help determine the severity of cavitation in the centrifugal pump. It was then applied to data from independent cavitation trials to determine threshold values indicating the onset of cavitation. Further investigation is needed to determine a robust threshold value that would best indicate the inception of cavitation across various pumps and under various conditions. In addition, investigation is needed to determine whether there is an ideal location for an accelerometer on a centrifugal pump to obtain the vibration readings needed to detect cavitation.

Acknowledgments This research was conducted within the CRC for Infrastructure and Engineering Asset Management, established, and supported under the Australian Government's Cooperative Research Centres Programme. The authors would also like to thank Professor Melinda Hodkiewicz for facilitating the vibration measurements.

References

1. McKee K, Forbes G, Mazhar I, Entwistle R, Howard I (2011) A review of machinery diagnostics and prognostics implemented on a centrifugal pump. In: Jay Lee JN, Mathew J (eds) Proceedings of the 6th world congress on engineering asset management
2. Čudina M (2003) Detection of cavitation phenomenon in a centrifugal pump using audible sound. Mech Syst Signal Process 17:1335–1347
3. Černetič J, Čudina M (2011) Estimating uncertainty of measurements for cavitation detection in a centrifugal pump. Measurement 44:1293–1299
4. Čudina M, Prezelj J (2009) Detection of cavitation in operation of kinetic pumps. Use of discrete frequency tone in audible spectra. Appl Acoust 70:540–546
5. Neil GD, Reuben RL, Sandford PM, Brown ER, Steel JA (1997) Detection of incipient cavitation in pump using acoustic emission. Proc Inst Mech Eng 211:267–277
6. Alfayez L, Mba D, Dyson G (2005) The application of acoustic emission for detecting incipient cavitation and the best efficiency point of a 60 kW centrifugal pump: case study. NDT and E Int 38:354–358
7. Athavale MM, Li HY, Jiang Y, Singhal AK (2002) Application of the full cavitation model to pumps and inducers. Int J Rotating Mach 8:45–56
8. Hofmann MaS B, Coutier-Delgosha O, Fortes-Patella R, Reboud JL (2001) Experimental and numerical studies on a centrifugal pump with 2D-curved blades in cavitating condition. In: CAV 2001: Fourth international symposium on cavitation. California Institute of Technology, Pasadena, CA USA, 20–23 Jun 2001
9. Uchiyama T (1998) Numerical simulation of cavitating flow using the upstream finite element method. Appl Math Model 22:235–250
10. Franz R, Acosta AJ, Brennen CE, Caughey TK (1990) The rotordynamic forces on a centrifugal pump impeller in the presence of cavitation. Trans ASME: J Fluids Eng 112:264–271

11. Rapposelli E, Cervone A, d'Agostino L (2002) A new cavitating pump rotordynamic test facility. In: 38th AIAA/ASME/SAE/ASEE joint propulsion conference and exhibit, Indianapolis, Indiana
12. Jensen J, Dayton K (2000) Detecting cavitation in centrifugal pumps. Orbit, Nevada
13. McKee KK, Forbes G, Mazhar I, Entwistle R, Howard I (2012) Modification of the ISO-10816 centrifugal pump vibration severity charts for use with Octave band spectral measurements. In Kotousov A (ed) 7th Australasian congress on applied mechanics, Adelaide
14. Standardization IOf (1998) ISO 10816-3: Mechanical vibration—evaluation of machine vibration by measurements on non-rotating parts—in Part 3: industrial machines with nominal power above 15 kW and nominal speeds between 120 r/min and 15 000 r/min when measured in situ, ISO, Switzerland
15. Standardization IOf (2009) ISO 10816-7: Mechanical vibration—evaluation of machine vibration by measurements on non-rotating parts. In: Part 7: rotordynamic pumps for industrial applications, including measurements on rotating shafts, ISO, Switzerland
16. Standardization IOf (1975) ISO 532: method for calculating loudness level. International Organization for Standardization, Switzerland
17. Wold S, Esbensen K, Geladi P (1987) Principal component analysis. Chemometr Intell Lab Syst 2:37–52
18. Blagrove L (2003) Cavitation detection in centrifugal pumps using vibration signals and fractal analysis. School of Mechanical and Chemical Engineering, University of Western Australia

Chapter 43
A Case Study on Condition Monitoring Based on Statistical Feature for Coil Break on Tandem Cold Rolling

Jun-Seok Oh and Hack-Eun Kim

Abstract Steel markets are very competitive and demand greater gauge precision and higher production rates. These growing requirements result in tandem rolling mill, which is of substantial interest to the steel industry, in order to improve quality and productivity. In such an environment, it is important to construct appropriate condition monitoring, which can lead to achieving the highest economic efficiency and avoiding equipment damage. This paper proposes a comprehensive condition monitoring methodology based on statistical feature extraction technique to increase the efficiency of feature extraction from high-dimensional feature space. It is examined that one can explore easily the effective features by using three-dimensional feature space for the condition monitoring. The method has been applied on condition monitoring of the stationary rolling in steel industry.

43.1 Introduction

Cold rolling is a process by which the sheet metal or strip stock is introduced between rollers and then compressed and squeezed. The thickness accuracy of cold-rolled products is determined by rolling force and strip tension to minimize thickness deviations during rolling process. A tandem rolling mill consists of a set of rolling mill stands. In order for efficient production, the strip is rolled on a continuous production line, passing from one mill station to another at high speed without stopping. Compared to a single mill, tandem rolling mill has high efficiency, good quality, and high automatic degree, so that it is suited for production with high strength. But tandem cold rolling mill has a very complex control

J.-S. Oh (✉) · H.-E. Kim
Korea Gas Technology Corporation, 839, Bongsan, Yuseong, Daejeon, Korea
e-mail: jsoh@kogas-tech.co.kr

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_43,

problems and complex correlation with undesired mechanical vibration. A limited method which is based on the time series or frequency of vibration has been applied to the condition monitoring of a tandem rolling mill [1–3]. Several studies have been performed to identify the limited problems such as third octave, fifth octave vibration, and vertical chatter vibration [4–6]. When considering the industrial importance, the studies of various aspects are required in rolling mill and its product. This paper proposes an approach to the development of a feature-based method [7, 8] on coil break of tandem rolling mill. The objective of this study is to compare the usability of various signals and extract the useful information among the signal features originated from rolling force, tension, strip thickness, temperature, motor current, and vibration. In order to remove existing irrelevant or redundant features, feature selection is considered, which often results in improved performance, especially when small dimensions are required [9–11]. This method can help identify the trend of selected feature in 3D features space [12–14]. The implementation of a three-dimensional visualization of features allows one to see the trend of three different features, which can help viewers digest monitoring data efficiently.

43.2 Tandem Cold Rolling Mill and Coil Break

The discussion in this paper focuses on a five-stand tandem cold rolling mill. Each mill stand has two work rolls and four backup rolls about two feet in diameter, which are driven by powerful electric motors. Also known as six-Hi arrangement, the rolls in contact with the strip are the work rolls and four backup rolls prevent the work rolls from being distorted by the force. The five-stand is operated at very high strip tension to reduce the vertical force in the roll and to remove any shape defects between stands. Constant speed and roll gap between each stands are maintained, and the volume rates of strip of rolls must be identical. Significant tension is maintained even between the last stand (Fig. 43.1).

A coil break may result in an unacceptable surface finish on parts, which are small creases present across some portions of the width of the coil. Sometimes, there are severe rolling accidents, due to weak rolled materials with internal defects. The defects of material surface can cause coil breaks accompanied by mill damage that can affect the rolled product. A coil break which occurred during steady-state operating speed is difficult to repair and requires longtime maintenance, which can lower productivity and generate higher cost. In addition, due to increasing production volumes and needs for various material conditions, coil break is frequently occurs during rolling process. To prevent any outage and avoid critical accidents, condition monitoring technique is required.

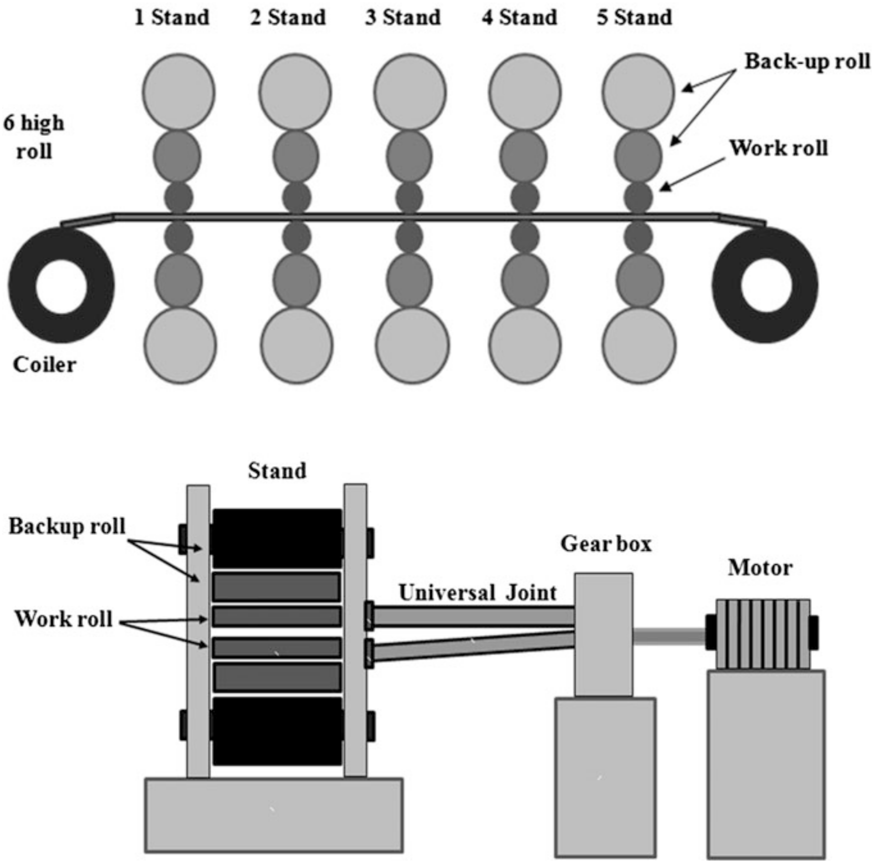


Fig. 43.1 Five-stand tandem cold rolling mill

43.3 The Proposed Condition Monitoring Method

Basically, rolling force and current consumption in motor are different depending on thickness and various material conditions. From pass to pass between each stand, the strip is reduced in thickness with decreasing rolling force. When the rolling mill is under fault condition, each stand cannot supply sufficient strip tension and maintain constant rolling force, which can change current consumption and lead to a fluctuation of motor current. Figure 43.2 shows current signal between normal and fault condition during rolling mill process. From start-up to shutdown, normal condition is shown as stable, while fault condition shows significant change of current signal.

In addition, the vibration of tandem rolling mill is comprehensive and common phenomenon. Vibration analysis can expose many mechanical and electrical problems. Further analysis of the vibration spectra is used to separate various

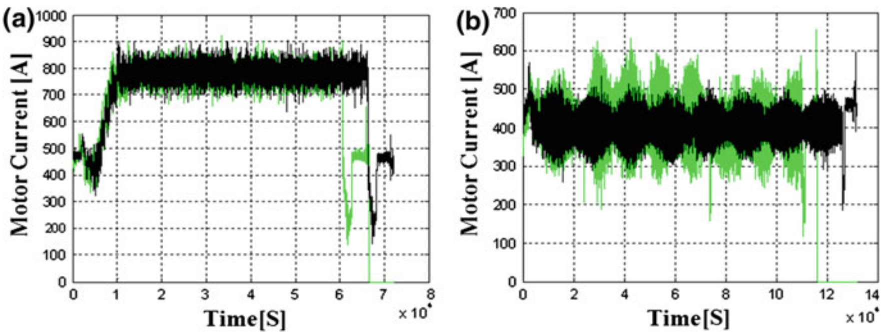


Fig. 43.2 Motor current signal at no. 4 stand; **a** Normal condition and **b** Fault condition

specific faults. Vibration research on tandem rolling mill is of great practical significance. Mechanical defects such as third octave, fifth octave, and chatter can be identified through vibration, which has been published in several literatures [4–6].

In this study, rolling force, tension, strip thickness, temperature, vibration, and motor current signals are measured from each stands of target machine. Too many signals make it difficult for the operating manager to clearly identify the machine condition. In order to reduce high dimensionality and extract the useful information from those signals, this paper addresses the application of statistical feature-based technique using measured signal. The process is summarized as below:

Step1. Measuring current from drive motor and vibration from upper and lower bearing of work roll at steady-state rolling speed.
Step2. Statistical feature calculation from measured raw data using Table 43.1.
Step3. Normalization from calculated feature vectors (feature normalization).
Step4. Select the features based on the squared Euclidean distance as described below.
Step5. Built up three-dimensional feature space for condition monitoring (Fig. 43.3).

Feature selection criteria

- Find the minimal feature subset that is necessary and sufficient to decide the classes.
- The same feature vectors should be constant value over time.
- Between the different states of the feature, vectors must have a clear distinction.
- Select a subset of M features from a set of N features, M < N, such that the value of a criterion function is optimized over all subsets of size M.

In this study, the squared Euclidean distance is used as an objective function to the selection of the most desirable feature value.

Table 43.1 Used features for feature calculation

Domain features	
	Mean
	RMS
	Standard deviation
Time (7)	Skewness
	Kurtosis
	Shape factor
	Crest factor
Frequency (5)	Frequency
	Mean square frequency
	Root mean square frequency
	Variance frequency
Entropy (10)	Entropy estimation
	Entropy estimation error
	Histogram lower bound
	Histogram upper bound
	Auto-regression coefficients (6)

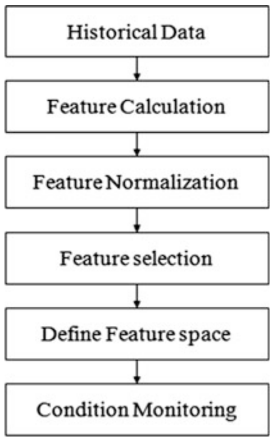


Fig. 43.3 Procedure of condition monitoring

$$d(x, y) = ||x - y|| = \sqrt{(x - y)^T (x - y)} \quad (43.1)$$

43.4 Result and Discussion

As described in Fig. 43.4, where X-axis is a number of features from stand 1 to 5 of rolling mill and Y-axis is feature value after normalization. Based on feature selection criteria, the dimension is reduced by removing irrelevant features from

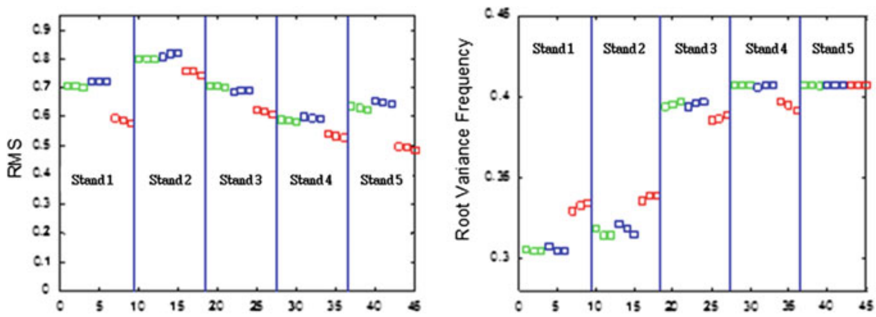


Fig. 43.4 Result of feature selection

original feature vectors. Dimensionality reduction through feature selection is important to highlight salient features of a machine's conditions and improving accuracy of distinguishable class. Through this procedure, 45 of the most important features are found, instead of the entire 110 feature set. Figure 43.4 descried the result of RMS and root variance frequency. Selected features has shown clear distinction between different states (green/blue: before coil break, red: at coil break).

Each reduced feature subset was then used to make feature space at three axes, respectively. Figure 43.5 shows where features are clustered in the reign of the three-axis feature space depending on what features are selected. In these results, selected features of motor current and vibration from lower roll are distinguished clearly between normal and coil break condition, while features of roll force, thickness deviation, and tension force are clustered in the same area or very close between different conditions.

Therefore, vibration and motor current are selected as fault parameters to distinguish a fault from normal condition effectively. It can be seen that RMS, entropy estimation, and root variance frequency of motor current have significant distance between two conditions, while kurtosis, crest factor, and entropy estimation of vibration signal contains significant difference between a normal and a coil break condition.

Figure 43.6 is the result of monitoring at stand 3 to 5 using motor current. Normal and fault conditions are clustered in specific area, respectively. In addition, differences are gradually increased over the stand from 3 to 5. Especially, this figure shows clear differences in the distance at stand 5 because physical characteristics under the rolling mill are showing that upper roll is more unstable than lower roll which is in condition of relatively higher loading, as it does not maintain constant rolling force on upper roll. Therefore, measured vibration from upper roll is very important to know corresponding movement of strip.

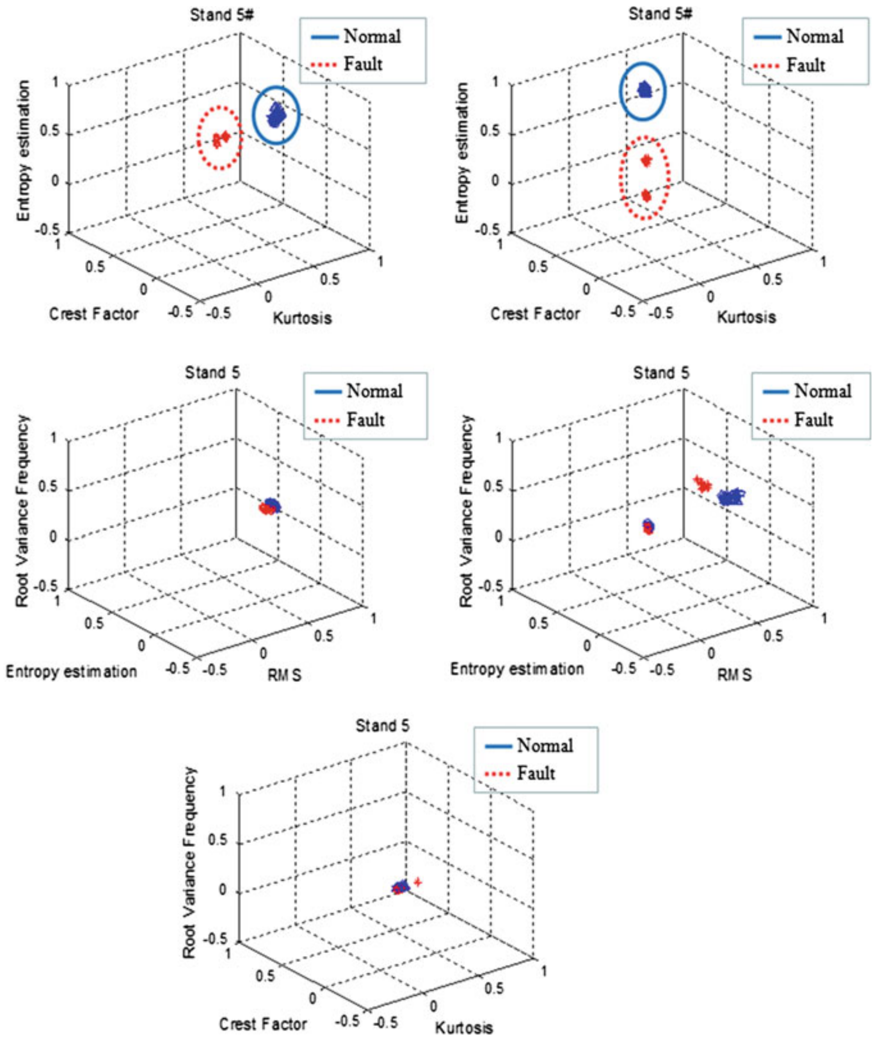


Fig. 43.5 Motor current signal and vibration

This approach importantly helps the viewers to more efficiently explore feature set and reduce complexity during condition monitoring. The three-dimensional view is of great importance to help user visually depict the trend tracking of features and understanding of machine condition effectively.

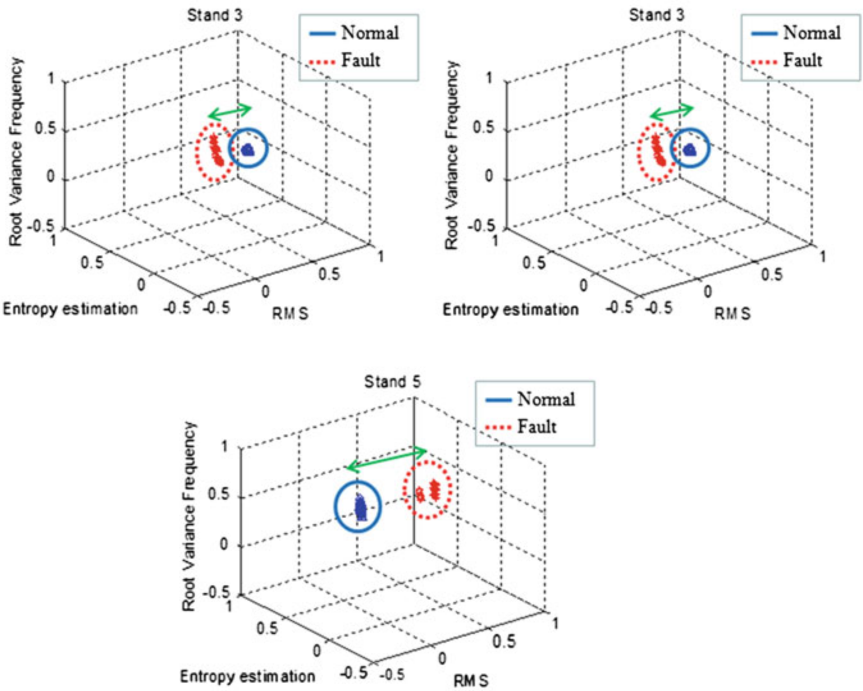


Fig. 43.6 Result of motor current

43.5 Conclusions

This paper presents a study on condition monitoring of coil break in multi-stand mill using motor current and vibration signal. In order to determine relevant features and reduce dimensionality, feature selection and three dimension of feature space are hired. Three-dimensional technique can lead to data visualization, which is a powerful way to simplify complexity and improve efficiency of explore features during condition monitoring. Our methodology proved that RMS, entropy estimation, root variance frequency of motor current and kurtosis, crest factor, and entropy estimation of vibration clearly show lager distance between normal and faulty conditions. As a result of this study, these features can be considered as a relevant feature for condition monitoring of coil break on rolling mill. This study reveals that the proposed statistical feature-based technique is useful as it has the advantage of condition monitoring using low-dimensional feature representations. Therefore, the proposed method of using feature-based condition monitoring is recommended and making it more suitable in various real applications.

References

1. Asch A, Hohn W (1999) Monitoring system for roll stand drives using strain gage technology. In: Proceedings of international federation of automatic control: automation in mining, mineral, and metal processing, pp 159–164
2. Cerv H, Kefller H, Luckmann F, Mackel J (1998) Mill diagnostic system (MiDaS)—a monitoring system with quality and maintenance-related diagnostic functions. Aluminium 74(3):148–154
3. Vinod DT, Jayant PM, Girish DM (2011) Vibration based condition monitoring of rolling mill. Int J Sci Eng Res 2(12):2229–5518
4. Donkle LB (1999) Fifth-octave chatter problem solved using vibration analysis. AISE Steel Technol 76(11):40–45
5. Mohammad RN, Mohammad RF, Mahmud S, Mohsen K (2012) Frequency analysis of chatter vibrations in tandem rolling mills. J VibroEng 14(2):1392–8716
6. Kimmura Y, Sodani Y, Nishiura N, Ikeuchi N, Mihara Y (2003) Analysis of chatter in tandem cold rolling mills. ISIJ Int 43(1):77–84
7. Widodo A, Yang BS (2007) Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motors. Expert Syst Appl 33(1):241–250
8. Yang BS, Han T, Hwang WW (2005) Application of multi-class support vector machines for fault diagnosis of rotating machinery. J Mech Sci Technol 19(3):845–858
9. Srilatha C, Ajith A, Johnson PT (2004) Feature deduction and ensemble design of intrusion detection systems. Comput Secur 24(4):295–307
10. Mandokht M, Glenn F, Jennifer GD (2010) From transformation-based dimensionality reduction to feature selection, appearing. In: Proceedings of the 27 the international conference on machine learning
11. Zhu Y, Shan X, Guo J (2005) Feature selection method based on variable similarity in intrusion detection system. Microelectron Comput 22(10):34–36
12. Widodo A, Yang BS, Han T (2007) Combination of independent component analysis and multi-class support vector machines for fault diagnosis of induction motors. Expert Syst Appl 32(2):299–312
13. Widodo A, Kim EY, Son JD, Yang BS, Tan CC, Gu DS, Choi BK, Mathew J (2009) Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine. Expert Syst Appl 36(3):7252–7261
14. Hwang WW, Yang BS (2004) Fault diagnosis of rotating machinery using multi-class support vector machines. Trans Korean Soc Noise Vibr Eng 14(12):1233–1240

Chapter 44
Embedded Systems Solutions for Fault Detection and Prediction in Electrical Valves

L. B. Piccoli, R. V. B. Henriques, E. Fabres, E. L. Schneider and C. E. Pereira

Abstract This paper proposes an embedded system architecture for fault detection and prediction in electrical actuators used in pipelines for oil and gas transportation. The proposed system incorporates a signal processing flow that requires low complex mathematical operations using ANSI-C language. However, when described in the hardware description language (HDL), it can be implemented in dedicated field-programmable gate array (FPGA) or ASIC. To prove its functionalities, a test bench was developed, which aims to reproduce in a laboratory some common faults and degradation processes that may occur in real-world field applications. A data acquisition equipment was used to collect the sensors information from specific points of the actuator. The sensor data collected and simulation were used to validate the propose fault detection methodology.

44.1 Introduction

In critical systems, adopting proactive maintenance strategies such as the monitoring, prognosis, and diagnosis together in an embedded system is of great importance. An example is petroleum transportation lines, in which one interruption or bad operation results in potential danger to the ecosystem and also to human lives that may cause damage to the environment or economic loss.

In general, the oil and gas industry activities are attached to reserves exploration, production, oil refining, and distribution [1]. In the refining stage, the hydrocarbons that form oil are separated, giving rise to distinct products (petroleum subproducts). The distribution refers to the derivatives transportation activities of petroleum until the sale points. The maintenance procedures in tanks, ships,

L. B. Piccoli (✉) · R. V. B. Henriques · E. Fabres · E. L. Schneider · C. E. Pereira
UFRGS—Control Automation and Robotics Group, Av. Osvaldo Aranha,
103 Porto Alegre-RS, Brazil
e-mail: leonardo@piccoli.eng.br

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_44,

and other kinds of contention systems usually demand the coating and control of the inner walls of these containers to avoid leakage, corrosion, and fire. Currently, the most maintenance activities are accomplished by the Petrobrás hat which includes corrective, preventive, and predictive actions [2].

Nowadays, we employ mainly two kinds of maintenance: the corrective and the preventive. The purely corrective maintenance consists of repairing a machine or equipment after the breakage or failure detection (so a maintenance action occurs in response to failure, this alternative is also sometimes called "reactive maintenance"). In this case, there is no time or resources spent between consecutives failures. Therefore, the cost of maintenance becomes higher and there may be serious consequences, such as augmenting the risk to employees and the necessary forced shutdowns in the production process. A preventive maintenance considers a certain level of machine degradation, usually based on statistics of previous similar failures occurred, and it requires a periodic maintenance. This practice can replace or repair equipment or pieces that could be in place for longer operation, but it does not exclude unexpected downtime caused by unpredictable failure. In both cases, the costs related can be high and also the downtime can be long.

In the case of a failure that could be detected proactively, the system could automatically make an auto reconfiguration in order to sustain operation, although in a degraded way, and by the same time issue an alert or request for maintenance. This reconfiguration feature allows, for example, engine torque or power limitation with the intention of reducing the strength on the components and thus prolonging useful lifetime of the system [3].

A proactive maintenance strategy requires an automation system consisting of following parts: sensors set, sensors data acquisition interface, signal processing unit, and an expert decision block. Since the operation of the system occurs in real time, all digital processing calculations must occur during the time of a sampling period in the data acquisition device. In this sense, the choice of algorithms such as discrete wavelet transform (DWT) and adaptive filter algorithm with least mean squares (LMS) that have low computational complexity and high performance [4] represents a breakthrough.

A test bench was designed and built to validate the proposed technique in real-world field application failure. This way, it was installed a brake disk system in a commercial fluid flow control valve system to simulate the open/close mechanical efforts. This approach makes possible to monitor the torque effort delivered by the actuator through the gear mechanism during driving operations for opening and closing the valve. This approach emulates typical initial failure situations on the gears, before any catastrophic faults occur.

It was used in the test setup a CSR6 model actuator from Coester company [5]. The valve actuator received a torque, and vibration sensors were installed at specific points to collect data at normal operating conditions and at failure as well. The torque signal was acquired for different operating conditions with variation on the dynamic behavior of the system under test, and it was applied to an embedded system with an algorithm implemented in ANSI-C language, employing digital signal processing techniques.

The study of failures in electric actuators has been developed by means of sensor fusion, correlating signals of torque and vibration, as well as through the use of self-organizing maps [6].

44.2 The Embedded System

The control, supervision, and automation in the processes of transport of petroleum with electric actuators require a computer system that is able to detect the equipment availability during operation. In dynamic operating situations, the signal behavior extracted from the actuator system is not known in advance and it behaves as a stochastic process. Thus, the system model cannot be obtained analytically. So, in this work is proposed a solution to the recognition of predictable failures by using digital signal processing techniques.

One approach possible is to use the wavelet transform. The advantages of the wavelet transform over traditional transforms, such as the Fourier transform, are already very well known [7]. Because of its strait representation in time and frequency domains, it is widely used for linear signal analysis. However, in environments with nonlinear and time-variant systems, the use of adaptive filters has become more attractive.

An LMS adaptive filter demands only multiply–add operations to be implemented, which facilitates their usage in embedded systems. Moreover, the mathematical complexity is approximately the same for a finite impulse response (FIR) filter. As FIR filters do not have feedback, the stability between input and output of a FIR filter is provided for any set of fixed coefficients and their implementation together with the LMS algorithm is relatively simple.

Regarding the open systems architecture for condition-based maintenance (OSA-CBM) model [8], this approach will employ the first four layers of the model: (1) data acquisition; (2) data manipulation; (3) state detection; and (4) health assessment. The system comprises the analysis of the signal properties in both the time and frequency domains simultaneously, characteristics extraction, and adaptive filtering.

44.2.1 Discrete Wavelet Transform

The wavelet transform of an arbitrary function $f(t)$ is defined by the convolution:

$$\omega(s,\tau) = \int_{-\infty}^{\infty} f(t) * \psi^{s,\tau}(t)\mathrm{d}t \tag{44.1}$$

where the wavelet mother function $\psi^{s,\tau}$ is dilated with a scale parameter s and translated in its domain by τ.

The DWT of a continuous function *f*(*t*) is defined as the discretization of the values sampled uniformly k in periods for the τ location of fixed scale in of s dyadic sequence 2^j [9]. It is known that the dilatation is multiplicative. Thus, the 2^j values with scales greater than 1 remove the details and produce low-frequency information, while values with scales smaller than 1 produce high-frequency information.

Then, the wavelet mother function ψ and the scaling function φ can be expressed as follows:

$$\psi(2^{j-1}, t) = \sum g_j(k) * \overline{\varphi}(2^j t - k) \tag{44.2}$$

and

$$\phi(2^{j-1}, t) = \sum h_j(k) * \overline{\varphi}(2^j t - k) \tag{44.3}$$

The coefficients of the function ψ are calculated from coefficients of the function φ as follows:

$$g_n = (-1)^n h_{L-1-n} \tag{44.4}$$

For $n = 0\ldots\ L - 1$
where L is the total number of coefficients. The set of coefficients h_j and g_i define the scaling functions of wavelet.

The analysis of wavelet consists in approximation and detail coefficients. It can be implemented from different orthogonal bases constructions [9]. The approximation represents the low-frequency components of the signal, and the details are its high-frequency components.

44.2.2 Extraction of Energy

Since Parseval equality condition exists for the wavelet transform [10], the total energy in discrete-time signals $x(n)$ for a window-N a N is determined by the following expression:

$$\sum_{n=-N}^{N} |x(n)|^2 \tag{44.5}$$

where N is the number of samples.

For periodic signals, it is more convenient to compute the average energy or power:

$$\frac{1}{(2N+1)}\sum_{n=-N}^{N}|x(n)|^2 \tag{44.6}$$

44.2.3 LMS Adaptive Filter

The implementation of the FIR filter in time domain only requires a few coefficients, and the filtering can be performed as a convolution of the input signal with these coefficients. This implementation is very efficient when the filter order is not so large. Due to this fact, in this work, the LMS algorithm was applied to this filter. The LMS method works in the time domain, and it uses a stochastic gradient algorithm to adapt the load of a filter.

Thus, the LMS algorithm works together with the FIR filter, adjusting the filter coefficients to minimize the error in the squares average sense. The coefficient μ (step size) affects directly how quickly the adaptive filter will converge toward the unknown system [11].

44.2.4 Proposed Signal Processing Flow

As mentioned in the previous subsections, this work is performed using the following signal processing steps: DTW, signal energy extraction, and a LMS adaptive filter, as shown in Fig. 44.1. The DWT domain (DWT-LMS) algorithm has a better performance because it has an improvement in convergence speed and a reduction in steady-state error [12]. Thus, this work proposes the development of an algorithm to be used in computers or microcontrollers-based systems to perform fault detection and diagnosis system.

The DWT block is responsible for the original signal decomposition into several components located in time and frequency domains. Depending on the desired result, the coefficient of approximation $l(n)$ or detail $h(n)$ is sent to the block **POWER**, where the signal energy is extracted. Then, an adaptive filter LMS is used to predict values of the input signal $s(n)$, which are the desired response delayed by a specific number of samples. The adaptation parameter step size μ controls the speed of convergence, stability, and steady-state performance of the adaptive filter algorithm and also the sensitivity to obtain a transient response variation in the time-domain system. It is important to note that in this application, the output signal (i.e., in other words, the steady-state error $e(n)$) represents the possible fault detection in the observed signal.

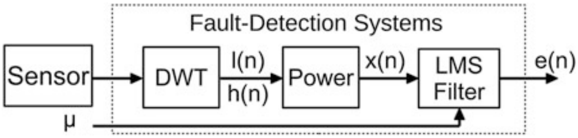


Fig. 44.1 Structure for the embedded system

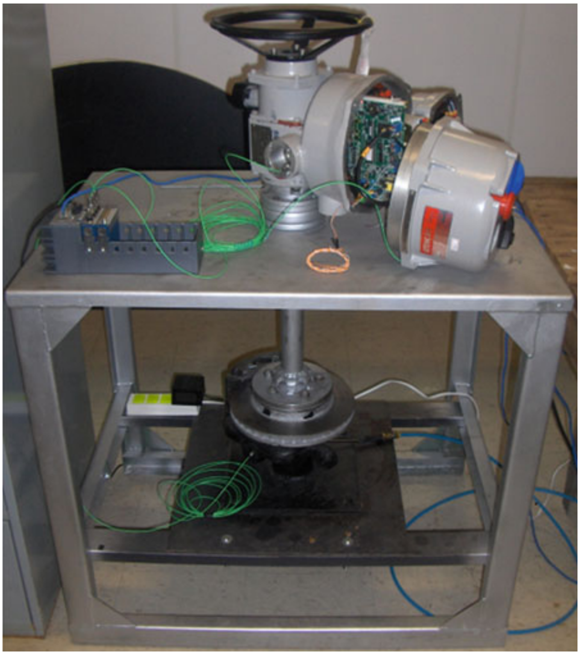


Fig. 44.2 The developed test bench

44.3 Description of Test Setup

The equipment CompactRIO [13] NI cRI0-9004 of National Instruments together with the software LABVIEW [13] was employed in the Control, Automation and Robotics Laboratory (LASCAR) for data acquisition and control of the opening and closing processes of the electric valve actuator (*Coester* Automation model CSR6), as shown in Fig. 44.2.

The test setup was used for the preparation of test cases emulating common behavior for the most common faults found in the real-world field application where the equipment becomes susceptible to the action of degradation such as aging, corrosion, cracks, and damages caused by operators.

It was used a disk brake in a test bench for fault injection. This is installed on the actuator stem, and it was driven and regulated by means of a pneumatic valve controlled by an auxiliary actuator, as shown in Fig. 44.3. Thus, it was possible to perform an emulation of possible efforts that the valve suffers according to the passage of fluid in the pipe.

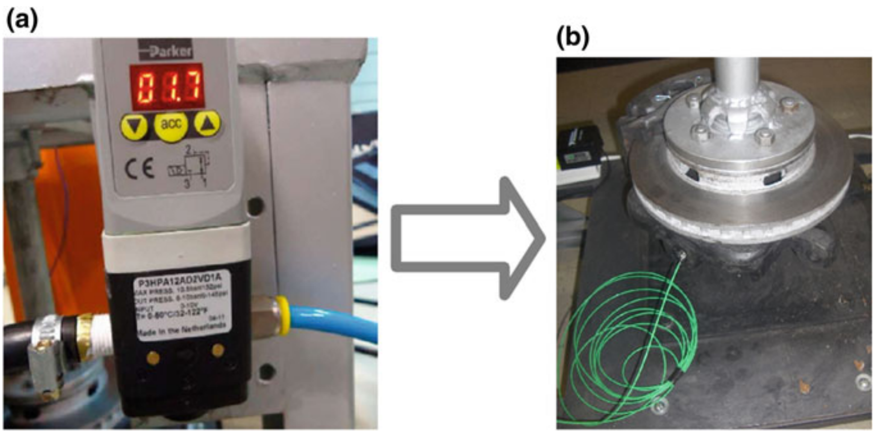


Fig. 44.3 Failure injection with disk brake. **a** Pneumatic valve, **b** disc brake

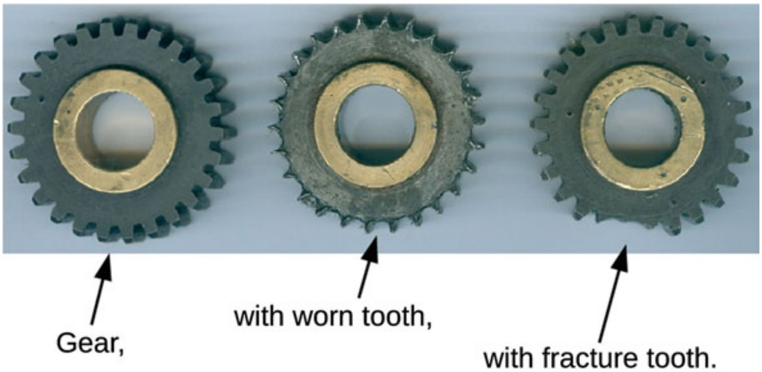


Fig. 44.4 Set of gears used in the tests

Figure 44.4 presents some possible gear aging conditions used to simulate several kinds of efforts and vibration delivered to the shaft depending on the gear health. Three different gears were employed, simulating field conditions for standard gear, aged gear, and fractured tooth gear.

For those all different operating conditions, sensor information was collected from the acquisition system using a vibration and torque sensor installed into the equipment as shown in Fig. 44.5, denominated as follows:

- **Sensor 1**: Torque on the motor;
- **Sensor 2**: Accelerometer on the motor.

In order to perform system tests, open/close test cycles of the valve were executed under normal and failure situations as it follows:

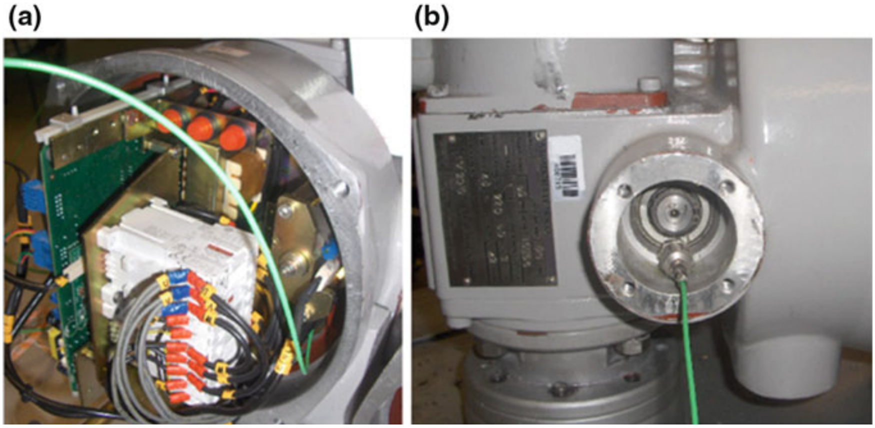


Fig. 44.5 Sensors on the engine compartment of the electric actuator. **a** Torque sensor, **b** accelerometer sensor

- **Normal**: Normal test cycle and without external effort on the system;
- **Fault type 1**: Test cycle with pressure equivalent to 3, 0 bar applied through the brake set;
- **Fault type 2**: Test cycle using two gears with aged tooth;
- **Fault type 3**: Performed using a test cycle of three gears with fracture tooth.

44.4 Results

For all test sequences, the following procedure was applied: These tests were initially performed a few seconds of normal system cycle without failure in opening and closing the valve, followed by a test cycle of specific failure. The data were collected according the proposed embedded flow. The value of adaptation parameter of LMS adaptive filter was adjusted to have a good coverage in the failure detection possibilities. The results with the information produced by the system were loaded into Scilab software [14] to produce the graphics contained in this article.

44.4.1 Test for Fault Type 1

In this test, sensor 1 was used as the input for the embedded system which executes the detection algorithm implemented in the ANSI-C language. Figure 44.6 shows the graphics generated by the tool Scilab [14], where the first signal presents the sensor value. The second signal shows the error, which is the response of the embedded system. In this case, the energy was extracted from the approximation

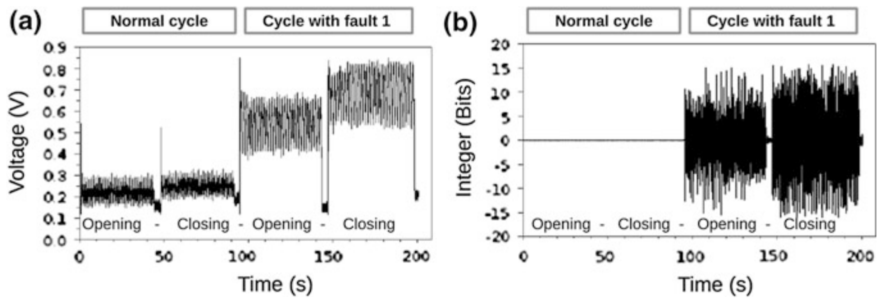


Fig. 44.6 Test with a sensor 1, fault type 1, approximation coefficient, and $\mu = 0.1$. **a** Input signal of the actuator torque valve, **b** output signal from the fault detection system

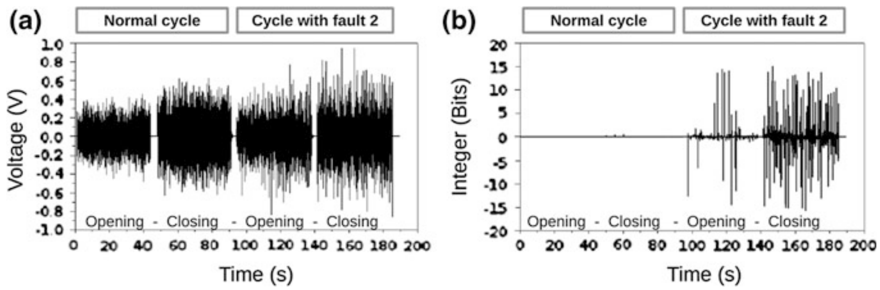


Fig. 44.7 Test with a sensor 2, fault type 2, detail coefficient, and $\mu = 0.05$. **a** Input signal of the actuator torque valve, **b** output signal from the fault detection system

coefficients of the input signal. It must be observed that the error was detected for this type of failure in both opening and closing valve processes. In this test results, the value used for parameter adaptation of the LMS adaptive filter was 0.1.

44.4.2 Test for Fault Type 2

Sensor 2 was used to generate the graphics of Fig. 44.7. Similar to the previous test, but now with the energy extracted from the detail coefficients signal, the error was also detected for this type of fault by the embedded system. In this case, the energy was extracted from the detail coefficients of the input signal.

44.4.3 Test for Fault Type 3

Similar to the previous test, the same sensor 2 was used and the same coefficient of the LMS adaptive filter and also the energy were extracted by the detail coefficients signal. The error also was detected for this type of fault by the embedded system, as shown in the graphic in Fig. 44.8.

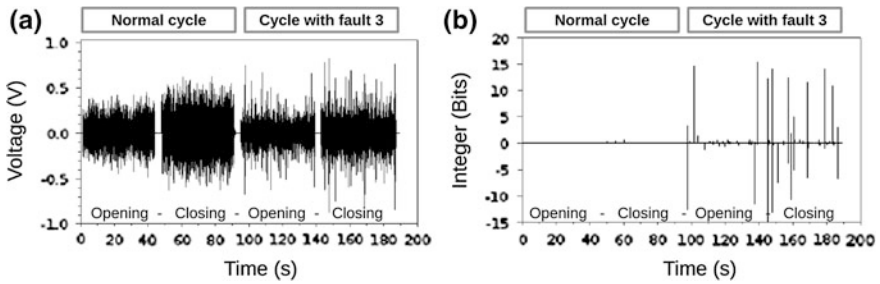


Fig. 44.8 Test with a sensor 2, fault type 3, detail coefficient, and $\mu = 0.05$. **a** Input signal of the actuator torque valve, **b** output signal from the fault detection system

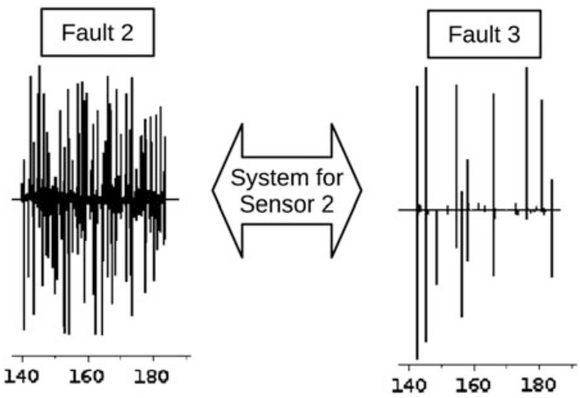


Fig. 44.9 Comparison of two faults on the system

44.4.4 Test Comparison

According to the characteristic of the error signal output by the system, it is possible to identify the probable cause of a system fault and its type. For example, in the graphic of Fig. 44.8, spikes occur intercalated in the signal, which characterizes the signature of the corresponding fault of the gear-fractured tooth. In the graphic of Fig. 44.7, there is a presence of noise, which can characterize a slack in the drive gear caused by aging. Figure 44.9 illustrates this situation by comparing these failures. Then, the operating state of the system can be determined and the equivalent output to the health index of the equipment monitored.

44.5 Summary and Conclusion

This work presented an embedded system description for fault detection using an instrumented test bench. It performed effort tests in the electric actuator coupled to a disk brake system and defective gears for fault injection in order to replicate conditions under field.

The embedded system in this work employs a technique using a computing device to perform analysis of the signal of a sensor by using signal processing and seeks the relation between the current and the future value, producing the fault detection in an iterative manner.

Furthermore, it may be implemented as a set of instructions executed by a general-purpose microprocessors or a digital signal processor (DSP). It can also be described using a hardware description language (HDL) such as VHDL or Verilog, to be applied as a set of logical operations in an field-programmable gate array (FPGA) or to generate a custom very large-scale integration (VLSI) chip.

The results were satisfactory, since all test conditions had their faults detected. It can be observed in the visualization of graphics that the sensors have different characteristics. Concerning the sensor 1, because the amplitude variation, it was decided to analyze the approximation coefficients. In the sensor 2, because the frequency variation, it was decided to analyze across the detail coefficients.

Therefore, in the efforts simulation, the use of a test bench together with the embedded system is of great importance to identify many kinds of fault situations and degradation as well the appearance of cracks and component aging. Moreover, it will be possible to validate the system with the data collected on field and to embed all the electronics on a chip.

References

1. Russo EER et al (2004) A realidade virtual na indústria de exploração e produção de petróleo, Editora Mania de Livro. Realidade Virtual: Conceitos e Tendências, São Paulo, pp 283–288
2. Figueiredo MG et al (2005) Reestruturação produtiva na bacia de campos—a terceirização e suas implicações para o trabalho petrolífero offshore, XXV Encontro Nacional de Engenharia de Produção, Porto Alegre, pp 2346–2353
3. Lee J (1996) Measurement of machine performance degradation using a neural network model. Comput Ind 30(3):193–209
4. Konezny M, Rao S (1995) Improving the DWT-LMS algorithm: boundary filter dwt matrix construction, signals, systems and computers. In: Conference record of the twenty-ninth Asilomar conference on, vol 1, pp 75–81
5. Coester (2001) Coester automação, manual Atuador—Linha CSR 12-CSR 25—CSR 50 integral
6. Architectures RL, Practices D (2009) CompactRIO developers guide, system
7. Gonçalves LF, Bosa JL, Balen TR, Lubaszewski MS, Schneider EL, Henriques RV (2011) Fault detection, diagnosis and prediction in electrical valves using self-organizing maps. J Electron Test 27(4):551–564
8. Carvajal REG (2011) Sobre técnicas para manutenção e diagnóstico inteligente de sistemas mecatrônicos. estudo de caso utilizando cálculo de ordem fracionária, Tese (doutorado), Universidade Estadual de Campinas. Faculdade de Engenharia Mecânica, Campinas, SP
9. Swearingen K, Majkowski W, Bruggeman B, Gilbertson D, Dunsdon J, Sykes B (2007) An open system architecture for condition based maintenance overview. In: Aerospace conference. IEEE, pp 1–8
10. Mallat S, Zhong S (1992) Characterization of signals from multiscale edges. IEEE Trans Pattern Anal Mach Intell 14(7):710–732

11. Walnut D (2002) An introduction to wavelet analysis, applied and numerical harmonic analysis, Birkhauser
12. Haykin S (2001) Adaptive filter theory, 4th edn. Prentice Hall, New Jersey
13. Murmu G, Nath R (2011) Convergence performance comparison of transform domain LMS adaptive filters for correlated signal. In: International conference on devices and communications (ICDeCom), pp 1–5
14. Sumathi S, Surekha P (2007) LabVIEW based advanced instrumentation systems. Springer, New York
15. Scilab Consortium (2011) Scilab: free and open source software for numerical computation. Scilab Consortium, Digiteo, Paris, France

Chapter 45
Application of Linear Prediction, Self-Adaptive Noise Cancellation, and Spectral Kurtosis in Identifying Natural Damage of Rolling Element Bearing in a Gearbox

C. Ruiz-Cárcel, E. Hernani-Ros, P. Chandra, Y. Cao and D. Mba

Abstract The ability to detect and diagnose faults in rolling element bearings is crucial for modern maintenance schemes. Several techniques have been developed to improve the ability of fault detection in bearings using vibration monitoring, especially in those cases where the vibration signal is contaminated by background noise. Linear prediction (LP) and self-adaptive noise cancellation (SANC) are techniques which can substantially improve the signal to noise ratio of a signal, improving the visibility of the important signal components in the frequency spectrum. Spectral kurtosis (SK) has been shown to improve bearing defect identification by focusing on the frequency band with higher level of impulsiveness. In this paper, the ability of these three methods to detect a bearing fault is compared, using vibrational data from a specially designed test rig that allowed fast natural degradation of the bearing. The results obtained show that the SK was able to detect an incipient fault in the outer race of the bearing much earlier than any other technique.

45.1 Introduction

Bearing faults are produced typically by the damage in the surface of the inner or outer race or in the rolling elements. When a damaged surface contacts another surface [1], a force impulse is generated, which excites resonances in the bearing and the machine. These successive impacts generate an impulsive vibrational

C. Ruiz-Cárcel (✉) · E. Hernani-Ros · P. Chandra · Y. Cao · D. Mba
Cranfield University, Cranfield, Bedfordshire MK430AL, UK
e-mail: c.ruizcarcel@cranfield.ac.uk

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_45,

signal which is easy to identify in the presence of low background noise. In a real machine, the background noise can sometimes mask the bearing fault components of the signal, especially in gearboxes because the gear meshing can generate a strong level of vibration [2]. For this reason, many different methodologies of signal processing have been developed in order to facilitate the detection of defects, particularly in bearings.

In this investigation, three diagnostic techniques [linear prediction (LP), self-adaptive noise cancellation (SANC) and spectral kurtosis (SK)] were applied in identifying a bearing defect in a gearbox where the bearing degradation was accelerated naturally in a specially designed test rig under conditions of relatively large background noise.

45.2 Theoretical Background

The estimation of a dynamic system output and its later analysis is one of the most important problems in signal processing. Different techniques have been employed by several researchers in a wide range of applications [3] such as neurophysics, electrocardiography, geophysics, and speech communication. LP is one of those methods which objective is to predict or estimate the future output of a signal $x(n)$ based on the past output observations. The complete mathematical development and a compilation of the different approaches of this well-known technique have been reviewed by Makhoul [3]. In vibration-based diagnostics, LP is presented by Randall [4] as a method that allows the separation of the deterministic or predictable part of a signal from the random background noise using the information provided by p past observations.

Adaptive noise cancelling (ANC) is another technique used to reduce the background noise in a signal, improving the visibility of the different signal components in the frequency spectrum. The first work in ANC [5] was performed by Howells and Applebaum at the General Electric Company between 1957 and 1960. Since then, this method has been successfully applied to different problems [5] including electrocardiography, cancelling noise in speech signals, and cancelling antenna sidelobe interferences. The main problem of this methodology comes from the fact that a reference signal (uncorrelated with the main signal) is needed in addition to the signal of interest. To solve this issue, a further development of ANC was formulated using a delayed version of the primary signal as a reference [5]; this latter version was named the SANC [6]. The capability of ANC and SANC to improve the signal to noise ratio of a signal and detect bearing faults masked by background noise has been demonstrated in several investigations [7–9].

Kurtosis gives a measure of how peaked a signal is, which is normally related with impulsive phenomena. In real applications, background noise often masks the signal of interest and, as a result, the kurtosis is unable to capture the characteristic peakness of a bearing fault signal. Therefore, the kurtosis is not useful as a global indicator of signal peakness, and it is necessary to apply it locally in different

Table 45.1 Bearing main dimensions

No. of rolling elements (n)	12
Ball diameter (B_d)	0.4063″
Contact angle (Φ)	40°
Pitch diameter (P_d)	1.811″
Input shaft speed (RPM)	710
Gear teeth	17

frequency bands [10]. This technique is named SK and was first introduced by Dwyer in [11]. The kurtogram [12, 13] is basically a representation of the calculated values of the SK as a function of the frequency f and the frequency band width Δf. The kurtogram allows us to identify the frequency band where the SK is maximum, and this information can be used to design a filter which extracts the part of the signal with the highest level of impulsiveness. An envelope analysis can be performed on the filtrated signal to identify bearing faults. This technique has been already applied successfully in bearing fault detection and diagnosis [14–16].

45.3 Experimental Setup

A gearbox test rig was constructed to carry out the experiments. This gearbox type (see Fig. 45.1) is used as part of the transmission driveline on the actuation mechanism of secondary control surfaces in civil aircrafts. One of the bearings in this gearbox failed in an endurance test, making it an ideal candidate for this investigation where fast natural degradation of the bearing occurred. The main dimensions of the bearing and the attached bearing defect frequencies can be seen in Tables 45.1 and 45.2, respectively. The transmission was driven by an electric motor with a nominal speed of 710 rpm. An electric load motor placed at the opposite side of the test rig was used to apply the different loads used in the experiment, simulating the actual load conditions during the flight.

The experiment ran continuously for 24 h a day over a duration of 36 days. Three accelerometers were mounted in the gearbox at locations identified in Fig. 45.1. The selected accelerometers (Omni Instruments model RYD81D) had an operating frequency range of 10 Hz to 10 kHz. These accelerometers were connected to signal conditioners (model Endevco 2775A) which were attached to a NI USB 6009 data acquisition device. Other than the vibration data, various parameters were monitored and stored at the same time and with the same sampling frequency: angular position of the input shaft, input and output torque, and shaft speed. The experiment started running on July 19, 2010, and the vibration measurements were taken on August 19, 2010, August 22, 2010, and finally on August 24, 2010. For each measurement case, a total 1,048,569 points were acquired at 5 kHz, resulting in a measurement length of approximately 3.5 min; sufficiently long to cover a whole loading cycle.

Table 45.2 Main defect frequencies and harmonics (Hz)

Harmonic	1X	2X	3X	4X	5X	6X
Shaft speed frequency (SS)	11.8	23.7	35.5	47.3	59.2	71
Gear mesh frequency (GM)	201.2	402	604	805	1,006	1,207
Inner race defect frequency(IRD)	83.2	166	250	333	416	499
Outer race defect frequency (ORD)	58.8	118	176	235	294	353
Cage defect frequency	4.9	9.8	14.7	19.6	24.5	29.4
Ball spin frequency	25.6	51.2	76.8	102	128	154
Rolling element defect frequency	51.2	102	154	205	256	307

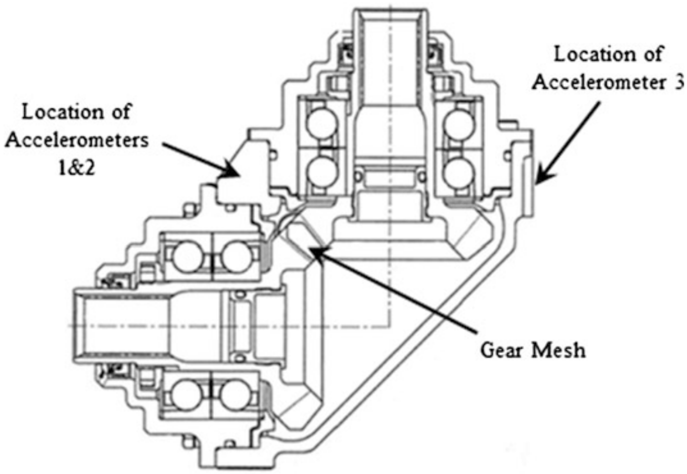


Fig. 45.1 Gearbox section

45.4 Results Obtained

Once the experiment was carried out, the data acquired was processed using the methodologies mentioned in Sect. 45.2. The results obtained for each measurement are plotted in Figs. 45.2, 45.3, 45.4 with the following format:

- amplitude spectrum of the original signal (*a*)
- amplitude spectrum of the signal obtained by LP (*b*)
- amplitude spectrum of the signal obtained by SANC (*c*)
- magnitude of the squared envelope of the signal obtained by filtration in the frequency band of maximum SK (*d*)

The spectrums of the original signal and the signals obtained by LP and SANC are represented twice. The first plot (left) corresponds to the spectrum covering frequency range of 0–2,500 Hz, which contains the gear mesh components and its harmonics. The second plot (right) covers the region of 0–500 Hz, where it is

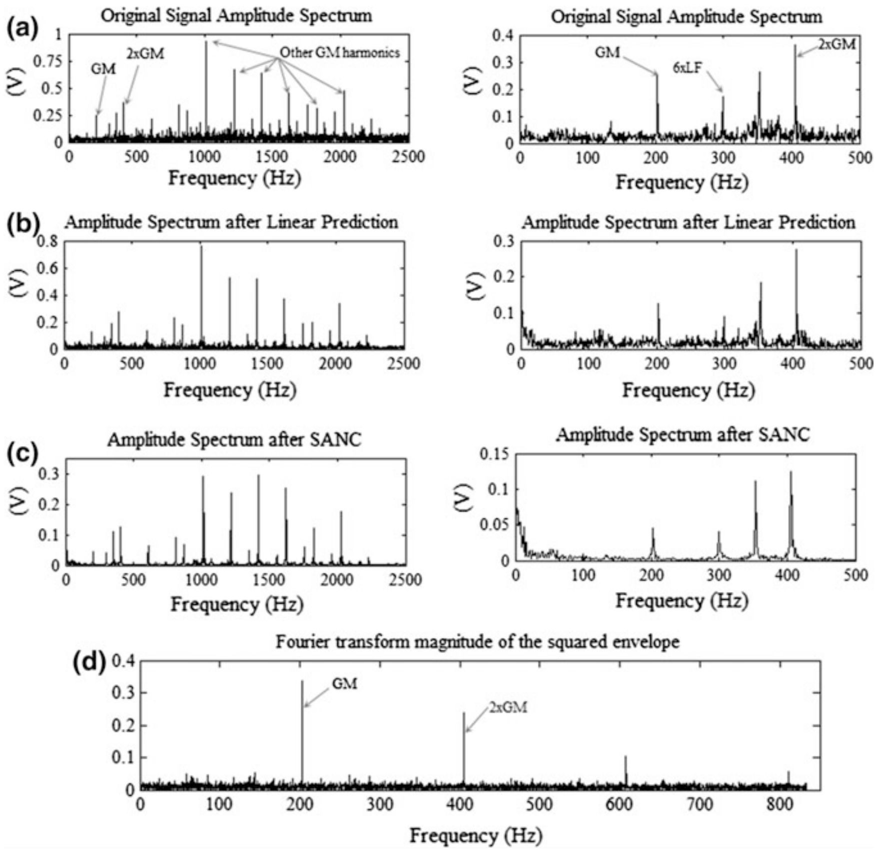


Fig. 45.2 Results obtained from the first observation (19/08/10)

easier to identify the typical defect frequencies. The main parameters used in each analysis are summarized in Table 45.3, where *p* represents the number of past observations used in LP; Δ, *H,* and μ are the time delay (in number of samples) filter order and forgetting factor [5] used in each SANC analysis. Finally *K* max, *fc,* and Δf represent the maximum kurtosis, the central frequency, and the bandwidth where it was found.

45.5 Results Discussion

From analysis of the first observation, (see Fig. 45.2) background noise was reduced by LP and especially by SANC, increasing clearly the signal to noise ratio (2.6 and 42.9 %, respectively) and facilitating the identification of the different signal components. The amplitude of the main peaks in the frequency spectrum

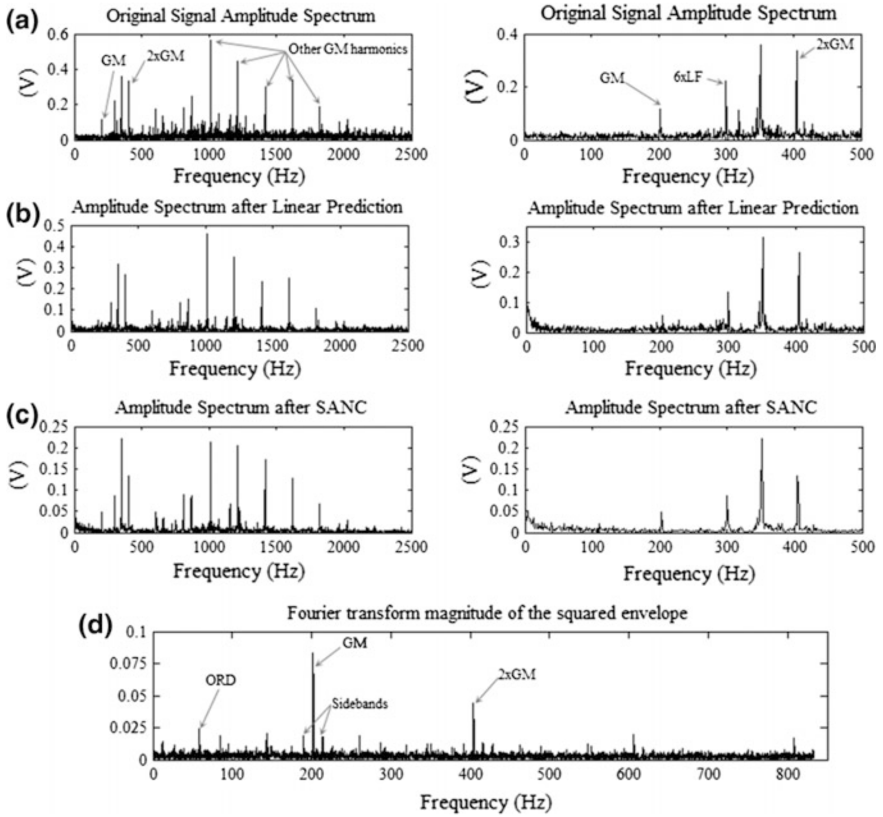


Fig. 45.3 Results obtained from the second observation (22/08/10)

was also reduced in magnitude, but this is not significant in terms of component identification, because the signal contains the same main components at a better signal to noise ratio. No new peaks masked by the background noise were identified in this observation. The envelope obtained after the signal filtration at the maximum kurtosis frequency band shows clearly that the signal is dominated by the gear mesh frequency, but at this stage, it does not provide any information of an incipient fault in the system.

The original spectrum of the second observation (see Fig. 45.3) shows more or less the same components as the first observation. No new signal components are identified by LP or SANC, despite the fact that the background noise reduction is considerable, with an improvement of the 14.4 and 16.5 %, respectively, in the signal to noise ratio. The most interesting result of this analysis is provided by SK: apart from the typical gear mesh frequency and harmonics, it is possible to identify a new peak at the frequency of 58.4 Hz, indicating an incipient fault in the outer race of the bearing in addition to sidebands around the gear mesh frequency: 190.2 and 214 Hz.

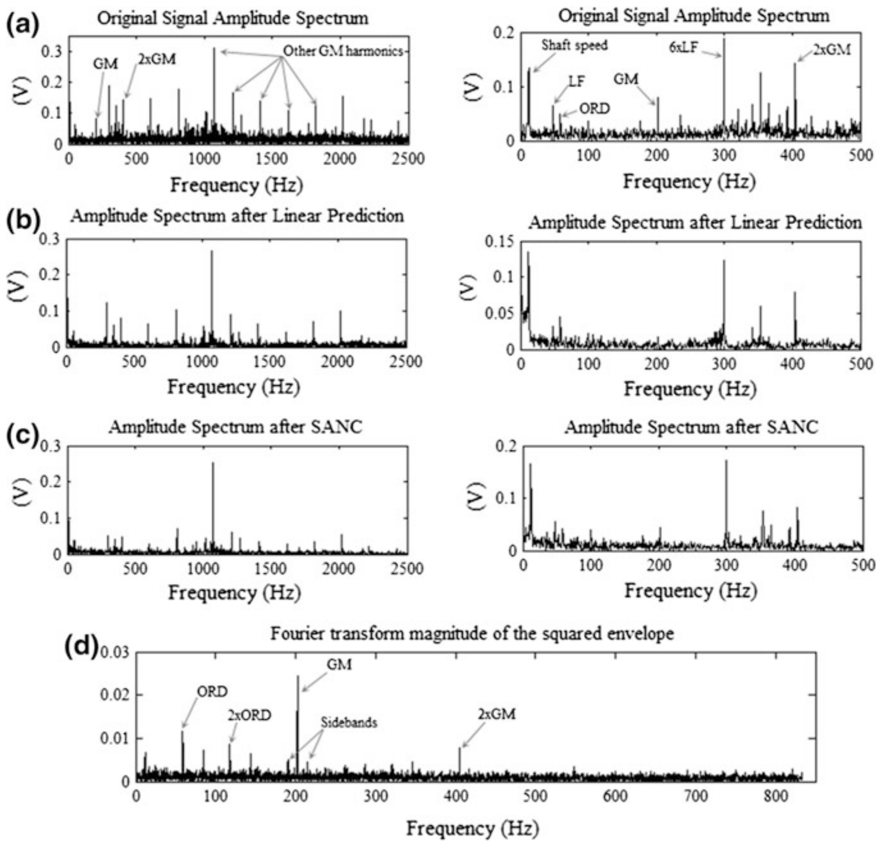


Fig. 45.4 Results obtained from the third observation (24/08/10)

Table 45.3 Parameters used in the analyses

Observation	LP	SANC			SK		
	p	*Δ*	*H*	*μ*	*K* max	*fc* (Hz)	*Δf* (Hz)
First	200	100	1,000	0.00001	2.4	2,083.33	833.3
Second	200	500	1,000	0.00005	2.3	2,083.33	833.3
Third	200	500	1,000	0.0001	1.6	2,083.33	833.3

On August 24, 2010, the last data capture was performed (see Fig. 45.4). All previously noted peaks in the spectrum of the original signal were evident, in addition to a clear peak at 58.8 HZ, indicating the defect in the outer race of the bearing. Moreover, several sidebands around the harmonics of the gear mesh frequency, separated by the shaft frequency, were noted. The spectrum of the squared envelope showed the peak at 58.8 Hz and a second harmonic of it at

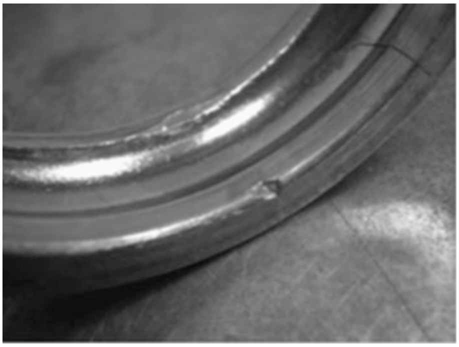


Fig. 45.5 Bearing outer race degradation after 36 days

117.9 Hz, indicating the fault in the outer race that was confirmed afterward by visual inspection of the component (Fig. 45.5).

45.6 Conclusions

This investigation shows the results of the application of three different vibration-based analysis methodologies for bearing diagnosis: LP, SANC, and SK. LP and specially the SANC showed its capability to reduce the background noise and facilitate the identification of the different components in the signal spectrum, but in this particular case, they did not help to identify the bearing defect earlier than the SK. The latter technique demonstrated its capability to identify the defect in the very early stage of degradation. This method is thus a very powerful tool for early detection of faults in bearings, even for those applications where strong background noise from other sources in the machine masks the characteristic fault components in the frequency domain.

Acknowledgments Financial support from the Marie Curie FP7-ITN project "Energy savings from smart operation of electrical, process and mechanical equipment—ENERGY-SMARTOPS," Contract No: PITN-GA-2010-264940 is gratefully acknowledged.

References

1. Ho D, Randall RB (2000) Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. Mech Syst Signal Process 14:763–788
2. Tan CC (1987) An adaptive noise cancellation approach for condition monitoring of gear box bearings. In: International tribology conference, Mel-bourne, Australia, 2–4 Dec 1987
3. Makhoul J (1975) Linear prediction: a tutorial review. Proc IEEE 63:561–580
4. Randall RB (2011) 3.6.3 linear prediction. In: John Wiley and Sons (ed) Vibration-based condition monitoring: industrial, aerospace and automotive applications, 2011th edn. Wiley, Singapore

5. Widrow B, Glover JR Jr, McCool JM (1975) Adaptive noise cancelling: principles and applications. Proc IEEE 63:1692–1716
6. Antoni J, Randall RB (2004) Unsupervised noise cancellation for vibration signals: part I—evaluation of adaptive algorithms. Mech Syst Signal Process 18:89–101
7. Khemili I, Chouchane M (2005) Detection of rolling element bearing defects by adaptive filtering. Eur J Mech 24:293–303
8. Dron J, Rasolofondraibe L, Chiementin X, Bolaers F (2010) A comparative experimental study on the use of three denoising methods for bearing defect detection. Meccanica 45:265–277
9. Patel VN, Tandon N, Pandey RK (2012) Improving defect detection of roll-ing element bearings in the presence of external vibrations using adaptive noise cancellation and multiscale morphology. In: Proceedings of the institution of mechanical engineers, vol 226, pp 150–162
10. Antoni J, Randall RB (2006) The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. Mech Syst Signal Process 20:308–331
11. Dwyer RF (1983) Detection of non-gaussian signals by frequency domain kurtosis estimation. In: Proceedings of international conference on acoustics, speech and signal processing (ICASSP). IEEE, vol 2, p 607
12. Antoni J (2007) Fast computation of the kurtogram for the detection of transient faults. Mech Syst Signal Process 21:108–124
13. Randall RB (2011) 5.3 spectral kurtosis and the kurtogram. In: John Wiley and Sons (ed) Vibration-based condition monitoring: industrial, aerospace and automotive applications, 2011th edn. Wiley, Singapore
14. Cong F, Chen J, Dong G (2010) Rolling bearing fault diagnosis based on spectral kurtosis in condition monitoring. In: Proceedings of the 23rd international congress on condition monitoring and diagnostic engineering management, Nara, Japan, 28 June–2 July 2010
15. Komgom CN, Mureithi NW, Lakis AA (2008) Application of time synchronous averaging, spectral kurtosis and support vector machines for bearing fault identification. American Society of mechanical engineers, pressure vessels and piping division (Publication), vol 7, p 137
16. Shi L, Randall RB, Antoni J (2004) Rolling element bearing fault detection using improved envelope analysis. IMechE Event Publications, vol 301

Chapter 46
Decision Support System for Infrastructure Sustainability in Operations

Gavin Shaw, Arun Kumar and David Hood

Abstract There is an increasing awareness of sustainability and climate change and its impact on infrastructure and engineering asset management in design, construction, and operations. Sustainability rating tools have been proposed and/or developed that provide ratings of infrastructure projects in differing phases of their life cycle on sustainability. This paper provides an overview of decision support systems using sustainability rating framework that can be used to prioritize or select tasks and activities within projects to enhance levels of sustainability outcomes. These systems can also be used to prioritize projects within an organization to optimize sustainability outcomes within an allocated budget.

46.1 Introduction

Civil infrastructure plays a key role in supporting and improving our current way of life. However, the assets can have a large impact on the region around them, which are both positive (usually for the purpose they are built) and negative (consequences and unintended effects). There is an increasing trend for society to place an importance on the role of sustainability to ensure that there is an environment suitable for future generations. In order to ensure future generations have a sustainable quality of life, it is increasingly important to integrate sustainability

G. Shaw (✉) · A. Kumar · D. Hood
Science and Engineering Faculty, Queensland University of Technology, Cooperative Research Centre for Infrastructure and Engineering Asset Management, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: gavin.shaw@qut.edu.au

A. Kumar
e-mail: arun.kumar@qut.edu.au

D. Hood
e-mail: davidahood@mac.com

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_46,

into the way industry operates, including the infrastructure industry. It is therefore important to undertake sustainability assessment of civil infrastructure projects. By having organizations undertake sustainability assessments of civil infrastructure projects both during construction and in operation, the industry can assist to drive outcomes and results that will benefit society and future generations and make their own operations more efficient.

With an increasing awareness of sustainability and climate change as well as increasing costs, infrastructure and engineering asset management has become an internationally recognized field, with the aim of reducing the costs of ongoing infrastructure operations. This field has now, in addition, taken on the challenge of reducing the impacts that infrastructure operations and maintenance have on the environment and society and hence the idea of sustainable infrastructure operations or 'sustainability in infrastructure operations.'

The purpose of this paper was to review the state of decision support systems in the field of infrastructure sustainability. This paper is structured as follows: in Sect. 46.2, the study is outlined; Sect. 46.3 presents the results of the study; and Sect. 46.4 sums up the key findings. Finally, the paper presents the concluding remarks in Sect. 46.5.

46.2 Study Overview

The primary focus of this literature review is decision support system research, application, and usage in infrastructure operations, where the driver is not just to support the economics of day to day operations and maintenance, but also to include these activities in a way that delivers sustainability outcomes, through the consideration of the core sustainability themes, environment, society/community, and economy (also known as the triple bottom line).

The scope of this review also covers:

- decision support systems in general
- decision support systems in the area of infrastructure including the design of the management/operation and emergency response
- decision support systems in the more general area of sustainability.

While the main purpose of this review is in the area of sustainable infrastructure operations, the separate specific areas of infrastructure and sustainability are also included. This review is based on existing published academic and research works from a variety of sources.

46.3 Results of Study

For the purposes of this paper, a common view of the triple bottom line, which covers environment, finance/economic, and societal/community, is used as the definition of sustainability.

Table 46.1 Breakdown of application areas and systems, frameworks, and processes

Application area	Type			
	Decision support system	Framework	Decision making process	Application area total
Infrastructure operations	4	1*	1*	5
Sustainability	3	1	0	4
Infrastructure sustainability	4	3	2	9
Type total	11	4/5	2/3	

* One of the reviewed works can be considered to present both a framework and a decision making process and thus belongs to both of these type categories

Surveys conducted in 2005 [1] and 2006 [2] on decision support systems examined a large collection of publications across several journals and observed that since 1994, the publication rate for work in decision support system (in general and not a specific application) has been decreasing up to the end of the surveys in 2001 and 2003. This trend is, however, disputed in [3] which observed that research is growing with an increase in publication rates. Further trends identified by Arnott and Pervan [1] and Eom and Kim [2] indicate that

- approximately half of the works reviewed did not use judgment and decision making reference research;
- an increasing trend over time in the use of case studies;
- the dominant application area for decision support systems is still production and operations, which can include infrastructure operations, part of the application area of interest to this study;
- systems and research based around optimization models is increasing and is the most popular field; and
- there is a gap in the identification of the clients and users of decision support systems.

From the research referred to in the surveys [1] and [2], it is apparent that decision support systems can be used for a variety of tasks and are available in many forms. It also appears that for just about every application where a decision support system is wanted, a new system is built.

Table 46.1 shows the breakdown by application area and system type of the unique proposals present in the publication covered by this study.

46.3.1 Infrastructure and Infrastructure Operations

A topic-specific survey focusing on infrastructure and in particular roads and bridge construction showed two issues of interest. Firstly, that infrastructure maintenance, considered to be part of the operation of infrastructure, does not appear to feature heavily in publications, with only one publication out of 22 making reference to maintenance [4]. Secondly, out of 17 publications related to decision making, only 6 made reference to or considered sustainability and in that referred to the application to buildings [4]. This implies that there is limited research in decision support systems in the areas of infrastructure operations and sustainability. It is stated by Zavadskas et al. [4] that there is no single decision making system, tool, or method that can satisfy all problem and scenarios. This would imply that each unique problem may potentially need a decision support system specifically designed for it or at the minimum an existing system would need to be modified. Following on from the survey [4], it could be assumed that there is limited research in decision support systems focusing on infrastructure operations.

Research in [5] and [6] cover decision support in infrastructure (separate to those works that cover infrastructure sustainability or sustainable infrastructure operations). Neither of the two research works examined infrastructure operations. Šelih et al. [5] developed a decision support system to determine the priorities of a maintenance schedule (which prioritises projects, but not how to undertake them) while [6] provides a decision support system to help design wastewater treatment plants. Both systems use multi-criteria (also known as multi-objective) analysis in order to develop and propose a solution that best meets multiple goals. It was noted by Hakanen et al. [6] that many previous works only considered optimizing one performance goal in the design, planning, or operation of infrastructure and that they were only able to identify two previous works in infrastructure that looked at a multiple criteria and neither were interactive systems. This implies that prior to [6], either there has only been one important criteria/goal for decision support systems with application to infrastructure or that other criteria/goals have been ignored. For decision support systems with multiple criteria or goals to meet or be measured against, multi-criteria approaches offer the best solution.

A popular application for decision support systems when it comes to infrastructure is in the area of emergency management. Of the publications reviewed on emergency response, only one made any reference to considering the environment during the response [7]. It appears that environmental and sustainability concerns are of minor importance in an emergency as the priority is naturally to help people out and restore some level of service provided by the infrastructure. The environment is most likely only considered when there is a legal requirement to do so during an emergency or disaster.

46.3.2 Sustainability

With sustainability becoming more central and having an impact on the way a business or organization (including governments) operates, it was only natural that decision support systems and tools would be developed for this field. This has given rise to the environmental decision support systems (EDSS).

The aim of an EDSS is to achieve the best environmental outcome, within reason, by improving the associated business processes. For example, the system in [8] is focused on improving the business processes related to the water use of the river system, which includes environmental uses. Similarly, systems can also focus on the management of natural resources [9], potentially with a framework to guide the process [10]. Either approach can result in sustainability improvements as they ultimately drive home the impact that the proposed actions have on the environment. It is not so much the approach taken, but rather the final outcome.

From the start, the entities of interest come from the natural environment (land, water, etc.), and thus, systems developed in this area do not focus on infrastructure, only the impact. The most common natural environment to feature in decision support system works in this area is water; [8, 9, 11, 12] are some examples; with the argument that this is due to its value in a social and economic context. While water management and sustainability have benefited from this focus, other areas have potentially been left behind.

A topic-specific survey conducted in 2007 focusing on the area of EDSS by Matthies et al. [9] identified trends that included

- a tendency toward better representation of reality in models, improving user interfaces, and an increasing focus on use in negotiation or group contexts;
- major research focus on water management, with land management being second; and
- there is an emphasis on decreasing the time taken to generate or discover a solution while providing higher accuracy and improvement in handling more complex problems or scenarios.

Achieving sustainability for the natural environment is said to be hampered by five challenges according to [11];

1. Lack of integrated tools
2. The segmentation of bodies responsible for sustainability
3. Limited stakeholder interest and participation
4. Lack and/or inability to undertake self assessment
5. Lack of interdisciplinary training

Finally, as sustainability and its achievement through a decision support system ultimately involves scientific aspects, there is the question of how science and its knowledge can best support decision making (through a decision support system) in order to achieve sustainability in the natural environment [12]. This question is yet to be fully solved. There are several issues that have prevented the successful integration of science and decision making and include

- scientific results are not always available and in a suitable form;
- the lack of uncertainty estimates to go with the scientific knowledge;
- differing values, interests, concerns, and perspectives; and
- lack of mutual understanding.

46.3.3 Infrastructure Sustainability

A major trend in decision support system research in this application area is the development of frameworks as opposed to an actual decision support system (e.g., an actual application). Instead of developing a system to support decision makers, frameworks outlining a decision making process are developed [12–19]. The advantage is that these frameworks could be applied into any system being developed, especially the more generic ones. However, there is a downside in that there often is no example of a decision support system implementing the framework to use as a reference. Also, their tendency toward being generic, either for multiple infrastructure classes or a single infrastructure class (e.g., transport) can mean that they do not necessarily provide the best criteria for a given infrastructure type (e.g., a highway).

A survey of sustainability frameworks for transportation systems was conducted in [13]. It was noted that there was no standard definition for transport sustainability (the survey provides a high-level view of 16 different sustainable transport initiatives); it was being defined through impacts that the system has on the economy, environment, and general social well-being and is measured through system effectiveness and efficiency. It then asked that if there is consensus on what a sustainable transport system is and thus how a plan for a sustainable system can be made [13]. While this seems to imply a focus on the planning and/or design aspects, if a sustainable transport system is not defined, how can one operate a transport system in a sustainable manner? Furthermore, while the survey by Jeon and Amekudzi [13] only looked at transport, this question could be asked of any infrastructure class or type.

The survey in [13] also indicated that there also was no standard framework for sustainability in transport and those existing frameworks could be classified into three groups:

- linkage-based frameworks that capture relationships between causal factors, impacts, and corrective actions;
- Impact-based frameworks that focus on the nature and extent of the various kinds of impacts; and
- Influence-oriented frameworks developed to keep in mind the relative level of influence that users, agencies, organizations, etc., have on activities and/or actions the effect progress toward sustainability.

These frameworks could be applied to any infrastructure and not just transport. However, no analysis or review was undertaken in [13] to compare the different framework categories and no examples or references to existing instances were provided.

The major limitation to the usefulness of the frameworks in the context of decision support system application put forward is that they are not actually decision support systems nor are they primarily made for one. Some frameworks, such as those in [15] and [17], focus more on the criteria or assessment used to determine sustainability rather than a system to support a decision maker. Having said this, some frameworks, for example, [16] and [18], provide a process for decision making on which a decision support system can be built.

One decision making process that considered sustainable infrastructure operations is proposed in [16], where a system for restoring environmental flows (more natural flow regimes) in rivers that have been dammed. Through modification of the standard operating practice of a dam, the environmental and, to lesser extent, social impacts caused by the dam could be reduced. The decision making process outlines what needs to be considered and that goals need to be clearly determined. This decision making process provides no guidance of how a decision support system might be implemented and what functionality it will have.

A further disadvantage to works that provide a framework only is that when there is no case study, it is difficult to determine the suitability or ability of the framework. For example, a strategic planning tool that provides a theoretical frame for understanding and determining transitions and impacts in a societal systems was presented in [19]. However, the lack of a case study or example of the tool in use makes it difficult to determine how it performs.

Despite the tendency of works in this area to be frameworks, decision support system applications have been developed. In [20–23], decision support systems are proposed to meet the requirements (including sustainability concerns) of various infrastructure types. However, these systems focus more on the design and planning stages with regard to infrastructure, rather than sustainable operations.

Similarly, a waste management system [20] also takes a long-term view and considers non-infrastructure aspects, as it looks across the whole waste management process. It was developed specifically for European Union countries to meet policies and requirements which offer obstacles in its adaptation elsewhere. The research by Boer et al. [20] also raises two interesting points with regard to internalizing external costs; first that doing so for those especially related to environmental effects is not always supportive of transparency in the outcomes and secondly that a methodological approach internalizing social effect cost in waste management does not exist. For other infrastructure types and classes, the same may also be true and even if not for all classes, the approach may be different. This is something that needs to be taken into account in a decision support system that is also going to report environmental and social effects through costing.

46.4 Key Points

From this review of previous research, we have identified five key points/issues in the area of decision support systems, especially in the application area of infrastructure sustainability. The key points identified in this review include:

- limited published academic and research works in the area of sustainable infrastructure operations which can be mitigated through ongoing searches, including a focused study on commercial systems and interaction with industry to gain knowledge directly;
- the individual and tailored nature of existing systems discovered through this review, with this key impact indicating that unless an existing decision support system that is very similar to the objective of the project, a tailored solution will have to be developed by the project;
- a potential perceived risk of the research area of decision support systems becoming irrelevant to industry which can be mitigated by ongoing communication and participation of industry partners to inform and demonstrate the both the importance and relevance of a decision support system;
- the lack of a standard definition of sustainability in decision support systems, especially those in the field of infrastructure operations mitigated through working closely with AGIC to ensure that the project and its developed decision support system aligns with their standards and definitions; and
- the trend from actual decision support systems (applications) to decision making frameworks and processes resulting in the need for the project to clearly define its objectives to avoid confusion and to communicate this with industry.

46.5 Conclusion

This paper undertook a review into the research of decision support systems and their application to the area of infrastructure sustainability. The results of this study are presented here in this paper.

From this review, we identified five key points/issues (presented in Sect. 46.4) that were discovered to be relevant. Our study details out the lack of published material on decision support systems in the key area of interest, infrastructure sustainability, and trends taking place, such as the move toward frameworks and processes. The study shows that for each problem instance, there is a custom solution and that no single system, framework, or process will fulfill and solve every problem. The logical conclusion from this is that to encourage and support industry in taking sustainability seriously, appropriate support needs to be in place and it needs to be tailored to solving the particular problem.

Acknowledgments This work was undertaken as part of The Sustainability Assessment for Infrastructure Project and is part of the larger CRC for infrastructure and engineering asset management. The CRC for infrastructure and engineering asset management (CIEAM) would like to acknowledge the financial support from the Commonwealth Government's Cooperative Research Centres Program and the contributions from our industry, government, and university participants. The overall support provided by the CEO, and Professor Joseph Mathew is gratefully appreciated.

References

1. Arnott D, Pervan G (2005) A critical analysis of decision support systems research. J Inf Technol 20:67–87
2. Eom S, Kim E (2006) A survey of decision support system applications (1995–2001). J Oper Res Soc 57:1264–1278
3. Burstein F, Holsapple CW (2007) Handbook on decision support systems 1: basic themes. Springer, Dordrecht
4. Zavadskas EK, Liias R, Turskis Z (2008) Multi-attribute decision-making methods for assessment of quality in bridges and road construction: state-of-the-art surveys. Baltic J Road Bridge Eng 3:152–160
5. Šelih J, Kne A, Srdić A, Žura M (2008) Multiple-criteria decision support system in highway infrastructure management. Transport 23:299–305
6. Hakanen J, Miettinen K, Sahlstedt K (2011) Wastewater treatment: new insight provided by interactive multiobjective optimization. Decis Support Syst 51:328–337
7. Merz M, Hiete M, Bertsch V (2009) Multicriteria decision support for business continuity planning in the event of critical infrastructure disruptions. Int J Crit Infrastruct (IJCIS) 5:156–174
8. Kalken TV, Szylkarski S, Berry D, Nachiappan N, Skinner J (2011) Improving river efficiency: a computer aided river management system (CARMS) for the Murrumbidgee river. Australian Water Management Review. Australian Publishing Resource Service, Australia
9. Matthies M, Giupponi C, Ostendorf B (2007) Environmental decision support systems: current issues, methods and tools. Environ Model Softw 22:123–127
10. Sieber S, Muller K, Verweij P, Haraldsson H, Fricke K, Pacini C, Tscherning K, Helming K, Jansson T (2008) Transfer into decision support: the sustainability impact assessment tool (SIAT). In: Helming K, Perez-Soba M, Tabbush P (eds) Sustainability impact assessment of land use changes. Springer, New York
11. Georgakakos AP (2006) Decision support systems for integrated water resources management with an application to the Nile basin. In: Castelletti A, Soncini-Sessa R (eds) Topics on system analysis and integrated water resources management. Elsevier, Amsterdam
12. Liu Y, Gupta H, Springer E, Wagener T (2008) Linking science with environmental decision making: experiences from an integrated modeling approach to supporting sustainable water resources management. Environ Model Softw 23:846–858
13. Jeon CM, Amekudzi A (2005) Addressing sustainability in transportation systems: definitions, indicators, and metrics. J Infrastruct Syst 11:31–50
14. Mcdaniels T, Chang S, Cole D, Mikawoz J, Longstaff H (2008) Fostering resilience to extreme events within infrastructure systems: characterizing decision contexts for mitigation and adaptation. Glob Environ Change 18:310–318
15. Sahely HR, Kennedy CA, Adams BJ (2005) Developing sustainability criteria for urban infrastructure systems. Can J Civ Eng 32:72–85
16. Richter BD, Thomas GA (2007) Restoring environmental flows by modifying dam operations. Ecol Soc 12
17. Liu KFR, Lai J-H (2009) Decision-support for environmental impact assessment: a hybrid approach using fuzzy logic and fuzzy analytic network process. Expert Syst Appl 36:5119–5136
18. Pearson L, Coggan A, Proctor W, Smith T (2010) A sustainable decision support framework for urban water management. Water Resour Manage 24:363–376
19. Haan JD, Ferguson B, Brown R, Deletic A (2011) A workbench for societal transitions in water sensitive cities. In: 12th international conference on urban drainage. Porto Alegra, Brazil
20. Den Boer J, Den Boer E, Jager J (2007) LCA-IWM: a decision support tool for sustainability assessment of waste management systems. Waste Manage 27:1032–1045

21. Makropoulos CK, Natsis K, Liu S, Mittas K, Butler D (2008) Decision support for sustainable option selection in integrated urban water management. Environ Model Softw 23:1448–1460
22. Cai YP, Huang GH, Lin QG, Nie XH, Tan Q (2009) An optimization-model-based interactive decision support system for regional energy management systems planning under uncertainty. Expert Syst Appl 36:3470–3482
23. Urich C, Bach PM, Hellbach C, Sitzenfrei R, Kleidorfer M, Mccarthy DT, Deletic A, Rauch W (2011) Dynamics of cities and water infrastructure in the DAnCE4Water model. In: 12th international conference on urban drainage. Porto Alegre, Brazil

Chapter 47
Sustainability in Infrastructure Asset Management

Gavin Shaw, Rick Walters, Arun Kumar and Antony Sprigg

Abstract The expectation to integrate sustainability aspects (social, environmental, and economic success) into the design, delivery, and operation of infrastructure assets is growing rapidly and globally. There are now several tools and frameworks available to benchmark and measure sustainable performance of infrastructure projects and assets. This paper briefly describes the infrastructure sustainability (IS) rating tool developed by the Australian Green Infrastructure Council (AGIC) that was launched in February 2012. This tool evaluates sustainability initiatives and potential environmental, social, and economic impacts of infrastructure projects and assets. The rating tool provides the following benefits to industry: a common national language for sustainability; a vehicle for consistent application and evaluation of sustainability in tendering processes; assists in scoping whole-of-life sustainability risks, enabling smarter solutions that reduce risks and costs; fosters resource efficiency and waste reduction, reducing costs; fosters innovation and continuous improvement in sustainability outcomes; and builds an organization's credentials and reputation in its approach to sustainability. The infrastructure types covered by this tool include transport, energy, water, and

G. Shaw (✉) · A. Kumar
Science and Engineering Faculty, Queensland University of Technology, Cooperative Research Centre for Infrastructure and Engineering Asset Management, 2 George St., Brisbane, QLD 4001, Australia
e-mail: gavin.shaw@qut.edu.au

A. Kumar
e-mail: arun.kumar@qut.edu.au

R. Walters · A. Sprigg
Australian Green Infrastructure Council (AGIC), Level 20, Tower 2, Darling Park, 201 Sussex St., Sydney, NSW 2000, Australia
e-mail: rwalters@agic.net.au

A. Sprigg
e-mail: asprigg@agic.net.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_47,

communication. The key themes of sustainability evaluation will be briefly presented in this paper, and they include management and governance; use of resources; emissions, pollution, and waste; ecology; people and place; and innovation.

47.1 Introduction

Civil infrastructure plays a key role in supporting and improving our current way of life. However, the construction, operation, and decommissioning of infrastructure assets can have significant impacts on society and the region around them, some positive (usually for the purpose they are built) and some negative (consequences and unintended effects). There is an increasing trend for society to place an importance on the role of sustainability to ensure that our world continues to be suitable for future generations. To achieve this, it is increasingly important to look at integrating sustainability outcomes into the way industry operates, including the infrastructure industry. Assessment using a sustainability rating scheme can be a highly effective means to build sustainability aspects into the design, construction, and operation of infrastructure assets. This can assist to drive outcomes and results that will benefit society and future generations and make operations more efficient.

47.2 Background

Civil infrastructure can be defined as the structural elements of economy and society which allow for production and distribution of goods and services without themselves being part of the production process. Such infrastructure includes roads, railways, bridges, tunnels, ports, airports, distribution grids/networks (such as pipes, poles, and wires used in water, sewage, electricity, communications, etc.), water and/or resource management, preparatory civil works, and more. Spending on infrastructure is massive and predicted to continue to increase. An example of the scale of spending is that across just 7 countries, there are plans or the need for nearly $5 trillion dollars' worth of investment [1].

Given the large amounts spent on infrastructure projects, the size and scale some projects can take on and the sheer volume of projects being undertaken, planned for or needed in the future (especially in countries undergoing modernization) civil infrastructure projects for better or worse will continue to have significant impacts on society, the surrounding environment, and communities, along with their residents.

With increasing awareness of sustainability and climate change in the community and increasing expectations of being more 'environmental friendly', organizations involved in delivery or operation of infrastructure assets (regardless of the stage) have to now take into consideration and face the challenge of also reducing the negative impacts and enhancing positive impacts from infrastructure assets on the environment and society, and hence, the idea of 'infrastructure sustainability' has been born.

Traditionally, the primary focus for an organization involved in an infrastructure project or asset has been to meet the needs/terms specified without exceeding a money or time limit (e.g., on time and on budget). The focus has been on economic performance and getting good value for money. More recently (gaining intensity over the last few decades), a shift toward also considering the environment has come about (often via government legislation) and many organizations now undertake activities such as environmental impact studies and assessments and aim to reduce negative impacts the infrastructure asset (and activities associated with it) may have on the surrounding environment. However, true sustainability considers not just the economic factors and environmental impacts, but also the impacts on society—the communities and people that will be affected, both positively and negatively, by the infrastructure asset or activities associated with it. This vision of sustainability (economic, environment, and social) is often known as the triple bottom line [2, 3]. Another version of sustainability also being put forward is the quadruple bottom line, which encompasses the same aspects as the triple bottom line, but also includes governance as a fourth aspect. The triple bottom line (or the quadruple bottom line) in itself is also just one part of a larger vision of corporate sustainability which includes relationships and responsibilities [4].

47.3 Key Issues in Sustainability

While it is becoming more important and critical for sustainability to be considered in the design, construction, and operation of civil infrastructure assets, it is also important to ensure that appropriate issues are taken into consideration when performing a sustainability assessment.

There are a small number of sustainability rating tools for infrastructure in existence globally and a selection were analyzed in [5], identifying some similarities and differences. Half the tools in the study were specific to the transport sector, while the remaining were more generic and suitable for all civil infrastructure (non-building). The rating tools showed a reasonable degree of alignment in terms of the key issues that were covered but differed in the weighting (value) assigned to each issue. The study in [5] also looked at decision support (tools to guide or help make a decision) in the transport sector and identified that of the 16 tool frameworks (which outline a decision-making process to follow, without assessing the final decision) included, only four included issues/criteria that combined covered all three aspects of the triple bottom line and that the majority of frameworks were heavily weighted toward environmental issues.

This paper presents the framework of AGIC's infrastructure sustainability (IS) rating tool. This framework comprises six themes and fifteen categories incorporating issues that organizations involved in civil infrastructure assets need to consider when undertaking sustainability assessment [6, 7].

47.3.1 Management and Governance

Governance involves the establishment and oversight of an organization's or a project's purpose, systems, structure and processes, and their implementation for the effective delivery and operation of infrastructure. Good governance in the context of IS is vital and is sometimes referred to as the fourth pillar of sustainability under a quadruple bottom line approach. The three categories within this theme are as follows:

- Management Systems: Good management systems help to ensure consistent and efficient operations and support during decision making. Management systems alone do not guarantee the achievement of sustainability, but they are considered to be a critical component. This category encourages sustainability to be comprehensively addressed within management systems from the policy level down to detailed processes.
- Procurement and Purchasing: This category assesses the extent to which economic, environmental, and social aspects and impacts have been considered in the evaluation, selection, and final procurement of goods and services for an infrastructure project or asset.
- Climate Change Adaptation: Until recently, infrastructure was designed on the basis of historical weather records, assuming that the climate will remain the same. Climate change means that this assumption is no longer valid, so that long life span infrastructure now needs to be designed, constructed, and operated to cope with the projected much hotter, drier, and stormier climatic conditions, with higher sea levels. This category facilitates the self-assessment, and thus, rating of the appropriateness and effectiveness of how climate change risk and adaptation issues have/will be addressed.

47.3.2 Using Resources

The infrastructure industry is a large and intensive user of materials, energy, and water. The ultimate goal is to reach a state where finite natural resources are consumed no faster than the planet can replenish them. The resources of most concern are those that are non-renewable. The three categories within this theme are as follows:

- Energy and Carbon: The majority of energy (especially in Australia) is derived from non-renewable sources the consumption of which generates greenhouse gas (GHG) emissions. Such emissions are increasing and are being linked to climate change events, which can impact negatively on the environment and society.
- Water: Fresh water is becoming increasingly scarce, and excessive consumption threatens ecosystem function. In regard to water, the aim of all infrastructure assets at each stage of their life cycle should be first to avoid or reduce water consumption per service output and then to replace potable water with effective reuse and recycling of locally appropriate alternative water sources.
- Materials: Infrastructure typically involves the consumption of large quantities of materials, a significant portion of which are derived from natural resources. The supplies of some of these resources are limited and are becoming increasingly scarce. To fully assess materials in a sustainability context requires consideration of a complex set of environmental, social, and economic factors across a life cycle perspective [8]. The intent of this category is to encourage design and practice that minimizes the consumption of precious resources, optimizes resource efficiency, and reduces the environmental impacts of infrastructure.

47.3.3 Emissions, Pollution, and Waste

This theme focuses on understanding and measuring emissions, pollution, and waste and their impacts, identifying and implementing feasible opportunities to reduce those impacts, and restoration to reverse past impacts. The ultimate goal is to reach a state where wastes are emitted no faster than the rate at which the planet can absorb them and also support them to aim for zero emissions and zero harm to the environment and society. The three categories within this theme are as follows:

- Discharges to Air, Land, and Water: Discharges and emissions often have negative and harmful impacts on the environment and society. This category assesses the level and effectiveness of management practices for preventing and mitigating discharges to air, water, and land over the life cycle of a given piece of infrastructure. It also seeks to encourage initiatives to enhance natural capital.
- Land (usage, conditioning, etc.): In Australia, high-value environmental or social land is scarce, and often land is under pressure from different industries and uses. This category focuses on the project-level decisions that flow on from good strategic land-use planning in relation to infrastructure.
- Waste: Waste generation globally is increasing, and recycling and reuse are not increasing at the same rate. Waste from construction and demolition is significant in Australia, and currently, 45 % of this is disposed to landfill. This category assesses the level and effectiveness of waste management practices for achieving the goal of zero waste to landfill over the life cycle of a given piece of infrastructure through recycling, reuse, design, optimization, and contract management.

47.3.4 Ecology

Ecosystems are considered to have important environmental and social functions that cannot be replicated artificially. Furthermore, Australia is considered to have globally significant ecosystems. The intent of this theme is not just to minimize or mitigate the negative impacts on ecosystems through all stages of the project life cycle, but also to foster infrastructure decisions that enhance ecosystem functioning.

47.3.5 People and Places

The people and place categories focus on effects on the well-being of communities and the users of the infrastructure, how infrastructure integrates with and enhances the surrounding urban and landscape environment, how the past is recognized and conserved and how stakeholders participate in infrastructure design, construction, and operation. The four categories within this theme are as follows:

- Community Health, Well-being, and Safety: The quality of both natural and built environments along with other factors has significant effects on the general well-being, health, and safety of communities. The design, construction, and operations practices of infrastructure can dictate a community's behavior which contributes to their health, well-being, and safety.
- Heritage: Heritage is the cultural significance that we inherit from the past that we value and want to pass on to future generations. This category focuses on how heritage is assessed and then managed through design, construction, and operation of infrastructure.
- Stakeholder Participation: Stakeholder participation refers to the processes and mechanisms that enable stakeholders who have a direct or indirect interest in infrastructure development to be part of decision making. This category focuses on developing a strategic and planned approach to stakeholder participation, managing, and monitoring implementation of the participation process, achieving a high level of participation for negotiable issues, effectively communicating, and effectively addressing community concerns.
- Urban and Landscape Design: Urban and landscape design is a process, and an outcome concerned with the arrangement, appearance, and function of places in suburbs, towns, cities and regions. This category focuses specifically on the analysis, planning, and design of the infrastructure asset within its community and environment.

47.3.6 Innovation

Often the sustainability issues facing industry and society are difficult and complex due to their interwoven social, economic, and environmental aspects. The innovation category supports pioneering initiatives, methods, processes, procedures, technologies, etc. in sustainability that assist in transforming the market or industry sector toward a more sustainable approach and answering the challenges faced.

47.3.7 Summary

The IS rating tool provides the following benefits:

- Provide a common national language for sustainability in infrastructure.
- Provide a vehicle for consistent application and evaluation of sustainability in tendering processes.
- Help in scoping whole-of-life sustainability risks for projects and assets, enabling smarter solutions that reduce risks and costs.
- Foster resource efficiency and waste reduction, reducing costs.
- Foster innovation and continuous improvement in the sustainability outcomes from infrastructure.
- Build an organization's credentials and reputation in its approach to sustainability in infrastructure.

It provides industry with a means to voluntarily assess performance and to be recognized for good performance. The assessment is facilitated by using the IS rating tool which consists of this technical manual and the tool scorecard (downloadable from www.agic.net.au). The tool builds on current guidance and practices by providing industry with an incentive and protocol for assessing, benchmarking, and 'labeling' the sustainability performance of infrastructure projects or assets at the planning and design, construction, and/or operations phases.

47.4 Sustainability and Infrastructure Phases

The typical infrastructure asset life cycle broadly includes planning and design, construction, operation (with maintenance), and decommissioning (see Fig. 47.1). For most infrastructure assets, the operations phase is the longest and may not have a predetermined ending date. An infrastructure asset will have an impact on the environment and society from the moment construction starts to the end of decommissioning (and perhaps beyond if the asset is not disposed of) and economic impacts in all phases, even during design.

In order to achieve the best sustainable outcomes, sustainability implications need to be considered at each phase. An infrastructure asset can be designed to be sustainable (and assessed during the design phase as having a sustainable design), but it could be constructed in such a manner that has large negative impacts on the

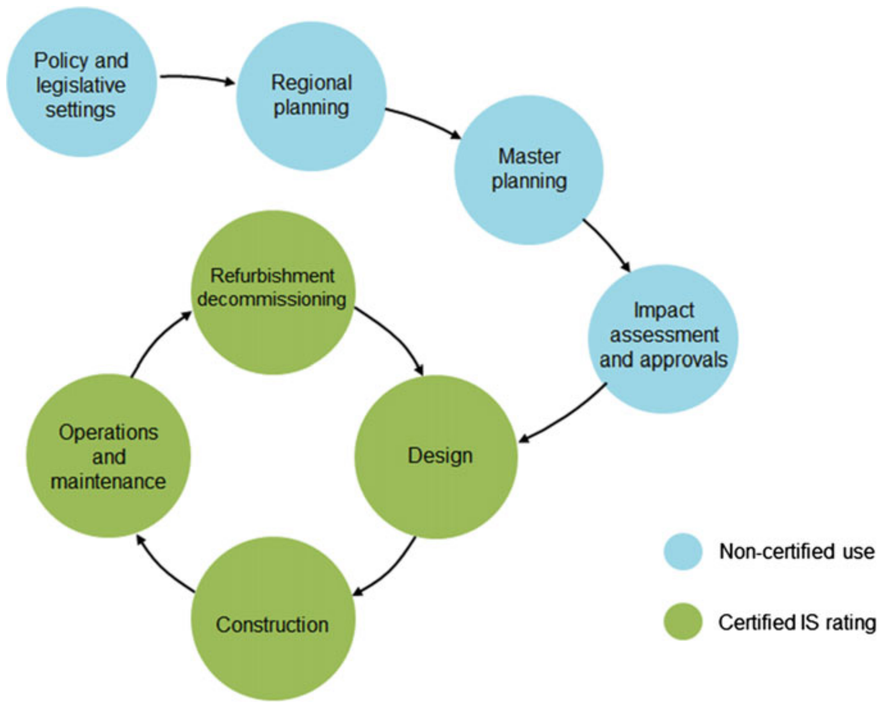


Fig. 47.1 Infrastructure asset life cycle and the phases rated by the IS rating tool (reproduced from Australian Green Infrastructure Council sources)

environment or society (such as the use of carbon-intensive materials, large-scale land, and community disturbance) or could be operated using a process that harms the environment (e.g., energy-intensive procedures and socially disruptive activities).

By assessing an asset and/or the associated project at each phase, an organization can ensure that it is designed with sustainability in mind, constructed in a manner that is sustainable with good practices used, operated, and maintained in a manner that ensures sustainability, and decommissioned in a sustainable manner. Failure to assess and monitor assets and projects during construction and later operations could easily result in an asset that was designed to meet sustainability outcomes, but then was constructed in a manner, where no concern was given to the environment or society and thus undermine the sustainable design and reduce the benefits achieved.

To help avoid this, rating schemes cover more than one phase of an asset's life cycle. From the study in [5], all but one of the rating schemes covered the design and construction phases. However, only one rating scheme (AGIC's IS rating tool) also covered the operations phase. The coverage of the operations phase is one of the strengths of the IS rating tool, as it gives greater coverage of an asset's life

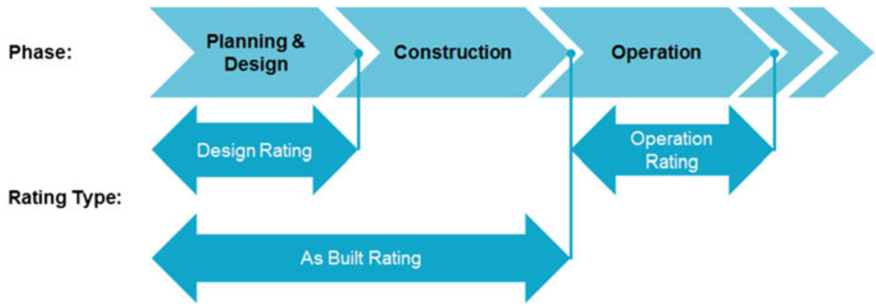


Fig. 47.2 Infrastructure life phases and IS rating types available (reproduced from Australian Green Infrastructure Council sources [9])

under a single consistent rating scheme. Figure 47.1 shows the stages that are assessable by the IS rating tool.

To align with the typical infrastructure asset life cycle phases, AGIC offers the rating types shown in Fig. 47.2. Note that the 'As Built' rating covers not just the construction phase of the asset, but also its design phase and thus looks at not only whether it was designed to be sustainability friendly, but also whether it was constructed in such a manner.

47.5 Conclusion

Infrastructure is such a vital part of the workings of society. It is all around us and involved in so many things that we do. Therefore, ensuring infrastructure is 'done' in a sustainable way is important for its contribution to sustainable outcomes and its symbolic role in a greater move toward sustainability. Over the years, sustainability has become increasingly more important to individuals, organizations, communities, and governments. Given the scale and impact of the infrastructure industry and the physical assets involved, it should not be a surprise that there is a real need for sustainability to become integral to the way the industry operates. In order to achieve real sustainability in the area of civil infrastructure projects and assets, it is necessary to perform sustainability assessment, with associated measures, metrics, and indicators, across all of the key areas and at all phases of a project's or asset's life. AGIC's IS rating tool represents a robust, comprehensive, and industry-supported means to achieve this.

Acknowledgments This work was undertaken as part of The Sustainability Assessment for Infrastructure Project and is part of the larger CRC for Infrastructure and Engineering Asset Management. The CRC for Infrastructure and Engineering Asset Management (CIEAM) would like to acknowledge the financial support from the Commonwealth Government's Cooperative Research Centres Program and the contributions from our industry, government, and university participants. The overall support provided by the CEO, Professor Joseph Mathew, is gratefully appreciated.

References

1. Urban Land Institute (2011) Infrastructure 2011: a strategic priority
2. Elkington J (1997) Cannibals with forks: the triple bottom line of 21st century business. Capstone Publishing Limited, Oxford
3. Group of 100 (2003) Sustainability: a guide to triple bottom line reporting. Melbourne
4. Stapledon T (2004) Offices as tools for organisational sustainability. PhD, The University of Sydney
5. Shaw G, Kenny J, Kumar A, Hood D (2012) Sustainable infrastructure operations: a review of assessment schemes and decision support. In: 25th ARRB conference. Perth, Australia
6. Australian Green Infrastructure Council (2012) Infrastructure sustainability rating scheme [Online]. Australian Green Infrastructure Council. Available: http://www.agic.net.au/ISratingscheme1.htm
7. Australian Green Infrastructure Council (2012) Fact sheet 1 infrastructure sustainability rating framework [Online]. Australian Green Infrastructure Council. Available: http://www.agic.net.au/fact_sheet_1.pdf
8. DTI 'Strategy for sustainable construction'—consultation events. Available: http://www.bis.gov.uk/files/file37172.pdf
9. Australian Green Infrastructure Council (2012) Fact sheet 2 rating types [Online]. Australian Green Infrastructure Council. Available: http://www.agic.net.au/fact_sheet_2.pdf

Chapter 48
Understanding the Business Case for Infrastructure Sustainability

Tony Stapledon, Gavin Shaw, Arun Kumar and David Hood

Abstract The relationship between corporate and sustainability performance continues to be controversial and unclear, not withstanding numerous theoretical and empirical studies. Despite this, views on corporate responsibilities "meet where management can show how voluntary social and environmental management contributes to the competitiveness and economic success of the company." This approach is fundamental to the business case for infrastructure sustainability. It suggests that beyond-compliance activities undertaken by companies are commercially justified if they can be shown to contribute to profitability and shareholder value. Potential public good benefits range across a wide spectrum of economic (for example employment, local purchasing, reduced demand for electricity generation), social (indigenous employment and development, equity of access), and environmental (lower greenhouse gas emission, reduced use of non-renewable resources and potable water, less waste, enhanced biodiversity). Some of these benefits have impacts that lie in more than one of the economic, social, and environmental areas of public goods. Using a sustainability rating schemes and potential business benefits from sustainability initiatives, this paper presents a brief summary of an online survey of industry that identifies how rating scheme themes and business benefits relate. This allows for a case to be built demonstrating which sustainability themes offer particular business benefits.

T. Stapledon · G. Shaw (✉) · A. Kumar · D. Hood
Science and Engineering Faculty, Cooperative Research Centre for Infrastructure and Engineering Asset Management, Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: gavin.shaw@qut.edu.au

A. Kumar
e-mail: arun.kumar@qut.edu.au

D. Hood
e-mail: davidahood@mac.com

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_48,

48.1 Introduction

Previous studies of the business case for sustainability in business generally, and the built environment in particular, have demonstrated that sustainability can have a positive impact on both short- and long-term firm performance and market value. However, little research has been focussed specifically on the benefits accruing to firms in the civil infrastructure industry. The research aims to address this shortfall through testing whether infrastructure sustainability initiatives also deliver benefits to these firms.

The paper briefly outlines the two stages of the research design that included:

1. Development of a framework that can be used to correlate sustainability performance, using AGIC's IS rating scheme, with the six key business benefits of sustainability.
2. Testing industry perceptions through an online survey and gathering case study examples of benefits arising from good sustainability performance to test the hypothesis that good sustainability performance delivers benefits to business.

The paper then provides a brief discussion of the findings, both from background literature and from the conducted online survey.

48.2 Background

Sustainability is a hot topic in the infrastructure industry, although there is a lack of understanding of what it is and how it can deliver benefits to firms. Essentially, infrastructure sustainability is the delivery of projects and management of infrastructure assets so that they meet the economic, environmental, and social needs of, and impacts on, key stakeholders. In the infrastructure business, key stakeholders may be owners, employees, the impacted community, authorities, and suppliers. Importantly, future generations are also a key stakeholder; much infrastructure is expected to have a long useful life. In infrastructure terms, sustainability is about balancing triple bottom-line trade-offs and extends beyond just addressing ecological concerns.

There have been a number of studies of the business case for sustainability both in business generally (for example, [1–3]) and the built environment in particular (for example, [4–6]). These and other studies have shown that there are five principal business benefits of sustainability (see [7]): cost savings, enhanced reputation, better risk management, stronger employee engagement, and new sources of revenue from new or expanded markets.

Confirmation of a firm's "social license to operate" is often argued as a further, sixth, benefit of corporate sustainability [8, 9]. Closely related to both risk reduction and reputation, this theory posits that society has granted the owners of companies limited liability in exchange for putting their capital at risk and receives public benefits in return [8, 10]. Porter and Kramer [9] argue that license to operate is particularly relevant to companies that depend on government consent and/or those that rely on good stakeholder relationships. Both these situations apply to the infrastructure industry.

While aspects of these studies are relevant to civil infrastructure, there has been little specific research into the benefits accruing to firms active in the ownership, construction, or operation of civil infrastructure assets. This research aims to address this shortfall.

Developing a business case for infrastructure sustainability requires an understanding of sustainability performance and the benefits that result, and the means to link the two.

It is important to note that the extent of benefits and costs of both corporate sustainability generally and individual initiatives varies across industries, business units, and projects—as well as across companies. Consequently, it is difficult to isolate critical variables that may apply generically to infrastructure. The industry embraces diverse asset types, geographies, and impacts, with a disparate set of resources required for both their delivery and use. Nonetheless, whole-of-industry performance measurement tools are now available, as is discussed next.

48.2.1 Measuring Sustainability Performance

Infrastructure sustainability performance is measured in particular practice areas, called "themes" by the Australian Green Infrastructure Council (AGIC) in its IS sustainability rating system [11]. The UK's CEEQUAL [12] and USA's envision [13] are similar performance rating schemes, albeit using different numbers of indicators and different terminology.

These infrastructure industry tools serve a dual purpose as performance measurement systems and decision-making frameworks. Subject to their particular design, they apply to all stages of infrastructure planning, design, construction, operation, and decommissioning. They all propose a set of meaningful indicators, arranged under sections or themes, and are designed to draw on and/or complement data collected in mainstream business systems. The AGIC IS system is designed to be applicable to the design, construction, and operation phases and to most types of civil infrastructure.

This research demonstrates how these tools can be used to link infrastructure sustainability performance to business benefits.

48.2.2 Measuring Business Benefit

A useful model for considering the business benefits of infrastructure sustainability is provided by Berns et al. [1]. This model, developed from a global survey and interviews with business leaders, proposes that sustainability efforts can influence all the levers companies use to create value, resulting in enhanced total shareholder return.

As the model suggests, there are a complications in measuring the benefits to business of infrastructure sustainability.

First, causal links between sustainability initiatives and corporate outcomes are often indirect and may also have ambiguous direction. Consequently, many of the benefits of sustainability initiatives do not impact financial performance directly but are intangibles that “need to be leveraged into higher cash flows” [2]. Second, the three different types of capital to be measured—economic, environmental, and social—“have different properties and thus require different approaches” [14]. Bringing intangibles and these three types of capital into the common corporate language of monetary value may not be possible. Indeed, many of these intangible assets are “valued by the market but overlooked by management” [15]. The result is that “management is prone to under-manage or even ignore what might…be the most significant portion of their company’s (market) valuation” to the extent that, in today’s knowledge economy, “traditional accounting assets still explain only about 25 % of market value” [15].

The consequence of these complications is that it is not possible for this research to directly translate some of the intangible business benefits expected from infrastructure sustainability—such as brand strength, customer loyalty, the ability to attract and retain employees, and lower risk premiums—into monetary value. Rather, it is necessary to rely on the Berns et al. [1] model and other literature to demonstrate that intangible assets leverage into shareholder value.

48.3 Key Issues in Sustainability

The research was conducted in two stages:

1. Development of a framework that brings together sustainability performance, using AGIC’s IS (infrastructure sustainability) rating scheme, and the identified six key business benefits of sustainability discussed above.
2. Gathering data from the industry to test the hypothesis that good sustainability performance delivers benefits to business that translate into improved stakeholder value.

The second stage of the research was further divided into two streams:

1. An online survey open to participants in the civil infrastructure industry, designed to measure industry perceptions of the business benefits of pursuing sustainability initiatives in each of the AGIC’s draft IS rating scheme themes: asset and project management and governance, economic performance, using resources, managing emissions, pollution and waste, ecosystems and biodiversity, people and place, and workforce. (Note that the final IS rating scheme

has slightly modified these themes and does not yet include the planned "economic" and "workforce" themes [11]).
2. A series of interviews with participants from the industry to gather data to be used in preparing case studies that illustrate how business benefits derive from particular sustainability initiatives.

These two streams will then be brought together to highlight examples of where infrastructure sustainability initiatives have resulted in business benefits to firms and how they relate to industry perceptions.

48.4 Survey Findings

The framework developed from the first stage of work is based on AGIC's draft IS rating scheme "themes" making up the columns and the six potential business benefits identified in [7–9] making up its rows.

This framework will allow firms to correlate business benefits (and costs) to specific sustainability initiatives in an industry-consistent and ordered manner. A particular advantage of this approach is that it uses sustainability performance data that businesses will be collecting as part of their normal operations when they opt to apply for an award under the AGIC rating scheme. This will minimize data collection costs while contributing to evolution of consistent industry language and a common approach to infrastructure sustainability business cases.

48.4.1 Summary of Survey Results

The second stage of research involved the execution of an online survey that invited industry to provide their view. At the closure of the survey, there were 99 valid respondents with the following breakdown:

- 61 % of respondents nominated their firm as a private company or registered business, 17 % as government, statutory authority or public utility, and 18 % as a public company.
- 65 % from firms that are either design or sustainability consultants.
- Across the spectrum of business size from less than 100 to more than 5,000 employees.
- Approximately 30 % from firms that are owners or operators of infrastructure assets.

Results of the online survey to date show that respondents consider that their firms would be most encouraged to pursue sustainability by cost savings, reduced risk, and improved image, reputation, and brand strength (Fig. 48.1).

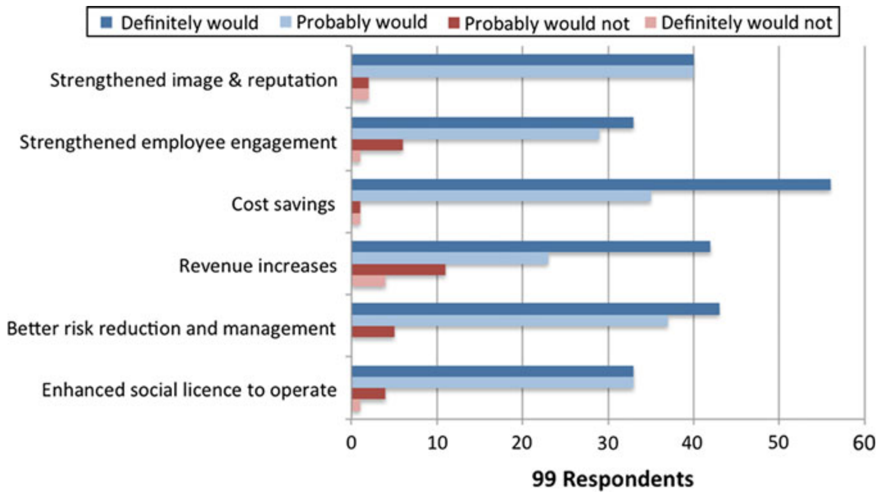


Fig. 48.1 Motivating benefits for pursuing sustainability in infrastructure

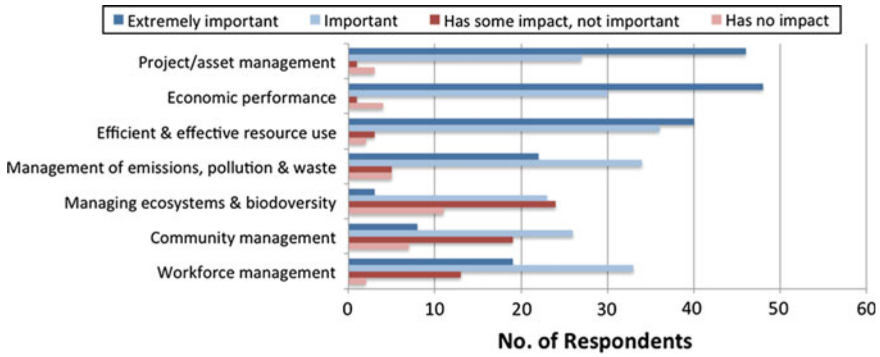


Fig. 48.2 Sources of cost savings from infrastructure sustainability

In the next questions, respondents were asked to rate how important they considered good performance in each of the IS themes to be in delivering the business benefits shown in the left-hand column of Fig. 48.1. While the research examined the full spectrum of six potential benefits, this paper briefly discusses the survey findings on the three principal motivations as identified in Fig. 48.1 due to space.

The survey respondents consider that cost savings from sustainability initiatives are primarily delivered by good economic performance, sound project/asset management, and efficient and effective resource use (Fig. 48.2).

Respondents consider that strengthened reputation results from good sustainability performance across all AGIC's IS rating scheme themes, with good

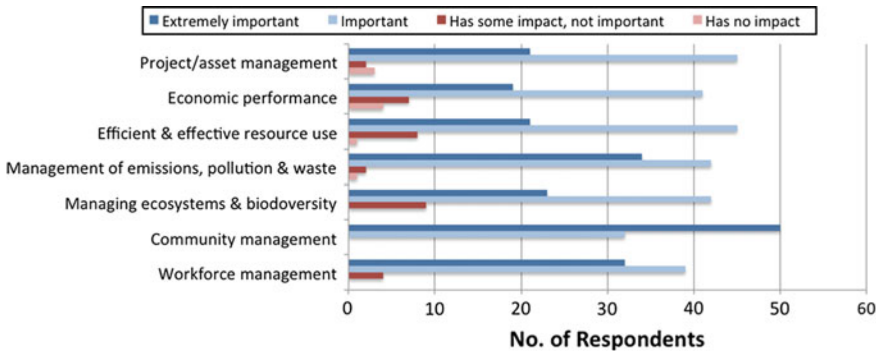


Fig. 48.3 Sources of reputational benefits from infrastructure sustainability

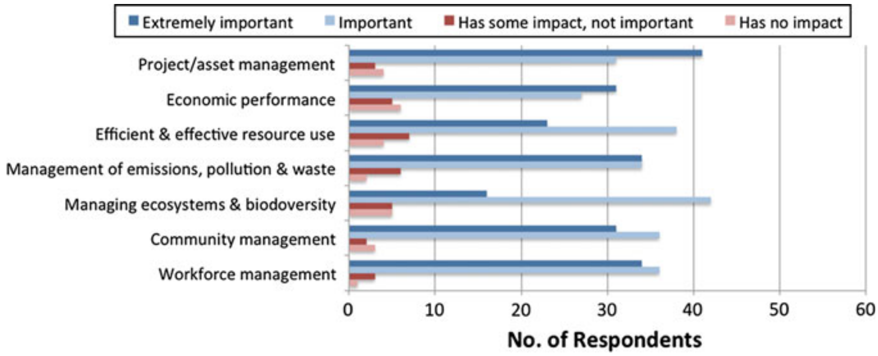


Fig. 48.4 Sources of risk management benefits

performance in community, workforce, and management of emissions, pollution, and waste seen to be particularly important (Fig. 48.3).

The survey respondents consider that risk management is improved by good sustainability performance across all themes, with good performance in project/ asset and economic management rated as most important (Fig. 48.4).

Other findings from the survey are that respondents consider that:

- Employee engagement is strengthened (understandably) primarily by good workforce management and also by good community management.
- Good performance in environmental, biodiversity and community management and in managing emissions, pollution, and waste enhances social license to operate.
- Across all the themes, only good economic management performance is important in contributing to increased revenue.

48.5 Conclusion

This research has developed a framework to correlate infrastructure sustainability performance with business benefits in an industry-consistent manner. It has also demonstrated the application of the framework through an online survey and gathering of case study examples. Data collected from the survey and case studies are being used to test whether infrastructure firms can enhance shareholder value through sustainability initiatives, as would be expected from generic studies of the impacts of corporate sustainability.

Findings reveal industry perceptions that cost savings, enhanced reputation, image and brand strength, and reduced risk with better risk management are the primary motivation for infrastructure firms to adopt sustainability. These benefits are delivered by good sustainability performance across the spectrum of possible sustainability initiatives.

Findings and outputs that support the business case for infrastructure sustainability are likely to be welcomed by the industry, which is making a significant investment in sustainability. Most infrastructure companies, irrespective of the role they play in the infrastructure life cycle, employ specialist sustainability staff, record their performance through sustainability reports, by other public documents or on their websites, and invest in sustainability initiatives both in-house and on their assets and projects.

However, while the public benefits are often described (e.g., lower carbon emissions, less potable water consumed, reduced negative environmental impacts), the benefits to the organizations themselves are often not defined or measured, meaning that their investment in sustainability may be wasteful, or at least inefficiently applied. This research helps to provide a framework that organizations can use to quantify and benchmark the business benefits of their sustainability efforts and, in future through case studies, will illustrate practical applications.

Acknowledgments This work was undertaken as part of the Sustainability Assessment for Infrastructure Project and is part of the larger CRC for Infrastructure and Engineering Asset Management. The CRC for Infrastructure and Engineering Asset Management (CIEAM) would like to acknowledge the financial support from the Commonwealth Government's Cooperative Research Centres Program and the contributions from our industry, government, and university participants. The overall support provided by the CEO, Professor Joseph Mathew, is gratefully appreciated.

References

1. Berns M, Townend A, Khayat Z, Balagopal B, Reeves M, Hopkins M, Kruschwitz N (2009) The business of sustainability. MIT Sloan, BCG, North Hollywood
2. Steger U, Ionescu-Somers A, Salzmann O (2007) The economic foundations of corporate sustainability. Corp Gov 7:162–177
3. Wilbard B (2002) The sustainability advantage: seven business case benefits of a triple bottom line. New Society Publishers, Gabriola Island

4. Bryson JR, Lombardi R (2009) Balancing product and process sustainability against business profitability: sustainability as a competitive strategy in the property development process. Bus Strategy Environ 18:97–107
5. Katz G (2003) The costs and financial benefits of green building: a report to California's sustainable building task force. Capital E
6. RICS (2005) Green value: green buildings, growing assets. The Royal Institution of Chartered Surveyors
7. Weber M (2008) The business case for corporate social responsibility: a company-level measurement approach for CSR. Eur Manag J 26:247–261
8. Bakan J, Burke T (2005) Corporate social responsibility. The Ecologist 35:28–32
9. Porter ME, Kramer MR (2006) Strategy and society: the link between competitive advantage and corporate social responsibility. Harvard Bus Rev
10. Lantos GP (2001) The boundaries of corporate social responsibility. J Consum Mark 18:595–630
11. Australian Green Infrastructure Council (2011) AGIC rating tool. Australian Green Infrastructure Council. Available: http://www.agic.net.au/
12. CEEQUAL (2011) What is CEEQUAL? Crane environmental. Available: http://www.ceequal.com/index.html
13. ISI (2011) The envision™ sustainability rating system. The Institute for Sustainable Infrastructure. Available: http://www.sustainableinfrastructure.org/index.cfm
14. Dyllick T, Hockerts K (2002) Beyond the business case for corporate sustainability. Bus Strategy Environ 11:130–141
15. Ballow JJ, Burgman R, Molnar MJ (2004) Managing for shareholder value: intangibles, future value and investment decisions. J Bus Strategy 25:26–34

Chapter 49
Advancing Risk Management in Nuclear Power Plant EPC Projects: An Empirical Evaluation of Risk Management Practices on Steam Generator Replacement Projects

Sola M. Talabi and Paul Fishchbeck

Abstract The nuclear power industry has historically been plagued with new plant deployment risks for engineering procurement and construction (EPC); project cost and schedule overruns present a risk to investors. These risks are anticipated to continue to hinder the growth of the nuclear industry, due to expected increased regulatory requirements due to the Fukushima Daiichi Nuclear Power Plant incident in March 2011. Although several risk management practices have been put in place, considerable cost and schedule excursions have continued to occur in the construction of recent nuclear power plant projects. We identify the limitations with current risk management practices by assessing the level of completeness of risk identification and accuracy of risk assessments on prior steam generator replacement (SGR) projects. SGR projects were chosen for this evaluation because their scope of work is characteristic of large nuclear power plant EPC projects.

49.1 Introduction

Engineering, Procurement, and Construction (EPC) projects for nuclear power plant construction have historically being plagued with significant cost overruns. Figure 49.1 shows that the actual cost of projects started between 1970 and 1984 were all above their estimated cost, with some projects in excess of 4 times the estimate prior to construction start. This pattern has continued with the recent wave of plants under construction, with cost overruns on some current projects

S. M. Talabi (✉) · P. Fishchbeck
Department of Engineering and Public Policy, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15101, USA
e-mail: sola@talabi.com

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_49,

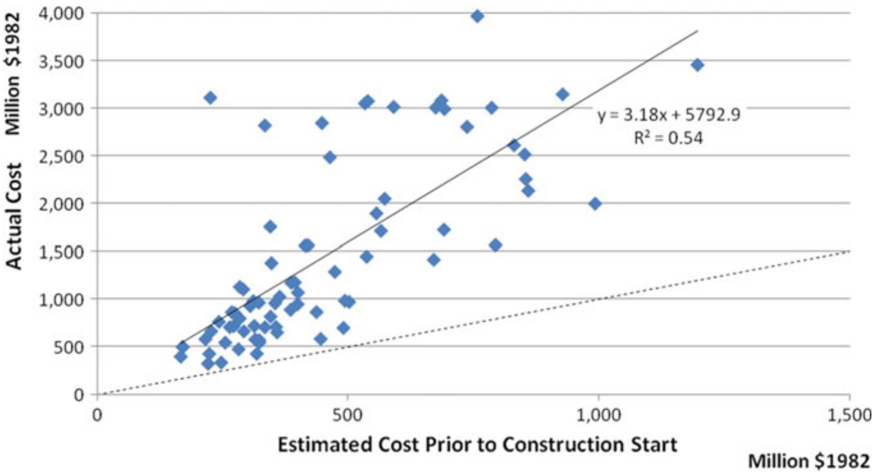


Fig. 49.1 Nuclear power plant construction cost comparison of planned to actual costs, for plants with construction start dates from 1970 to 1982. *Source* [2]

such as the V.C. Summer Units 2 and 3 project exceeding the estimated cost by over 50 %, with an increase from $4 billion at the start of construction, to an estimate at complete (EAC) of approximately $6 billion [1].

Despite over 40 years of nuclear power plant construction experience, the nuclear industry has failed to see a learning curve effect in nuclear plant construction cost. As shown in Fig. 49.2, the learning curve that was predicted by the Atomic Energy Commission (AEC) in 1964 never materialized; the AEC assumed construction costs by 1980 would be under $1,000/kWe, but as shown, some costs were more than 10 times greater than the AEC had predicted.

In order to identify the causes of cost overruns in nuclear power plant projects, the US Energy Information Administration (EIA) carried out a statistical analysis of nuclear plants completed between 1966 and 1977 [2]. The study explains that one of the most statistically significant variables that explains the variation in the effort and material required to build a nuclear power plant, as measured by the real overnight cost, is the construction period. The study further explains that although the construction period statistically explains the cost increase, it is not a cause. This is an important distinction, as both cost and schedule performance are consequences and not causes. The study makes this assertion by explaining that construction duration acts as a surrogate for a number of immeasurable factors that influence costs, such as design changes, safety, and environmental retrofits required by regulatory change, and labor productivity. We define these factors as risk event occurrences that invalidate the baseline assumptions that went into the initial planning of schedule and cost. The EIA study suggests that the risk assessment for the nuclear power plant construction was inadequate, as expressed by the lack of correlation between the expected lead times and expected cost. The EIA study makes this assertion by demonstrating that although the estimated time to complete the projects increased,

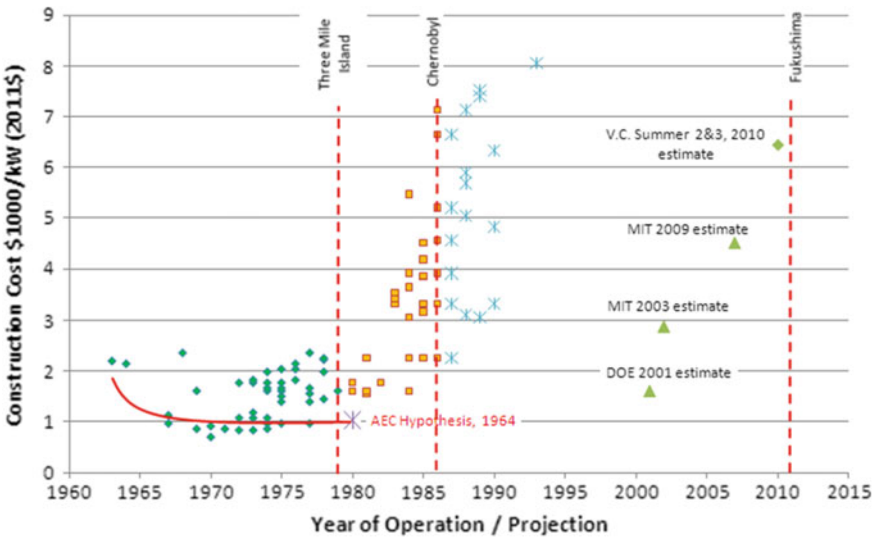


Fig. 49.2 Nuclear power construction costs, showing realized cost for plants completed before 1995, and projected costs and year of projection after 1995, figure modified from 'Nuclear Power—The Energy Balance,' Ceedata Consultancy, 2007 (Yr. 2011$)

the associated time-dependent cost estimates did not increase in a corresponding manner. Hence, there was inadequate assessment of the impacts of delays on cost overruns. The study proposed that additional work should be performed to establish the effects of managerial practices on cost overruns. This study addresses the role of risk management as a managerial practice in cost overruns for large nuclear infrastructure projects, by evaluating the historic records of risk identification and documented risk occurrences in the performance of nuclear power plant steam generator replacement (SGR) projects.

49.2 Overview of Current Nuclear Power Plant Project Risk management Practices

Nuclear power plant EPC project managers have to contend with a variety of risk types, including technological, quality, commercial, construction, supply chain, regulatory, organizational, and process. In addition, projects usually possess high levels of uncertainty derived from their often compressed schedules, inadequate or uncertain budgets, designs that are near the feasible limit of achievable performance, and frequently changing requirements [3].

To manage these risks effectively, the scope of project management has expanded to include risk management as a specialized discipline. The industry has responded to the uncertainty in nuclear power plant delivery by employing various risk management practices to reduce the uncertainty associated with nuclear EPC projects. Some of these practices include streamlining the regulatory process with the Nuclear Regulatory Commission, an example of this is the development of the inspection, tests, analysis, acceptance criteria, (ITAAC) in the case of US reactors, standardizing designs, and modularizing physical structures and components to improve construction and constructability. Along with these practical measures, the industry has also adopted robust project risk management frameworks that include dedicated personnel, processes, and tools for risk management. Regarding processes and tools, the industry has adopted the use of the Project Management Institute (PMI) standards for risk management [4], as described in the project management book of knowledge (PMBoK). Although these risk management processes have been put in place, there are still considerable cost and schedule excursions that have occurred in the construction of recent nuclear power plant projects. A recent example is the Finnish Olkiluoto 3 nuclear plant currently under construction, where there have been quality-related problems, nuclear safety culture shortcomings, schedule and cost overruns, and the resulting public concern [5].

A historic review of the development of project management showed an increase in the appreciation of a need for a dedicated project management competency post-World War II [6]. The initial formulations of project management were led by the US Department of Defense and NASA with the development of various internal guidelines and procedures that were developed due to necessity, and empirical observations [7]. Much of the early practice of project management included the use of tools and techniques such as network scheduling and performance measurement. Project management societies started becoming formalized in the late 1960s to early 1970s and began to provide professional forums for communication on the discipline.

In recognition of the significance of project management to effective nuclear project delivery, the International Atomic Energy Agency (IAEA) developed a guideline for effective nuclear project management [8]. The nuclear industry has adopted the recommendations of the IAEA, established strong project management organizations, and adopted industry standards for project management. Our observations of the practices of a major Nuclear Safety Steam Supply vendor showed the establishment of dedicated project management organizations that have adopted the PMI standards. The PMI standards include project management elements of cost, schedule, and risk.

PMI is a not-for-profit organization that provides project management standards and practices. The PMI's PMBoK was established in 1976 based on the observation that there were several project management practices that were common across projects. The PMBoK structure includes the following knowledge areas: integration management, scope management, time management, cost management, quality management, human resource management, communications management, procurement management and risk management. Our primary focus in this study is the application of PMI risk management standards for risk identification and risk assessment.

Most organizations supporting a nuclear EPC project have a formal framework for project risk management [9] which includes the people, processes, and analytical tools [10]. Examples of risk management methods as defined by PMI include risk identification methods such as brainstorming, checklists, influence diagrams, and cause and effect diagrams. Risk analysis methods include probability and impact matrices, event tree analysis, sensitivity analysis, simulation, Delphi techniques, and expert judgment. Risk-response tools include influence predictability matrix. In support of these analytical methods, there are software packages that have been developed to perform the required analysis. Some of these packages include Active Risk Manager (ARM), @Risk, Crystal Ball, PERT-Master, and several others.

49.3 Method for Assessment of Risk Management on Steam Generator Replacement Projects

49.3.1 Review of Risk Assessments and Lessons-Learned in a NSSS Vendor Firm

Risk assessments are performed prior to commencement of SGR projects. The goal of the risk assessment is to identify risks to the project before the project is started, thereby allowing for appropriate cost and schedule planning. A risk assessment includes a list of identified risks, probability of postulated event occurrence, and estimated impact assessments, given a risk occurs. Lessons-learned are performed at the end of the SGR projects. The goal of the lessons-learned process is to identify risks that occurred and affected the project's planned performance baseline in terms of cost and schedule. By comparing the expected risks (in the risk assessments) to actual risks that occurred (in the lessons-learned), we are able to evaluate the completeness and accuracy of risk assessments. We gathered risk assessment documentation on 20 SGR projects that spanned 12 years from 1999 to 2011. The risk assessment documents provided the following categories of risks: organizational and process, resources, technical, schedule, commercial, external, supply chain, quality assurance, customer, licensing, and schedule. Using these categories for each project, we documented the risks that were identified before the project started, and the risks that were documented at the end of the project. We estimated the cost of the expected risks based on stated risk values in the assessment documents. For the actual risks that occurred, the values of impacts were determined by finding the difference between the budgeted cost at the start of the project and the final reported costs. The following three quantitative measures were used to assess the risk identification and risk assessment effectiveness:

1. Count of Risks Identified and Risks Occurred: This provides a comparison between the number of risks identified for each category of risk and the number of documented risks that occurred in the lessons-learned documents.
2. Estimated Value of Postulated Risks and Actual Documented Cost of the Risks: This provides a comparison between the estimated value of the risks, given they occurred, and the actual documented cost of the risks.
3. Estimated Probabilities of Occurrence of Postulated Risks and Actual Frequency of Occurrence: This provides a comparison between the estimated value of the likelihood of the risks to occur and the actual documented frequency of occurrence.

Background on SGRs: Steam generators are used in nuclear power plants that use pressurized water reactor (PWR) technology. They are used to generate steam that turns a turbine and generates electricity in a nuclear power plant. Over the typical 40- to 60-year service life of a nuclear plant, SGs are replaced at least once because of reduced performance caused by various degradation mechanisms. SGR projects are typically performed during planned refueling outages. Typically, an outage will require 35–60 days to replace steam generators. SGR costs run approximately $150 million, but can vary significantly depending on site-specific considerations [11]. SGR projects were chosen because their scope of work is representative of a nuclear power plant EPC work scope as it includes all the major scopes of design, analysis, manufacturing, procurement, installation, construction, licensing, and plant start-up.

49.4 Results

49.4.1 Evaluation of Risk Identification on Steam Generator Replacement Projects

Figure 49.3 shows a plot of the rate of risk identification relative to risk occurrence for SGR projects. The risks are categorized by the functional groups supporting the projects (e.g., supply chain, technical). The vertical axis represents the percentage of projects for which the stated categories of risks occurred ('% Risk Occurrence') and the horizontal axis represents the percentage of projects for which the stated categories of risks were identified ('% Risk Identification'). Hence, as the figure indicates, supply chain-related risks occurred on 100 % of projects but were identified only on 55 % of these projects. All categories of risks below the diagonal line in the figure have a higher rate of occurrence than identification, and all categories above the line have a higher rate of identification than occurrence. With the objective of risk identification being the reduction of risk occurrence, we describe the categories above the line as over-estimated or well-managed. Over-estimation of risks in the categories above the line may be caused by a heuristic bias of availability where familiar issues tend to be over-estimated. These

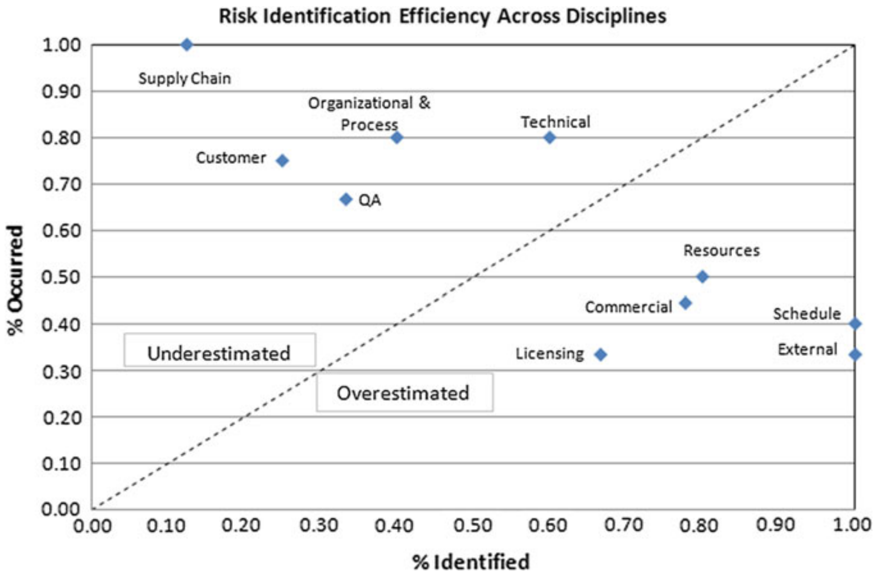


Fig. 49.3 Risk identification and risk occurrence relationship across functions

categories of risks may also be above the line because they are well-managed, thereby reducing the rate of occurrence. The results in Fig. 49.3 show a higher level of identification relative to occurrence for external, schedule, commercial, resource, and licensing risks, and a lower level for technical, organizational and process, customer, quality assurance (QA), and supply chain.

49.4.2 Evaluation of the Accuracy of Risk Impact Estimation on SGR Projects

Effectiveness of assessment of the cost of the risks is the ability to accurately estimate the cost impact of risks that actually occur. We compared the value of risks that occurred during the projects to the estimated cost of the risks that were identified for each category prior to the projects starting, as shown in Fig. 49.4. The vertical axis is the ‘actual cost of risks that occurred’ and the horizontal axis is the ‘estimated cost of risks identified’. An objective of efficient risk management is to have adequate reserves to fund risk occurrence; hence, if we assume a conservative risk-averse utility function is assumed, then the objective is for the estimated cost of the of risks to be on average equal to, or greater than the actual. The results showed that for all categories of risks, the estimated costs were less than the actual and were hence underfunded, meaning that the cost of the risks that occurred were higher than the estimated cost of the risks when identified. The

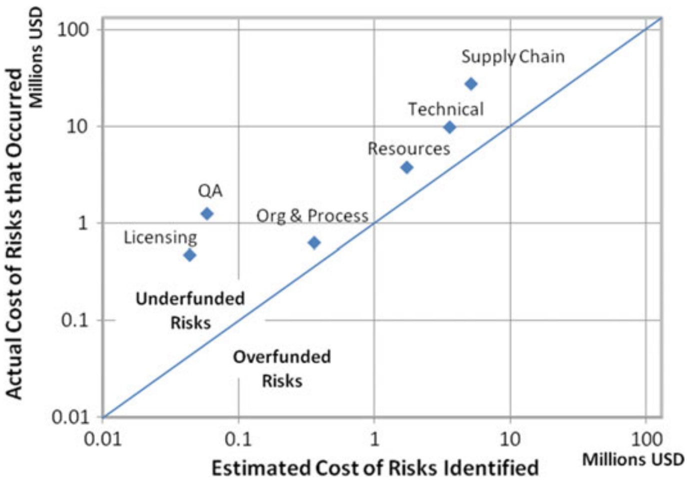


Fig. 49.4 Actual and estimated cost of risk occurrences by category

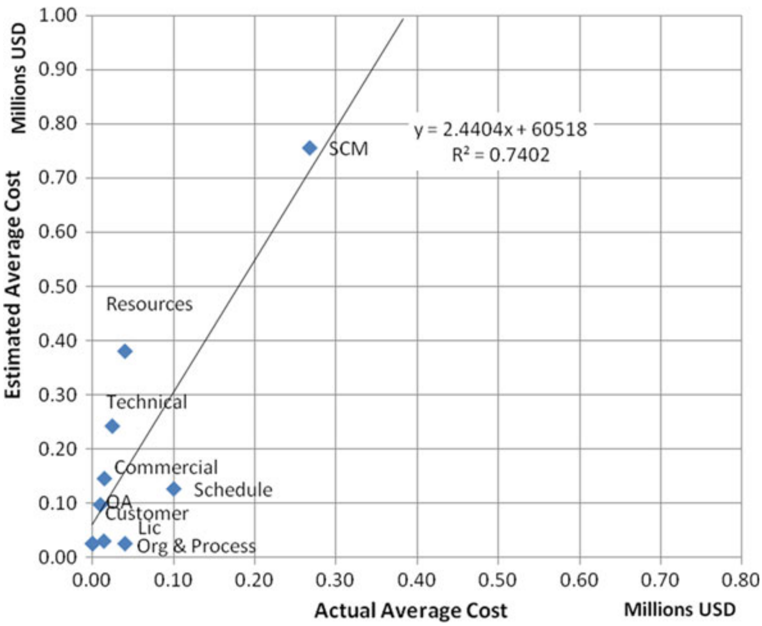


Fig. 49.5 Actual and estimated average cost of risks for each category

relationship between the cost of estimated risks and actual risks is approximated in Fig. 49.5, which is a plot of the average estimated and average actual cost of the risks in each category.

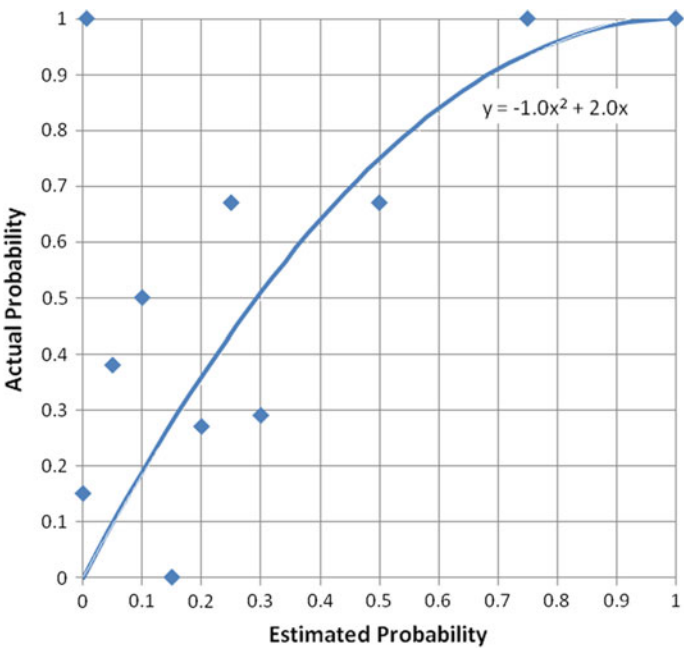


Fig. 49.6 Estimated and actual predicted probabilities for specific risks

49.4.3 Accuracy of Predicted Probability Values

This assessment was based on a selection of a subset of 71 SGR risks that were well defined in terms of properly documented probabilities of occurrence and impact, and also well-documented risk occurrences. Hence, we were able to directly align the estimated probability of occurrence against the documented frequency of occurrence. Figure 49.6 provides a chart of the 'estimated probability' (horizontal axis) values which are estimated for postulated events prior to the commencement of the project, and 'actual probability' (vertical axis) values which are calculated based on frequency of event occurrence. The 'actual probability' was calculated by dividing a count of the actual number of occurrences by the number of predicted occurrences. For example, there were a total of 27 risks that were predicted to have a value of 0 %, if these assessments were correct, none of those risks would occur. In actuality, 5 of those risks occurred. Hence, the actual probability was calculated as $5/27 = 0.19$. The relationship between the estimated probability and the actual probability is approximated by the curve in the figure.

	Estimated Impact																
Estimated Probability	10	12.5	25	37.5	50	62.5	75	87.5	100	112.5	125	137.5	150	162.5	175	187.5	200
0.01	16.9	14.5	9.7	8.1	7.3	6.8	6.5	6.2	6.1	5.9	5.8	5.7	5.7	5.6	5.5	5.5	5.5
0.02	16.8	14.4	9.6	8.0	7.2	6.7	6.4	6.2	6.0	5.9	5.8	5.7	5.6	5.6	5.5	5.5	5.4
0.05	16.6	14.2	9.5	7.9	7.1	6.7	6.3	6.1	6.0	5.8	5.7	5.6	5.6	5.5	5.4	5.4	5.4
0.07	16.4	14.1	9.4	7.8	7.0	6.6	6.3	6.0	5.9	5.7	5.6	5.6	5.5	5.4	5.4	5.3	5.3
0.10	16.2	13.9	9.3	7.7	7.0	6.5	6.2	6.0	5.8	5.7	5.6	5.5	5.4	5.4	5.3	5.3	5.2
0.12	16.0	13.7	9.1	7.6	6.9	6.4	6.1	5.9	5.7	5.6	5.5	5.4	5.3	5.3	5.2	5.2	5.2
0.15	15.7	13.5	9.0	7.5	6.8	6.3	6.0	5.8	5.6	5.5	5.4	5.3	5.3	5.2	5.2	5.1	5.1
0.17	15.5	13.3	8.9	7.4	6.7	6.2	5.9	5.7	5.6	5.5	5.4	5.3	5.2	5.1	5.1	5.1	5.0
0.20	15.3	13.1	8.8	7.3	6.6	6.2	5.9	5.7	5.5	5.4	5.3	5.2	5.1	5.1	5.0	5.0	5.0
0.22	15.1	13.0	8.7	7.2	6.5	6.1	5.8	5.6	5.4	5.3	5.2	5.1	5.1	5.0	5.0	4.9	4.9
0.25	14.9	12.8	8.5	7.1	6.4	6.0	5.7	5.5	5.3	5.2	5.1	5.1	5.0	4.9	4.9	4.8	4.8
0.27	14.7	12.6	8.4	7.0	6.3	5.9	5.6	5.4	5.3	5.2	5.1	5.0	4.9	4.9	4.8	4.8	4.7
0.30	14.5	12.4	8.3	6.9	6.2	5.8	5.5	5.3	5.2	5.1	5.0	4.9	4.8	4.8	4.8	4.7	4.7
0.32	14.3	12.2	8.2	6.8	6.1	5.7	5.5	5.3	5.1	5.0	4.9	4.8	4.8	4.7	4.7	4.6	4.6
0.35	14.1	12.0	8.0	6.7	6.0	5.6	5.4	5.2	5.0	4.9	4.8	4.8	4.7	4.7	4.6	4.6	4.5
0.37	13.8	11.9	7.9	6.6	6.0	5.6	5.3	5.1	5.0	4.9	4.8	4.7	4.6	4.6	4.5	4.5	4.5
0.40	13.6	11.7	7.8	6.5	5.9	5.5	5.2	5.0	4.9	4.8	4.7	4.6	4.6	4.5	4.5	4.4	4.4
0.42	13.4	11.5	7.7	6.4	5.8	5.4	5.1	4.9	4.8	4.7	4.6	4.6	4.5	4.4	4.4	4.4	4.3
0.45	13.2	11.3	7.6	6.3	5.7	5.3	5.0	4.9	4.7	4.6	4.5	4.5	4.4	4.4	4.3	4.3	4.3
0.47	13.0	11.1	7.4	6.2	5.6	5.2	5.0	4.8	4.7	4.6	4.5	4.4	4.4	4.3	4.3	4.2	4.2
0.50	12.8	11.0	7.3	6.1	5.5	5.1	4.9	4.7	4.6	4.5	4.4	4.3	4.3	4.2	4.2	4.2	4.1
0.52	12.6	10.8	7.2	6.0	5.4	5.0	4.8	4.6	4.5	4.4	4.3	4.3	4.2	4.2	4.1	4.1	4.1
0.55	12.4	10.6	7.1	5.9	5.3	5.0	4.7	4.6	4.4	4.3	4.3	4.2	4.1	4.1	4.1	4.0	4.0
0.57	12.1	10.4	7.0	5.8	5.2	4.9	4.6	4.5	4.4	4.3	4.2	4.1	4.1	4.0	4.0	4.0	3.9
0.60	11.9	10.2	6.8	5.7	5.1	4.8	4.6	4.4	4.3	4.2	4.1	4.0	4.0	4.0	3.9	3.9	3.9
0.62	11.7	10.0	6.7	5.6	5.0	4.7	4.5	4.3	4.2	4.1	4.0	4.0	3.9	3.9	3.8	3.8	3.8
0.65	11.5	9.9	6.6	5.5	4.9	4.6	4.4	4.2	4.1	4.0	4.0	3.9	3.9	3.8	3.8	3.7	3.7
0.67	11.3	9.7	6.5	5.4	4.9	4.5	4.3	4.2	4.1	4.0	3.9	3.8	3.8	3.7	3.7	3.7	3.6
0.70	11.1	9.5	6.3	5.3	4.8	4.4	4.2	4.1	4.0	3.9	3.8	3.8	3.7	3.7	3.6	3.6	3.6
0.72	10.9	9.3	6.2	5.2	4.7	4.4	4.2	4.0	3.9	3.8	3.7	3.7	3.6	3.6	3.6	3.5	3.5
0.75	10.7	9.1	6.1	5.1	4.6	4.3	4.1	3.9	3.8	3.7	3.7	3.6	3.6	3.5	3.5	3.5	3.4
0.77	10.4	9.0	6.0	5.0	4.5	4.2	4.0	3.9	3.7	3.7	3.6	3.5	3.5	3.5	3.4	3.4	3.4
0.80	10.2	8.8	5.9	4.9	4.4	4.1	3.9	3.8	3.7	3.6	3.5	3.5	3.4	3.4	3.4	3.3	3.3
0.82	10.0	8.6	5.7	4.8	4.3	4.0	3.8	3.7	3.6	3.5	3.5	3.4	3.4	3.3	3.3	3.3	3.2
0.85	9.8	8.4	5.6	4.7	4.2	3.9	3.8	3.6	3.5	3.4	3.4	3.3	3.3	3.2	3.2	3.2	3.2
0.87	9.6	8.2	5.5	4.6	4.1	3.9	3.7	3.5	3.4	3.4	3.3	3.3	3.2	3.2	3.1	3.1	3.1
0.90	9.4	8.0	5.4	4.5	4.0	3.8	3.6	3.5	3.4	3.3	3.2	3.2	3.1	3.1	3.1	3.1	3.0
0.92	9.2	7.9	5.2	4.4	3.9	3.7	3.5	3.4	3.3	3.2	3.2	3.1	3.1	3.0	3.0	3.0	3.0
0.95	9.0	7.7	5.1	4.3	3.9	3.6	3.4	3.3	3.2	3.1	3.1	3.0	3.0	3.0	2.9	2.9	2.9
0.97	8.7	7.5	5.0	4.2	3.8	3.5	3.3	3.2	3.1	3.1	3.0	3.0	2.9	2.9	2.9	2.8	2.8
1.00	8.5	7.3	4.9	4.1	3.7	3.4	3.3	3.1	3.1	3.0	2.9	2.9	2.9	2.8	2.8	2.8	2.8

Fig. 49.7 Error correction chart for estimated impact and estimated probability combinations

49.4.4 Establishing Error Correction Factors for Risk Assessments

We establish an error correction matrix by approximating a mathematical relationship between estimated and actual values for both probability and impact values in risk assessments. Hence, for any combination of estimated probability and impact values, we can calculate the 'actual' values, and by diving, 'actual' by estimated, we calculate the correction factors, as shown in Fig. 49.7.

EMV_{actual}: Expected monetary value calculated by multiplying 'actual' corrected probability and impact values
$EMV_{estimated}$: Expected monetary value calculated by multiplying estimated probability and impact values
I: Impact if event occurs
P: Estimated probability that postulated event occurs
$EMV_{estimated} = P_{estimated} * I_{estimated}$
$EMV_{actual} = P_{actual} * I_{actual}$
$\text{Correction Factor} = \frac{EMV_{actual}}{EMV_{estimated}}$

49.5 Discussion and Conclusion

This study has provided a quantitative assessment of the practice of risk management on major nuclear power plant EPC projects. The study specifically evaluated four quantitative measures for risk identification, impact estimation, probability assessment, and estimation error for a given combination of impact and probability. The study was performed with historic data from SGR EPC projects conducted by a major NSSS vendor. The study also includes a documentation of anecdotal information from interviews and observations by the researchers on the practice of risk management within the NSSS vendor's organization.

On the issue of effectiveness of risk identification, the study shows that there is an inverse relationship between risk identification and risk occurrence. Hence, there is evidence that the more risks are identified, the less they occur, as demonstrated in Fig. 49.3. There are two possible explanations for this trend; increased awareness of risks generates action that prevents the risks from occurring, or, there may be 'over-identification' of risks, that may not be applicable, and hence do not occur. The anecdotal evidence based on observations of the researchers suggests that both issues are present. This is based on the observation that there was very low risk identification and high risk occurrence, in the supply chain functional areas that was not a core-competency of the NSSS organization. Hence, there may be an availability bias, at play, where risks are identified more in areas where the organization focuses on, and less in areas where the organization has less focus and less expertise. To address this issue, a recommendation is offered for dedicated risk assessment workshops for each major function and category of risks. Our observations show that risk assessment workshops have been carried out from a project perspective and not a dedicated functional perspective. Hence, there within these workshops, there may be a tendency to focus the attention on issues that the organization is familiar with, which may not necessarily be the areas with the highest risk exposure.

On the issue of accuracy of risk assessments, the study shows that the assessment of cost of risks prior to occurrence is ineffective, as all categories of risks evaluated were underfunded. We observed the importance of risk assessment accuracy in the project planning process for the NSSS vendor, as these risk assessments are used to develop the project contingency funds. Recommendations to address this issue include ensuring that the development of cost estimates for risks is performed with the recognition of uncertainty. There are uncertainties in terms of the occurrence of events that can invalidate baseline expectations of cost performance, and uncertainties in terms of variability of impact of events that occur. Some of the risk assessments reviewed only captured the uncertainty of event occurrence, in terms of a probability, and used a single-point estimate for the cost impact. The use of three-point estimate techniques that capture 'minimum', 'most-likely,' and 'maximum' scenarios allows for a better representation of the uncertainty and hence leads to improved estimates.

Regarding the accuracy of probabilistic estimation, the study shows that there is an optimistic bias in the estimation, which may be caused by heuristics such as anchoring, conditioning, and over-confidence. To address this issue, heuristic biases are discussed and well-established methods such as the 'betting-bar' are used to reduce the bias. A concurrent study by Talabi et al. shows that probability estimation for nuclear power EPC risk assessments may be influenced by providing both historical information on the occurrence of the risk, and the assessment opinions of other experts (note that this study is still in progress, and this comment is based on an initial evaluation of the study results).

The study further provides a matrix for potential error correction in probability and impact assessments. This matrix provides a reference chart that shows the historic error in risk assessments based on estimated values, as compared with actual. This chart shows that the errors are higher with items that are considered to be lower probability and lower impact, and less with higher probability and higher impacts.

The limitations of this study are predominantly based on the completeness and correctness of the documented risk assessments and lessons-learned used to develop both estimated and actual probability and impact values. Another limitation is the researcher's interpretation in categorizing some of the risk assessments and lessons-learned that were not cataloged according to the organization's standard set of risk categories. To improve future studies of this type, a recommendation is made to have a standard risk register that provides a structured set of categories, risks, and risk drivers, which limits individual interpretation and variability risk definition.

Acknowledgments This paper was developed as research conducted at Carnegie Mellon University's Department of Engineering and Public Policy. The authors gratefully acknowledge the financial support provided by the Department.

References

1. International Atomic Energy Agency (IAEA) (1988) Nuclear power project-management. A Guide Book, pp 10–12
2. Kerzner H (2009) Project-management: a systems approach to planning, scheduling, and controlling. Wiley, OH, pp 57–63 (10th d. Berea)
3. Kline A, Renn O (2002) A new approach to risk evaluation and management: risk-based, precaution-based, and discourse-based strategies. Risk Anal 22(6):17–18
4. KPMG International Cooperative (2009) Construction risk in new nuclear power projects—Eyes Wide Open, pp 3–5
5. Levitt R (2011) Toward project-management 2.0. CRGP Working paper #0066:36-37
6. Morris P (2011) Updating the project-management bodies of knowledge. Center for Research in the Management of Projects, pp 23–24
7. Project-management Institute, PMI (2009) Construction extension to the PMBOK guide, Chap. 11. Risk-management, pp 34–35
8. Southworth FH et al (2003) The next generation nuclear plant (NGNP) project. Idaho National Engineering & Environmental Laboratory, Global, pp 12–23
9. Voetsch RJ (2004) The current state of project risk-management practices among risk sensitive project-management professionals. George Washington University PhD, AAT 3112236, pp 2–3
10. Walewski J, Gibson G (2003) International project risk assessment: methods, procedures, and critical factors. A Report of the Center for Construction Industry Studies, The University of Texas at Austin, pp 13–14
11. Du Y, Parsons JE (2009) Update on the cost of nuclear power. MIT, Center for Energy and Environmental Policy Research, pp 43–44

Chapter 50
Towards a Comprehensive Service Business Management Environment for Equipment Manufacturers

Juha Tiihonen, Jukka Borgman and Marko Nieminen

Abstract Transformation of equipment manufacturers to become providers of industrial services is a trend that involves deep transformation of companies. Equipment manufacturers can potentially provide services significantly more efficiently and effectively by utilizing information about their installed base (IB). We aim to recognize decision support functionality and required information contents for a system that supports service business in the context of services like 'full service', lease of maintained equipment, extended warranties, and operations and maintenance. A further supported function is product and service development. Eight cases from Finnish machine-building and telecommunications industries provide a basis for our view of business needs in decision support of operational and tactical levels of service business management. We present a vision of a decision-making tool for service business of equipment manufacturers—the service business management environment (SBME)—and identify decision-making situations ('use cases') of SBME. SBME integrates ideas and information content from engineering asset management, condition-based maintenance, standards to represent and exchange life cycle, reliability and maintenance data, and e-maintenance. If successful, SBME can support transformation of equipment manufacturers to service business. This takes place by aggregating information from whole IB, which enables better accuracy in health assessment, prognostics, cost estimation, awareness of service contract status in terms of defined KPIs, and analysis of profitability of service offerings. Furthermore, SBME supports product development by providing information on field reliability, maintainability, recognition of epidemics, and other inputs.

J. Tiihonen (✉) · J. Borgman · M. Nieminen
School of Science, Department of Computer Science and Engineering, Aalto University, P.O. Box 15400, 00076 Aalto, Finland
e-mail: juha.tiihonen@aalto.fi

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_50,

50.1 Introduction

50.1.1 Background and Motivation

Services are an increasingly significant part of the global economy. In this context, transformation of equipment-manufacturing companies (*suppliers* for brevity) to become providers of industrial services is an important trend that involves deep transformation of companies, see, e.g. [1].

The *installed base* (*IB*) of a supplier is formed by the set of delivered product individuals that are still in use [2]. Many companies offer services such as maintenance to product individuals manufactured by competitors or suppliers of complementary products. Therefore, we call this extended set of product individuals in scope of service activities *service base* (*SB*).

Primary activities towards the SB include provision of spare parts, repairs and maintenance, operation, upgrade, and replacement. These primary activities, their management and support, and ownership and financing can be allocated in different proportions to the customer, supplier, or third parties. Different service propositions may be created that divide responsibilities and risk between the related parties. In order to gain a larger share of customer's business, the service provider must be able to produce services with a better combination of cost and quality than the customer as self-service or through other partners can achieve.

Many of the activities towards SB can be performed with rudimentary information about the SB, but more extensive and better quality information—when applied effectively—can improve quality and cost of service provisioning. Simultaneously, advances in availability and price of telecommunications, and digital control and embedded intelligence of products make it increasingly feasible to collect and maintain up-to-date product individual-level information about the SB. Product manufacturers have a significant, unleashed key benefit of *fleet view* to IB, which can give a remarkable benefit over its customers or third parties. Combined with superior knowledge of products designed by the company, more comprehensive, accurate, and accessible information creates potential to provide services more efficiently and effectively compared to customers themselves or third parties. For example, a customer may have 30 similar valves in a factory and can compare their condition or maintenance records. However, the supplier can compare those 30 valves with the whole IB of that type of valves, e.g. several thousand individuals being used in various environments and in different points of their life cycles. This gives a competitive edge because the supplier is able to predict more accurately and reliably the maintenance needs.

50.1.2 Goals, Materials, and Methods

Our research in the context of eight cases from Finish machine-building and telecommunications industries shows that the state of the practice of management of information about SB has significant potential for improvement. We aim to recognize decision support functionality and required information contents for a system that supports service business of suppliers in the context of such services as ‘full service’, facilitator contracts (e.g. lease of maintained equipment), extended warranties, and operations and maintenance. This support requires adequate information about the SB—in terms of both scope and quality. Product and service development benefits from the same information.

We present a vision of a decision-making tool for service business of suppliers—the *Service Business Management Environment* (*SBME*). We identify decision-making situations (‘use cases’) of SBME in service business and product development. Our vision of business needs (and challenges) for SBME has been constructed based on both the literature and ideas inspired by the eight cases.

In order to support decision making in service business, we include decision support methods and required data contents from previous work. We take into account the literature on engineering asset management (especially data warehousing approach, e.g. [3]), Condition-Based Maintenance (e.g. [4]), ways to represent and exchange life cycle, and reliability and maintenance data [5–7]. We take the point of view of a company that manufacturers equipment and offers services to its IB, possibly offering services also to product individuals manufactured by competitors.

This paper is structured as follows. Section 50.2 is an overview of SBME; Sect. 50.3 outlines ‘use cases’: service offerings and activities where SBME provides decision support. We continue in Sect. 50.4 by outlining main information contents of SBME. Finally, Sect. 50.5 provides discussion—brief comparison with previous work, challenges awaiting implementation, future work, and conclusions.

50.2 SBME Overview

SBME aims to facilitate decision making to enable more efficient and effective service operations by optimizing the costs and value of service operations. Informed decision making is based on collecting and analyzing a wide range of information about SB. We call this information *service base information* (*SBI*). SBI includes both technical and commercial information on each product individual, see Sect. 50.4. SBME processes SBI into a form that delivers the right information, to the right person in the right place and time, in the right way—an application of the informatics approach.

Product individuals communicate regularly with the SBME to keep the corresponding information in SBME up to date, alleviating challenges of manual data collection. Furthermore, product individuals should be adequately instrumented and empowered by the intelligence and communication capabilities necessary for collecting and transmitting the relevant usage, environment, and condition data. From the data and information management point of view, essential SBME activities would cover the following:

- Acquire data from product individuals of SB.
- Acquire data from customers, their value production processes, and customer satisfaction.
- Process data (diagnostics, prognostics, and decision support).
- Provide decision support for service business management purposes.

A major functionality of SBME is to support services towards products manufactured by competitors. An extreme example from our case companies showed a maintenance site where all product individuals under maintenance were manufactured by a competitor. In another case of hundreds of thousands of product individuals in SB, about 1/3 are ‘own’ manufacture, 1/3 are manufactured by acquired competitors, and 1/3 manufactured by current competitors. Therefore, rapid modelling of competitor products is essential.

SBME is to support different personnel groups by generating actionable information (recommendations) and by providing analysis tools that make accessing relevant information straightforward. As outlined in [6], recommendations may include prioritized operational and maintenance actions, capability forecast assessments, and changes in operations, e.g. to allow a production run to complete. Some recommendations are to be acted on immediately such as notification of operators of an alert and recommended corrective actions. Other recommendations are longer term such as a warning production planning function about the high risk of failure of a production line due to a soon-to-fail critical product individual. Capability forecast assessments provide likelihood of accomplishing a specific goal, e.g. for deciding whether or not to accept certain orders or where to assign the work.

Some analyses to be supported by SBME take place on product individual level, e.g. optimization of operation or maintenance activities in case of a foreseeable failure. Other analyses concern the whole IB or a subset of it, e.g. warranty cost prediction or detection of patterns of usage, environment, early symptoms, etc., that may provide an early warning of a foreseeable failure.

50.3 Use Cases: Service Offerings and Activities Supported by SBME

Equipment is used in customer's processes to provide some function(s). How optimally these functions are achieved depends on how the equipment is used and on its condition that may limit availability, quality of output, or even safety. As a whole, operation is a highly complex domain and relevant decisions, constraints on decision variables, and optimization criteria depend on the context.

Sales of many service contracts depend on reliable estimated costs of service provisioning and risks. Among key information is estimated equipment reliability, maintainability, estimated spare part consumption, and preventive and corrective maintenance required. All these should be determined in the context of estimated usage and environmental conditions.

After a service contract has been established, service delivery needs to be managed and SBME should support related decision making and reporting both towards customer and for internal purposes. Financial performance monitoring of contracts, products, and services is required [2].

Table 50.1 summarizes the main use cases for service provisioning, some of their key performance indicators, main decisions, and potential support from SBME.

Spare parts management Based on SBI (actual usage and conditions, preventive maintenance programs for each product individual, new and expiring contracts, and location information), spare part consumption can be estimated more accurately than based on historical demand. Similarly, SBI can provide inputs to strategic design of the spare part network and tactical decisions such as determining spare part inventory levels [8]. A functionality of SBME is to identify compatible spare parts. This is seemingly easy, but complexity arises from product changes, product individual-level differences in structure caused by configuration and engineer-to-order mode of operations, discontinuation or substitution of component types, and modifications and upgrades of a product individual.

Basic maintenance contracts provide adequately skilled resources for the customer so that they perform corrective and preventive maintenance as planned and ordered by customer. The service provider is responsible for the availability and competence of workforce. Skills management and assignment of sites to regions are of central importance. Key decisions of *corrective maintenance* include allocating service jobs to field engineers and identifying necessary spare parts. Service can be expedited by informing the engineers on access requirements and by providing information to identify equipment [2].

Planning of maintenance: Planning of preventive maintenance (*PM*) *activities* is a major area of decision making to be supported by SBME. For example, it is possible to *optimize preventive maintenance intervals* based on globally aggregated relationship of usage, environment, and field reliability of components or sub-systems. In addition, field engineer travelling routes optimization is possible based on equipment locations. Effectiveness and efficiency of maintenance and

Table 50.1 Summary of use cases, their KPIs, decision, and SBME support

Service promise	Responsibility and KPIs of equipment manufacturer	Key decisions of equipment manufacturer and SBME outputs
A: I provide spare parts to you *No specific contract*	Quick delivery of correct spares parts	Spare part compatibility with product individuals, availability of spare parts, inventory levels, and locations
B: I maintain equipment for you *'Basic maintenance contract'*	Corrective and preventive maintenance planned and ordered by customer	Availability and competence of workforce. Skills management, assignment of sites to regions, and assignment of jobs to engineers
C: I plan maintenance for you '(*Predictive*) *maintenance planning contract*'	Correct advisory in maintenance planning	When to perform which maintenance activities Remaining useful lifetime (RUL) of equipment Optimization of product–individual-specific PM schedules Remote and condition monitoring. Fault and maintenance event recording and analysis.
D: I take the risk of equipment failures from you '*Extended warranty contract*'	Minimum failures of equipment	Managing risk of equipment failures. Is equipment used as agreed? Is equipment maintained as agreed?
E: I guarantee availability of your equipment by adequate maintenance '*Full maintenance contract*'	Short and infrequent failures → availability	Decisions of A, B, C, and D. Availability. Cost of service provisioning (work and spare parts)
F: I operate and maintain equipment for you '*Operations and maintenance contract*'	Maximum economic throughput of equipment	Align equipment economic and technical performance with customer processes What we do when a product failure is foreseeable? Overall equipment effectiveness (OEE). See also A, B, C, D, and E

maintenance programs can be evaluated based on KPIs that are based on failure and maintenance data. Example KPIs provided in [7] include field reliability and availability of equipment, failure rate and hazard rate, maintainability and repair rate, adherence to preventive maintenance programs, and cost of maintenance. *Remote monitoring* enables experts of the supplier to support the customer's value production processes. For example, troubleshooting, optimization of operation, and support for evaluation of equipment condition may be provided. An example is (*predictive*) *maintenance planning contract*, where the service provider advices on timing of maintenance based on condition and individual preventive maintenance plans. Timing can be based on remaining useful lifetime (RUL), condition indicators, preventive maintenance program criteria, records of failures and alarms, and required or estimated usage of equipment. For example, RUL simulations with real options approach can be applied for decision making for timing of maintenance [9].

Extended warranty (*EW*) is an agreement to extend a base warranty provided at time of purchase of the equipment. EW may include conditions, such as use of original spare parts or a maintenance contract between a customer and the manufacturer. In order to manage EW costs and profits, the manufacturer needs to monitor product reliability and related usage.

Full service a.k.a. performance contracts In a performance contract, a supplier usually guarantees a certain availability of equipment owned by the client and is responsible for the design of the maintenance concepts and the planning and control of the maintenance activities [10]. The contracts are complex to define and require significant trust and mutually agreed ways of measuring performance such as availability. Optimization of preventive maintenance intervals, optimization of spare part inventories, applying condition monitoring, etc., are of interest to constrain costs. Experiences from and information on global IB can increase accuracy of these decisions.

In a *facilitator contract*, the client uses systems owned and maintained by the supplier. The customer pays for usage, e.g. in terms of time or units of production. Lease contracts are a well-known example of facilitator contracts [10]. Different KPIs from full service may apply.

In *results-oriented product service systems* [11], the key is selling a result or capability where the producer maintains ownership of used equipment and the customer pays only for the agreed results or capability. '*Operations and maintenance*'-type service contracts differ from results-oriented product service systems in that the customer retains ownership of the equipment. Supply of materials and other production inputs can be allocated to customer or the service provider. According to [12], some typical decisions include performing corrective maintenance, compensating equipment, executing an intervention ('a field operation'), and performing no service. Compensating the equipment involves modifying the operational settings or adjusting the degraded or faulty system, sub-system, or component so that the equipment can continue operating safely but at reduced level of performance. An intervention could be to instruct an operator based on analysis that reveals an incorrect or sub-optimal control setting. Performing no service in the case of a degraded component implies running the equipment to failure, or may allow a production run to be completed before full failure. A key decision concerns *timing of maintenance activities that require shutdown or operation at reduced capacity* such as major overhauls or required inspections. Scheduling may take into account condition indicators, PM program criteria such as time, amount of usage, occurred stress, maintenance history, and records of failures, and alarms. Another major input is required or estimated usage of equipment, for example to time operations to a period of low demand or to avoid critical periods. Further optimization is possible by allocation of additional individual preventive maintenance activities to these maintenance breaks by deferring or advancing them in an informed way. One more main decision is to *determine maintenance strategy* for types of equipment or even for product individuals—run-to-failure, preventive maintenance (and its interval), or condition-based maintenance. This decision depends, among others, on equipment criticality, effect of maintenance on reliability, cost of a failure, and cost and effectiveness of preventive maintenance, and cost, availability and effectiveness of condition-based maintenance. For example, plotting and analyzing equipment reliability since repairs/maintenance might reveal that preventive maintenance does not increase reliability of specific equipment. Thus, change of strategy into run-to-failure might be appropriate [13].

Product development SBI can be utilized in product development at least in two ways. First, specifications and design details can be more precisely determined based on actual usage patterns, environmental factors, stress, and failure history of a current product generation. Field reliability and maintainability of existing systems and components can help to prioritize product changes. SBI makes it possible to identify and analyse needs of existing customers in the context of a specific product, service, or application. SBI can also support proactive customization of product and service offers and performance monitoring [2]. Second, identification of a developing 'epidemic' within a product type can be supported by SBME. In these cases, unexpected corrective maintenance is needed repeatedly among individuals of a product type. The company needs to determine why the product individuals fail and whether it is really an epidemic or just product abuse. The number of expected repairs can be estimated as well as possibly escalating warranty costs. If the product type is sold in large quantities or if it is used in critical processes, it is essential for the supplier to find out the root cause. SBME can support identifying the root cause by enabling comparison of data of faulty product individuals to find out whether the cause is the product itself, usage, specific environmental conditions or a quality problem of a specific component supplier, etc.

In addition, potential use cases of SBME stem from management of 'normal' warranties (e.g. fraud detection), HR activities (required skills development and sizing of field workforce), and sales (opportunity detection for new sales, upgrades and modernizations, spare parts, and analysis of needs).

50.4 Main Information Contents of SBME

Service Base Information (*SBI*) collected by SBME contains current, historical, and also future (planned, predicted) information on product individuals that form the IB. SBI contains technical (basic information, operational, and resource data) and commercial (customer, service contracts, cost, and value) points of view. Main views of SBI are discussed below. In some of these, alternative depths and level of detail are applicable, and some points of view are not always needed. Furthermore, some information may only be available in the context of a service contract. Metadata on information quality is required. Note that SBME needs and aggregates information on several levels ranging from product individuals through site or fleet level to global IB level.

As usual, we distinguish between types and individuals, although the prototype model approach [14] may be adopted to add capabilities of modelling. Note that SBME models also product individuals made by competitors, including structure, usage, and maintenance events to provide similar support than for 'own' products. Less detail in terms of structure may be available.

As-maintained compositional structure with *roles* reflects the current compositional structure of a product individual on the level of component individuals. Whenever a component individual is changed for any reason, this change is reflected in the as-maintained structure. A component individual in a product individual occupies a specific *role*. The concept of role is similar to (a low-level) *functional location* of SAP or *segment* of MIMOSA OSA/EAI [5], but roles are utilized also at a detailed level of product structure. For example, 'Intake_filter' of a compressor individual is a role. Originally, this role could have been occupied by individual C#123 of type INTKF791 and then by an unknown component individual (e.g. unrecorded change), and currently from a specific point of time (say, T3), it is occupied by component individual C#936 of type INTKF795. Roles serve multiple purposes. The most fundamental of these is that the as-maintained structure can be represented as an association between a role of a whole-component individual and a part-component individual during a specific period of time. This time-based approach enables collecting information about useful lifetime and reliability of components. Additionally, a role may be used to associate information about component types that can occupy the role, i.e. eligible spares. This information can be carried over from type level, but changed locally as needed. Components in the as-maintained structure are individuals—each component individual has its own unique identity. This identity remains even if the component individual is removed from a product individual, possibly refurbished, and then installed to another product individual. Thus, it is possible to track the 'life cycle story' of a component individual, e.g. number of refurbishments, successful and unsuccessful identification of faults, material or testing certificates, etc. In practice, it seldom makes sense to store the as-maintained structure as individuals down to the 'bolts and nuts' level. It is a common industrial practice that no 'machine-readable' record is created when component individuals are replaced with spare parts. It is therefore impossible to determine field life of most component individuals and corresponding component types. We suggest to support time-stamped history of component individual changes—updating this is less tedious than information for serial number tracked component individuals. Furthermore, it should be possible to include indications of uncertainty—it should be possible to record a change with no clear indication of time or a rough estimate. Finally, many companies face competition from third-party spare parts. Information about competitive offerings can be associated with type level with roles.

Failure and maintenance data definitions of ISO 14224:2006 [7] can be applied as a basis for SBME:

- Failures are *associated with product individual and component that failed*. Actual failure data includes the following:
- *timestamps*,
- *failure impact*,
- *operating condition at failure* (e.g. running, start-up, testing, idle, and standby).
- *Failure mode* is the effect by which a failure is observed on the failed item such as desired function is not obtained (e.g. failure to start). Applicable failure modes are specific to each type of equipment, but [7] defines 3 main failure modes and numerous sub-modes.
- *Failure mechanism* expresses the physical, chemical, or other processes which have led to a failure. Again, 6 main modes and numerous sub-modes are specified.
- Sometimes, also *failure cause* (root cause) is determined.
- Finally, 10 *failure detection methods* are coded, e.g. periodic maintenance or production interference.

Maintenance data [7] is recorded in case of both preventive and corrective maintenance events. Each maintenance event includes the following:

- *identification data* (e.g. equipment and items and related failure if any),
- *timestamps*,
- *maintenance category* (preventive (testing/inspection, condition monitoring, and periodic) and corrective),
- *maintenance activity category* (e.g. replace, repair, and inspection),
- *impact of maintenance on operations* (zero, partial, and total),
- *maintenance resources usage* includes maintenance man-hours per discipline and utility resources applied,
- *maintenance times* include active maintenance time, downtime, and issues that extended the time used.
- We agree with [2]—maintenance data should include person(s) who performed the activities to identify experts or colleagues that can provide support.

Operational, *monitoring*, *and environment data* come from digitized sources including sensors and instruments, the controller of equipment, laboratory equipment, and manual inputs [12]. Here, *alarms and alerts*, *conditioning monitoring data*, and *usage data* are captured. Usage data include how (control parameters and relevant process parameters) and how much the product has been used, including start-up, shutdown, and idle periods. Time series of relevant parameters are made available to SBME, either directly, as censored data, or as locally calculated features and statistics and some only on demand.

Failure logic information created with product development time techniques such as cause–consequence trees and failure mode and effect analysis (FMEA) can provide support to life cycle management in the SBME. Some failure modes are related to specific patterns of usage (e.g. extreme loads in specific conditions).

Often, these can be predicted only through experience such as occurred failures. Global management of SBI has the largely unexploited potential to reveal such patterns and to enable prediction of potential failures in cases that match the pattern. This allows reacting in a pre-emptive manner before negative effects of a breakdown occur.

Service contracts are modelled to be aware of performance indicators in terms of promises made to customers (e.g. availability and response times), contract period, costs, and other internal key performance indicators (KPIs). *Customer's value production process* is modelled sufficiently to understand the role of product individuals, their performance and availability in customer's value production. *Customer stakeholders and their roles* may be complex, and several customer stakeholders may be related to a product individual. Examples of stakeholder roles include owner and user. When these are organizations, responsibilities may be further allocated to several stakeholders. Decision making is often separated from users or maintenance technicians and delegated to numerous stakeholders.

Site and location or fleet information describes physical, logical, or geographical locations and is associated with product individuals. Sites may be associated with an owner, operator, or other customer stakeholder roles. Physical location where a piece of equipment resides at a (fixed) site is often identified with a 'tag number', code of the physical location on a grid (e.g. B7), and/or functional location such as process ID (PID). Locations are independent of product individuals occupying them. When the IB consists of moving equipment, such as ships, it may be challenging to know where each product individual is physically located, but membership in a fleet or association with customer (owner, operator) is normally maintainable. The customer may make the schedule of a vessel known to allow for planning of service events, such as repair in the next port. Access information is tied to location, including access procedures, safety procedures, and instructions for reaching the location.

Service tasks specify job contents ('recipes') in terms of amount and type of human resources required (including adequate competences, such as technical qualifications and skills, experience, and language). Furthermore, spare part needs and special resources are indicated. Performing a task triggers updates to the as-maintained structure and creates appropriate maintenance events.

50.5 Discussion

Related work Our work differs from previous work that often takes the point of view of an asset owner (site, plant, or fleet owner). SBME's main functionality relates to the three highest blocks of the architecture of ISO 13374 'Condition monitoring and diagnostics of machines. Data processing, Communication, and presentation', see [6]. These blocks combine monitoring technologies in order to assess the current health of the machine, predict future failures, and provide recommended action steps to operations and maintenance personnel.

We aim to model SBI on a well-founded conceptualization that takes into account existing standards. In view of [2], central pieces of SBI are records on items, customer locations, and service events. Mimosa OSA-EAI [5] presents a broader technical view that implements ISO 13374-2 [6]. OSA-EAI represents measurement data, concrete maintenance plans, diagnosis results, and equipment identification. It does not provide means to perform analysis, and the business view to SBI is absent. Asset management literature includes the business view, [3] presents one of the most comprehensive conceptualizations, but is limited in terms of presentation of the technical view.

Challenges Manual collection of data about IB is challenging. For example, a current industrial challenge is maintaining as-maintained structures up to date, because component changes are often not registered, and their visibility to supplier is poor. Even if personnel of the supplier perform these activities, a 'computer-understandable' coded event is usually not created. Similarly, maintenance and failure events are not registered or are registered in a way that leaves many aspects unclear. For example, in some of our cases, it is common that about half of failure classification codes are 'other' or 'miscellaneous'. In our experience, currently, the utilization of SBI is hampered by scattered availability, insufficient quality, and lacking comprehensiveness.

Future work Implementation of SBME is a major undertaking, and stepwise implementation of the vision is required. The large scope of information causes challenges, and issues of 'big data' become relevant due to the potentially huge amount of SBI. Currently, SBME-related information may be scattered in several (e.g. 5–10) systems, and some of it is not available at all. SBME will not replace these systems. There are numerous challenging aspects of integration.

Conclusions We presented a vision of SBME—a comprehensive decision-making tool to be developed from the point of view of service business of equipment manufacturers. The requirements originate from experiences from eight cases, and the vision utilizes a wide range of information to provide decision-making support using state-of-the-art methods. The SBME vision promises significant gains on efficiency and effectiveness of service operations of equipment manufacturers gained through better utilization of fleet view of the SB. Realization of the SBME concept requires significant future work.

Acknowledgments We thank TEKES, FIMECC, and related companies for financial support; this work has been funded by TUPASU research project and FIMECC SHOK program Future Industrial Services (FutIS). We also express our gratitude to companies that have offered us access in the context of TUPASU and FutIS research.

References

1. Oliva R, Kallenberg R (2003) Managing the transition from products to services. Int J Serv Ind Manag 14(2):160–172
2. Ala-Risku T (2009) Installed base information: ensuring customer value and profitability after the sale. Doctoral dissertation. Industrial Engineering and Management, Helsinki University of Technology

3. Mathew AD (2008) Asset management data warehouse data modelling. Doctoral dissertation, Queensland University of Technology
4. Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech Syst Signal Process 20(7):1483–1510
5. Mimosa (2007) MIMOSA's open system architecture for enterprise application integration (OSA-EAI) version 3.2 technical architecture summary
6. ISO (2007) ISO 13374-2:2007 condition monitoring and diagnostics of machines—data processing, communication and presentation—Part 2: data processing
7. ISO (2006) ISO 14224:2006(E) petroleum, petrochemical and natural gas industries—collection and exchange of reliability and maintenance data for equipment
8. Dekker R, Pinçe Ç, Zuidwijk R, Jalil MN (2013) On the use of installed base information for spare parts logistics: a review of ideas and industry practice. Int J Prod Econ 143(2):536−545
9. Haddad G, Sandborn PA, Pecht MG (2012) An options approach for decision support of systems with prognostic capabilities. IEEE Trans Reliab 61(4):872
10. Martin HH (1997) Contracting out maintenance and a plan for future research. J Qual Maintenance Eng 3(2):81–90
11. Baines T, Lightfoot H, Evans S, Neely A, Greenough R, Peppard J et al (2007) State-of-the-art in product-service systems. Proc Inst Mech Eng Pt B J Eng Manuf 221(10):1543–1552
12. Dausch M, Hsu C (2006) Engineering service products: the case of mass-customising service agreements for heavy equipment industry. Int J Serv Technol Manage 7(1):32–51
13. Moseley S (2012) Asset management basics. Maintworld 4(1):6–8
14. Peltonen H, Männistö T, Alho K, Sulonen R (1994) Product configurations—an application for prototype object approach. In: Tokoro M, Pareschi R (eds) Proceedings of object-oriented programming. 8th European conference, ECOOP'94, Bologna, Italy, 4–8 July 1994

Chapter 51
Engineering Asset Life Span Evaluation Using Logistic Regression

Amy J. C. Trappey, Charles V. Trappey and Wan-Ting Tsao

Abstract Large-sized engineering assets, such as electrical transformers, are critical parts of the power supply chain. This research identifies the key parameters influencing transformer optimal operating conditions and asset life span management. Engineering asset research has developed few life span-forecasting methodologies for managing transformers in operation with respect to their maintenance, repair, and replacement policies. Using logistic regression (LR) models, this research develops an innovative transformer life span-forecasting approach and verifies the models with data from series of 161-kV transformers to improve the accuracy and efficiency for industrial applications. The transformer supplier and their stakeholders benefit from reliable and accurate asset life prediction, which enhances the reliability of decision making for engineering asset management.

51.1 Introduction

The tasks and processes of engineering asset management (EAM) are defined as distributed and collaborative due to disbursed asset locations and maintenance chain complexity. Traditional EAM often relies on experiential rules of thumb and lessons learned from periodic repair and maintenance increasing the difficulty to prevent emergency shutdowns and costly interruptions of services. In particular, for the electrical power generation industry, power outages cause great dissatisfaction with the public and potentially devastating losses to the manufacturing industry. Thus, to ensure a stable and reliable power supply is a critical issue for scholars and society. Chaidee and Tippachon [1] used a systematic record of

A. J. C. Trappey (✉) · C. V. Trappey · W.-T. Tsao
Department of Industrial Engineering and Engineering Management,
National Tsing Hua University, Hsinchu 30013, Taiwan
e-mail: trappey@ie.nthu.edu.tw

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_51,

scattering failure statistics and their analysis to determine critical risk components of transformers. The authors then estimated the life time of the critical components to prevent failure and support maintenance activities. Chen and Egan [2] used Perks' hazard function to model transformer aging and then applied the Bayesian method to choose the best fit distribution allowing engineers to estimate the service lives of large electric transformers and improve the regional transmission expansion-planning process.

In this research, we study power transformers as the asset of interest due to its critical role in power supply system. Large transformers may be set in remote areas, so their health conditions may not be detected immediately even though routine repairs and maintenances are conducted. There is a lack of emergency response, flexibility, and maintenance capability to prevent sudden shutdowns. Thus, this research uses the large transformer's existing condition data to predict the transformer's operating status. Real-time data including the amounts of combustible gases and furfural concentrations generated in the insulating oil, parameters associated with the transformer condition, and installation information are collected. These device data are used to judge the health status of the transformer and predict its remaining life. Thus, decision makers can monitor and maintain these engineering assets more effectively, prevent unexpected failures, and extend the life of a transformer toward optimal usage.

51.2 Literature Review

In this section, the research literature related to insulation deterioration of oil-immersed transformers and life-forecasting applications of logistic regression (LR) are reviewed and described.

51.2.1 Insulation Deterioration of Oil-Immersed Transformers

Oil-immersed transformers are made of materials including copper, aluminum, silicon, steel, stainless steel, insulating oil, insulating paper, pressed cardboard, and other related materials. Of all the materials, insulating paper and insulating oil have distinct rates of deterioration over time. After a period of normal operation, a transformer and its internal insulation will gradually deteriorate. The causes of deterioration are summarized as follows.

1. Deterioration of the insulation paper: When there are abnormalities during transformer operations, the insulation paper discharges gases such as CO, CO_2, H_2, and C_2H_4. When the insulation paper is heated under low temperature for long periods of time, CO and CO_2 are the primary gases discharged. Further, the insulation paper will produce H_2, C_2H_4, and other indicator gases during high

temperatures. Because these gases are flammable, they can cause the transformer to malfunction and even explode. The rapid increase of these gases is a degenerative indicator definitely linked to transformer performance and life.

2. Deterioration of the insulating oil: The insulating oil also decomposes and releases various types of gases, including CH_4 at low temperatures and H_2 at high temperatures. If the insulating oil has an electrical current, H_2, C_2H_2, C_2H_4, and CH_4 may also be produced [3].

Dissolved gas-in-oil analysis is considered the most effective diagnostic approach for checking a transformer's health status. At present, the method of calculating the content ratio and the types of gas-in-oil is based on the Rogers and Doernenburg IEEE standard [4] and the Duval triangle IEC standard [5].

Since the insulation paper cannot be replaced and has a deterioration effect, many researchers have established a baseline for the expected life of a transformer based on the condition of the insulating paper. Stebbins et al. [6] developed a prediction model using furanic compounds and the degree of polymerization of the insulation paper to estimate the remaining life of transformers. Pradhan and Ramu [7] used carbon monoxide, carbon dioxide, 2-furaldehyde, total furan contents, 2-acetylefuran, and 5-methyle-2-furfural to calculate the aging of insulating paper to estimate transformer life. Ariffin and Ghosh [8] used CO, CO_2, CO_2/CO, and 2-furaldehyde to establish a nonlinear regression model to estimate the degree of polymerization of insulation paper and calculate the remaining life of 33/11-kV transformer insulation paper. Wouters et al. [9] also proposed a life model based on the degree of polymerization of insulation paper. Therefore, this paper extends the previous research theory by using combustible gases and furfural concentration as variables to establish the life model of the oil-immersed transformer.

51.2.2 Application of Logistic Regression

Logistic regression is known as a generalized linear regression model for fitting a model to a data set which may be continuous, discrete, or of mixed variables. Trappey et al. [10] applied a simple logistic model to predict RFID technology maturity based on the cumulative numbers of RFID patents filed over a period of years. The cumulative numbers of patents are plotted to fit an S-curve, which is used to explore the life cycle of technology trends. Since the S-curve has symmetry at the inflection point, the inflection point may be used to identify and predict the maturity of research and development strategies. Yan and Lee [11] used wavelet packet decomposition and Fisher's criteria to extract critical features. By inputting the features into LR models, performance is evaluated to identify possible failure modes of an elevator door system. Since LR allows for binary dependent variables, this paper uses LR to predict the remaining life of a given transformer based on the values of dependent variables under normal or abnormal conditions.

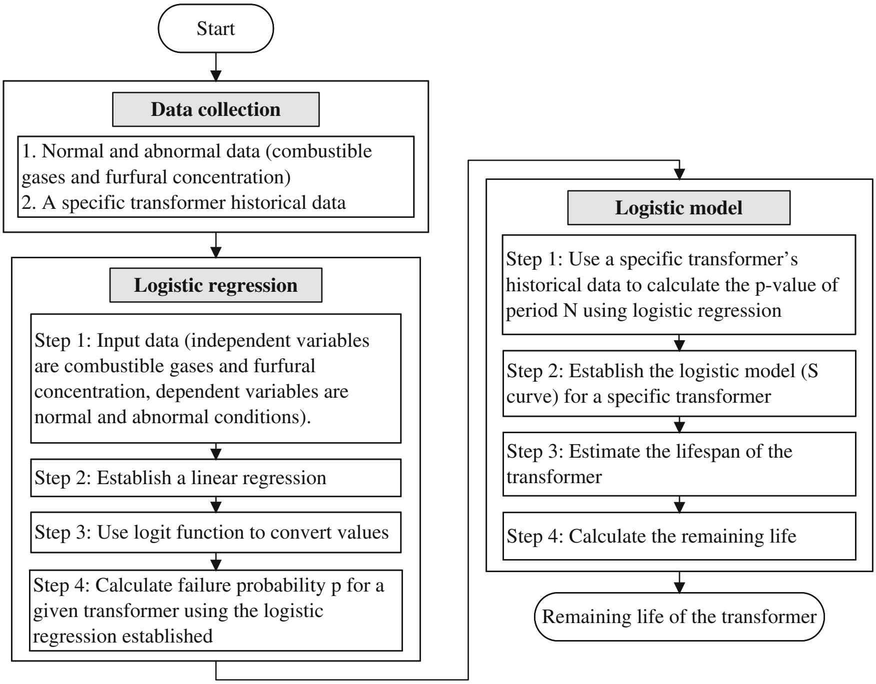


Fig. 51.1 The process of evaluating a transformer's remaining life

51.3 Methodology

In this study, the normal and abnormal operating condition data of a transformer are fit to a LR model. The LR calculates the failure probability value *p* for a set of data and then computes the past *p* value for a given transformer. The logistic model is used to estimate the transformer expected life as shown in Fig. 51.1.

51.3.1 Logistic Regression

Logistic regression as proposed by Berkson [12] is similar to regression analysis but treats the dependent variable as a binary variable. The advantages of LR are managing two categories (nominal variables) of dependent variables that can predict the odds ratio of events. The LR curve yields an S-function for the analysis of data as a probability model, and it generates output variables as 0 or 1 ("normal" and "abnormal"). Further, the odds ratio is defined as the probability that one thing will happen divided by the probability that it will not happen. If the probability $P(Y) = 0.5$ is the cutoff value, then values above 0.5 are identified as 1 (abnormal).

When the dependent variable is a binary variable, then the result is either normal ($Y = 0$) or abnormal ($Y = 1$), and the general linear regression model Y is expressed as the function:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \tag{51.1}$$

where X_k represents the kth item of value from the variables. However, using the general linear regression model to predict binary dependent variables, the dependent variable estimate may fall outside (0, 1) and not match the actual situation. Therefore, a LR transfer function known as the logit function, $\ln\frac{p}{1-p}$, is used to convert the result to a linear solution, as shown in Eqs. 51.2 and 51.3:

$$g(\mathbb{X}) = \ln\left[\frac{p(\mathbb{X})}{1 - p(\mathbb{X})}\right] = \sum_{i=0}^{k} x_\mathrm{i}\beta_i, \quad i = 1, 2, \ldots, k \tag{51.2}$$

or

$$p = \frac{1}{1 + \mathrm{e}^{-g(\mathbb{X})}} \tag{51.3}$$

where p is the failure probability, the failure probability divided by normal probability $\frac{p}{1-p}$ is called the odds ratio, and $\ln\frac{p}{1-p}$ is called the odds logarithm. In other words, let $\mathbb{X} = (x_1, x_2, \ldots, x_k)$, and assume $\mathbb{X}$ is failure probability belonging to the logistic distribution. Then,

$$p(\mathbb{X}) = p(Y = 1|\mathbb{X}) = F(g(\mathbb{X})) = \frac{1}{1 + \mathrm{e}^{-g(\mathbb{X})}} \tag{51.4}$$

where $g(x)$ and $p(x)$ are positively related and the logit function converts the output p value to (0, 1).

51.4 Model and Life Span Evaluation

In this section, we use historical data collected from 121 transformers to establish the LR. LR calculates the historical p value of period N for a specific transformer, and then, the established logistic model is used to estimate the remaining life of the transformer.

Table 51.1 The accuracy of the logistic regression applied to the 161-kV transformer series

		Number of predictions		
		Normal	Failure	Accuracy
Number of observations	Normal	614	9	
	Failure	33	23	
	Overall accuracy (%)			94

*When $p < 0.5$, the transformer condition is normal, and $p \geqq 0.5$ represents failure

51.4.1 Logistic Regression Model

From 121 operating units of 161-kV transformers, there are 679 data points (including 623 normal data points and 56 failure data points) used to fit the LR model. The independent variables use 9 combustible gases, the sum of combustible gases, 5 furfural concentrations, and the sum of furfural concentrations. These variables are oxygen (O_2), nitrogen (N_2), carbon dioxide (CO_2), carbon monoxide (CO), hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2), total combustible gases (TCG), 5-hydroxymethyl-2-furaldehyde (5-HMF), 2-furaldehyde (2-FAL), 2-furfuryl alcohol (2-FOL), 2-acetylfuran (2-ACF), 5-methyl-2-furaldehyde (5-MEF), and total furanic concentrations (TFC). The dependent variable is a binary variable defining normal ($Y = 0$) or abnormal ($Y = 1$) conditions. Using predictive analytics software (PASW), only six condition variables are used for backward (conditional) linear regression as shown in Eq. 51.5. Further, comparing real observation failure modes with the corresponding prediction outputs, the accuracy of the LR method reaches 94 % (Table 51.1). Table 51.2 shows the partial data set from a series of 161-kV transformers. The independent variables (column A–P) and dependent variables (column Q) are used to build the LR model as shown in Eq. 51.5. Equation 51.5 calculates p values (column Y), and then, the predictions (column X) can be determined by comparing the threshold 0.5.

$$\begin{aligned} g(\mathbb{X}) &= -4.497 - 0.016[C_2H_4] + 0.685[C_2H_2] - 0.008[CO] + 0.01[TCG] \\ &\quad + 0.244[2-ACF] - 0.564[5-MEF], \\ \mathbb{X} &= (C_2H_4, C_2H_2, CO, TCG, 2-ACF, 5-MEF) \end{aligned} \tag{51.5}$$

51.4.2 Life Span Evaluation for a Specific Transformer

From a given transformer T1, the failure probability p values are determined using LR over time as shown in Table 51.2 (column Y). Using the p values (column Y) and age (column Z) as input to PASW, the best fit model (Eq. 51.6) is evaluated

Table 51.2 Some 161-kV transformer series data samples, e.g., combustible gases (ppm), furfural concentrations (ppb), transformer condition, *p* value, and age

Unit[a]	(A) O_2	(B) N_2	(C) CO_2	(D) CO	(E) H_2	(F) CH_4	(G) C_2H_6	(H) C_2H_4	(I) C_2H_2	(J) TCG	(P) TFC	(Q)[b] 0/1	(X)[c] 0/1	(Y) *p*	(Z) age
T1	3,566	12,602	531	39.58	11.7	20.7	12.61	2.52	0	87	0	0	0	0.018	1
T1	5,287	18,154	29	2.6	3.2	0	0	0	0.5	6	0	1	0	0.016	2
T1	2,812	18,505	2,942	122.6	107	78.8	57.9	25	0	392	0	1	0	0.125	3
T1	5,332	35,998	771	152.6	34.5	23.4	14.4	18.1	0	243	0	0	0	0.028	4
T1	3,959	20,874	1,159	41.4	2.2	46.2	93.4	32.2	0	215	0	0	0	0.039	5
T1	7,285	37,932	750	19.2	1.2	35.2	52.6	28.9	0	137	0	0	0	0.023	6
T1	5,871	32,426	430	32.1	48.2	53.5	36	23.6	0	193	0	0	0	0.039	7
T1	3,318	41,782	1,189	45.4	4.4	94.4	80.3	32.2	0	257	0	0	0	0.057	8
T1	3,088	21,712	324	26.2	1.7	71.9	53.9	8.2	0	162	0	0	0	0.039	9
T1	5,211	26,990	2,217	50.1	5.5	97.1	143.2	16.2	0	312	0	0	0	0.116	10
T2	4,595	30,772	1,833	119.6	54.1	950	256.2	1114	0	2494	0	1	1	0.763	–
T3	5,767	39,018	2,004	165.2	63	1356	290.7	1241	0	3115	0	1	1	0.993	–
T4	6,101	37,611	2,653	574.1	21	379.6	143.9	343.4	5.6	1468	19.8	1	1	0.983	–
T5	3,876	45,784	1,866	103	34.7	150.5	264.9	40.7	0.08	594	0	1	1	0.505	–
T6	2,949	23,935	1,111	74	59.5	76.6	32.6	11.5	0	254	31.5	0	0	0.029	–
T7	4,488	17,338	682	71.2	25.4	14.1	16.7	12.4	0	140	7.8	0	0	0.016	–
T8	2,715	21,962	562	33.1	11	49.7	21.6	6	0	121	7.5	0	0	0.024	–
T9	3,717	14,407	399	88.16	17.87	9.5	1.73	1.3	0	119	0	0	0	0.018	–
T10	2,080	7,937	599	54	36	15.0	13	5	0	123	0	0	0	0.023	–

[a] TN represents transformer unit *N* of the 161-kV series
[b] Transformer's condition including the normal (0) or abnormal (1)
[c] The result of predictions when the threshold is 0.5 ($p < 0.5$ represents normal, and $p \geqq 0.5$ is abnormal)

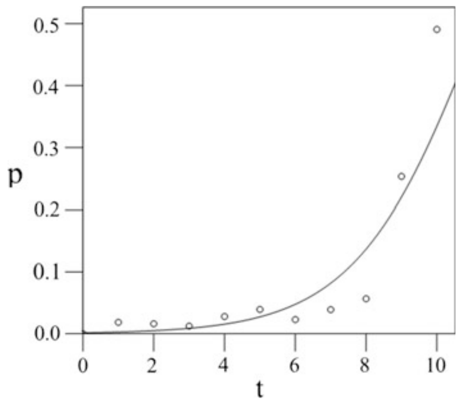


Fig. 51.2 The data and logistic model for the T1 transformer case

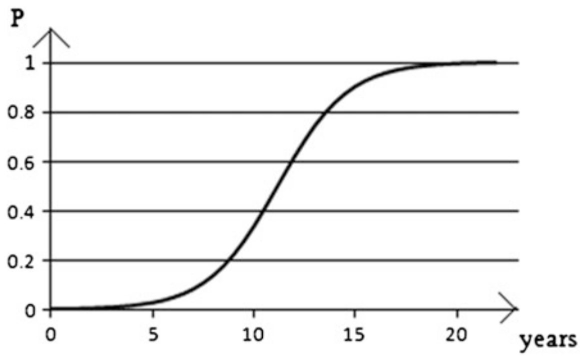


Fig. 51.3 The life span prediction for transformer T1

for time t and the probability p. In Fig. 51.2, the transformer years (age) corresponds to the failure probability, and it depicts the trend values of the Eq. 51.6. Similarly, Eq. 51.6 plots the S-curve as shown in Fig. 51.3 and predicts transformer T1's life span. To predict the transformer will have a breakdown when the failure probability reaches 0.99 ($p = 0.99$), Eq. 51.6 is used to show that transformer T1 will have a life expectancy of about 19 years ($t = 19$).

$$p = \frac{1}{1 + 639.28 \times 0.561^t} \tag{51.6}$$

The independent variables selected for the LR are used to determine whether the normal and abnormal states yield acceptable results. The LR failure probability and data life values fit the logistic model and predict the trend observed values.

51.5 Conclusion

This study was conducted for a series of 161-kV large electrical transformers. Using available information and LR to establish a suitable model determines the current status of the transformer. Other series of transformers can apply the same approach to further test the model. Furthermore, the logistic model estimates the remaining life of the transformer and helps engineers to make decisions. For near-death transformers, engineers pay intensive attention to the health of the transformer and can use early replacement before a major disaster. Given that the life of the transformer is very long, engineers can continue maintenance based on current maintenance plans unless an anomaly is indicated by the data. Therefore, as long as we have the relationship data of the transformers, we can help engineers make judgments about the health of the transformer, and then, the transformer can be salvaged and maintained. In this paper, the model for determining transformer health was established and the transformer life span was calculated.

Acknowledgments This research was partially supported by Taiwan National Science Council research grant. The authors would like to thank Fortune Electric Co., Ltd., for co-sponsoring the research and providing practical experience and valuable feedback.

References

1. Chaidee E, Tippachon W (2011) Failure statistics and condition evaluation for power transformer maintenance. In: Asia-Pacific power and energy engineering conference (APPEEC), Wuhan, China, pp 1–4, 25–28 March 2011
2. Chen Q, Egan DM (2006) A Bayesian method for transformer life estimation using Perks' Hazard function. IEEE Trans Power Syst 21(4):1954–1965
3. Zheng YC (1993) The temperature inhibition and life diagnosis of oil-immersed transformer. Fortune Electr Mag 2:34–40 (in Chinese)
4. Institute of Electrical and Electronics Engineers (2009) C57.104-2008 IEEE Guide for the interpretation of gases generated in oil-Immersed transformers, pp 1–27 (Feb 2)
5. International Electrotechnical Commission (2007) IEC 60599 Ed. 2.1 b: 2007 mineral oil-impregnated electrical equipment in service—guide to the interpretation of dissolved and free gases analysis
6. Stebbins RD, Myers DS, Shkolnik AB (2003) Furanic compounds in dielectric liquid samples: review and update of diagnostic interpretation and estimation of insulation ageing. In: Proceedings of the 7th international conference on properties and applications of dielectric materials, Nagoya, pp 921–926, 1–5 June
7. Pradhan MK, Ramu TS (2005) On the estimation of elapsed life of oil-immersed power transformers. IEEE Trans Power Delivery 20(3):1962–1969
8. Ariffin MF, Ghosh PS (2007) Estimating the age of paper insulation in 33/11 kV distribution power transformers using mathematical modeling. In: 19th international conference on electricity distribution, Vienna, pp 1–4, 21–24 May 2007
9. Wouters PAAF, Schijndel AV, Wetzer JM (2011) Remaining lifetime modeling of power transformers: individual assets and fleets. IEEE Electr Insul Mag 27(1):45–51

10. Trappey CV, Wu HY, Taghaboni-Dutta F, Trappey AJC (2011) Using patent data for technology forecasting: China RFID patent analysis. Adv Eng Inform 25:53–64
11. Yan J, Lee J (2005) Degradation assessment and fault modes classification using logistic regression. J Manuf Sci Eng 127:912–914
12. Berkson J (1944) Application of the logistic function to bio-assay. J Am Stat Assoc 39(227):357–365

Chapter 52
Corrosion Identification of Gas Pipe Risers in Buildings Using Advanced Ultrasonic Guided Waves

Peter W. Tse and Xiaojuan Wang

Abstract In modern cities, natural gas is a major source of fuel that is needed in each resident's daily life. Gas pipe risers are used to distribute natural gas to each household unit. They penetrate into each unit through concrete walls to connect the main outdoor gas pipe to the indoor residential gas pipes. After long periods of operation, part of the pipe riser embedded in concrete is prone to corrosion due to the combined effect of corrosive materials in concrete and humid environment. If the deterioration caused by the riser corrosion is ignored, pipe rupture and poisonous gas leaks are possible. In the worst-case scenario, catastrophic gas explosions can lead to human fatalities. Therefore, there is an urgent need for effective and reliable non-destructing testing methods that could enable the identification of corrosion in in-service through-wall riser so that the safety and proper maintenance of pipeline system in building can be guaranteed. This paper presents the application of advanced ultrasonic guided waves for this purpose. The problems and difficulties involved in the inspection including construction of transduction system and signal analysis of complex reflection wave signals will be identified and discussed. The experiments will be conducted to demonstrate the effectiveness of developed techniques. With the help of all of the achievements, the building is expected to become safer to live and hundreds of million dollars per year will be saved by avoiding potential gas explosions.

52.1 Introduction

In modern cities, natural gas is a major source of fuel for heating and cooking that is needed in residents' daily life. Gas pipe risers are used to distribute natural gas to each household unit, as illustrated in Fig. 52.1. The pipe risers are erected along

P. W. Tse (✉) · X. Wang
Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong
e-mail: meptse@cityu.edu.hk

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_52,

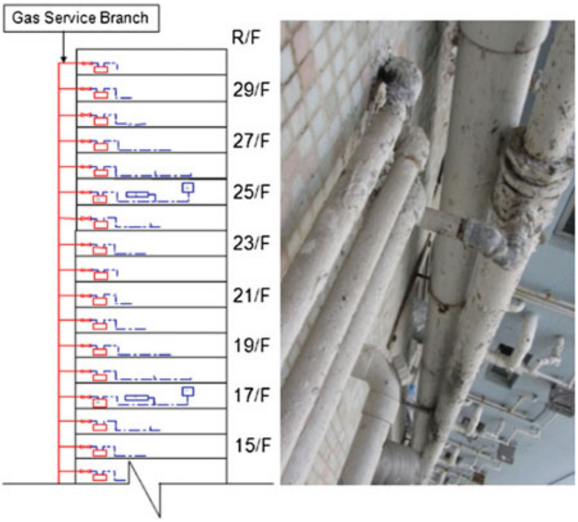


Fig. 52.1 Illustration of through-wall gas pipe riser

the external walls of the building and penetrate into each unit through concrete walls so that they can connect the main outdoor gas pipe to the indoor residential gas pipes. They constitute the most crucial infrastructure of any building. According to recent statistics provided by the Hong Kong and China Gas Co. Ltd. (Towngas), over 12.9 million household units in Mainland China and 1.75 million in Hong Kong have such risers. After long periods of operation, the part of the in situ riser that is embedded in concrete wall is prone to corrosion, as shown in Fig. 52.2, due to the combined effect of corrosive materials in concrete and ambient humid environment. If the deterioration caused by the riser corrosion is ignored, pipe rupture and poisonous gas leaks are possible. In the worst-case scenario, catastrophic gas explosions can lead to human fatalities. In 2007, a fatal explosion took place in a Hong Kong building, causing 2 deaths and 7 injuries. On December 30, 2008, a gas explosion occurred in Shanghai due to a gas leak from a through-wall riser. Two people were killed and 8 injured. It is obvious that the corrosions in pipe risers in a building are always intolerable and unaffordable for the integration of civil systems given increasing concerns for human safety, cost savings, and environmental protection. Therefore, it is in prime concern to provide advanced warnings and effective inspection prior to fatal failures of in situ pipe risers in building.

52.2 Inspection of Corrosion in Riser Using Guided Waves

A number of techniques have been considered by public utilities for inspecting pipe structures to meet the continuous requirements for building safety issues. They mainly include the methods based on acoustic, ultrasonic, and eddy current.

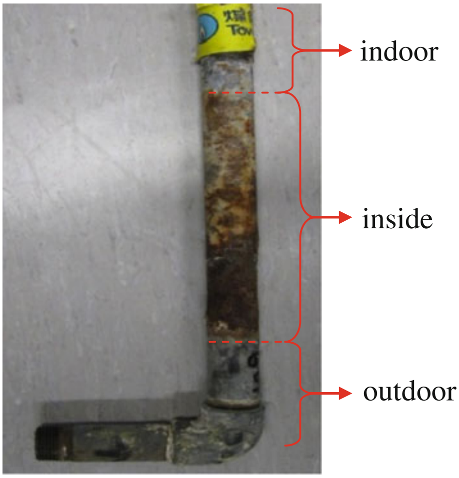


Fig. 52.2 Illustration of corrosion in pipe riser

Unfortunately, although these methods have achieved some success, they still have many limitations. Most of these methods are based on a passive mechanism, that is, only when an existing defect reaches a certain stage of deterioration or a break/leakage occurs do they become useful. Moreover, in these methods, inspection can only be enabled through scanning along the inspected structure in a point-by-point manner, so a comprehensive coverage of the inspected objects, especially the through-wall pipe risers that are impossible to access, is always difficult to achieve. It is thus obvious that these inspection procedures tend to be time-consuming. Therefore, conventional inspection methods for pipe risers in buildings are generally inefficient and prohibitively expensive for practical applications. There is currently no effective means of detecting corrosion in through-wall gas pipe risers.

Recent years have witnessed the evolution of the active inspection method that is based on ultrasonic guided waves [1], and it shows great potential for non-destructive testing (NDT) in inspecting various structures including pipeline, stand, and cable [2, 3]. Significant research has been conducted on topics related to structural inspection to meet the continuous requirements of civil and industrial users. This technique is potentially attractive for the inspection of the pipe riser used in buildings because of its line-to-line inspection, high speed, and high sensitivity. Moreover, it does not need to wait for the inspected structure to leak; rather, it actively excites an ultrasonic guided wave that can penetrate through the structures and find defects or impurities that exist in the structures. It is of particularly interest for applications where the areas of concern are not accessible. The guided waves technique is therefore expected to alleviate the aforementioned disadvantages of the conventional art of pipe inspection and enable accurate, reliable, and efficient NDT practices. However, few research has focused on inspecting short gas pipe risers embedded in concrete walls.

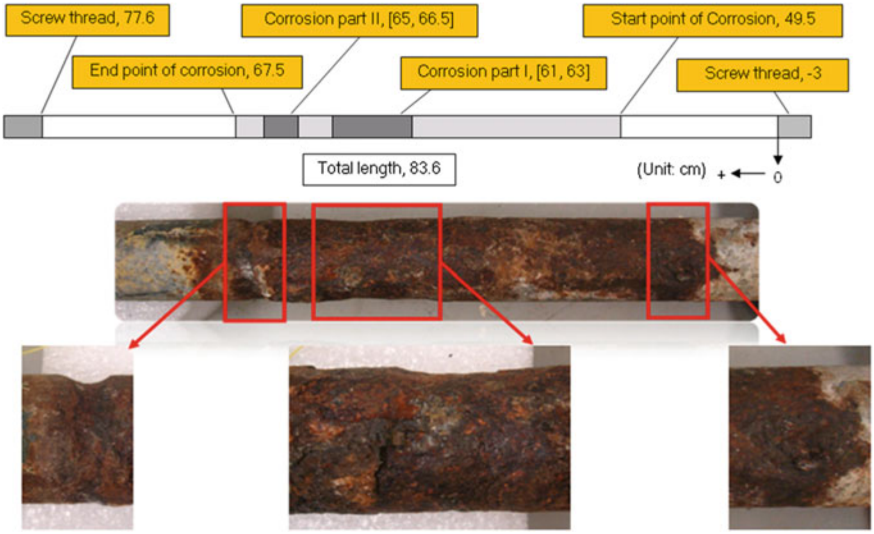


Fig. 52.3 Illustration of the tested corroded pipe riser sample and the corrosion in the pipe riser

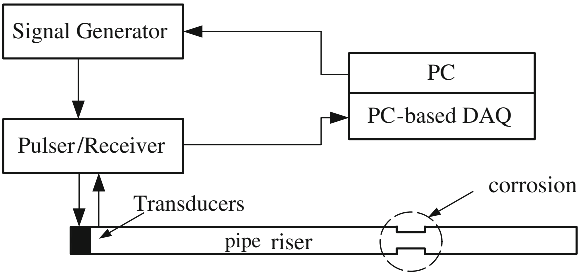
52.3 Analysis of Vibration Signals from the Piston Slap Test

To investigate the inspection and identification of corrosion in pipe riser, experiments were conducted on a real gas riser. The riser was provided by a natural gas supplier in Hong Kong, and it had a naturally developed corrosion with the class 5, which is a commonly occurring defect in the field. The profile of the corroded riser and the picture of corrosion are shown in Fig. 52.3. The riser sample had an external diameter of 34 mm, a wall thickness of 4.5 mm, and a length of 83.6 mm. There was a large corroded area in this riser, presenting a complex corrosion pattern. Three cracks could be clearly observed in the corroded area, as marked in Fig. 52.3. Two cracks were axisymmetric type of corrosion, and they had relatively large radial depth, one of which had large axial extent. The third crack was a pitch corrosion, the severity of which was not high compared to the other two corrosions.

52.3.1 Experimental Setup

The longitudinal L(0, 2) mode was adopted in this research because it is easy to excite and has relatively simple acoustic fields. The frequency of 135 kHz was chosen because it could provide enough resolution in inspecting short riser and the L(0, 2) mode excited at this frequency was non-disperse. The experimental setup

Fig. 52.4 Illustration of corrosion in pipe riser



and instruments used for the work in this research are depicted schematically in Fig. 52.4. A windowed tone burst consisting of five cycles at the chosen frequency was delivered through an arbitrary signal generator. The transduction system was carefully designed to excite a single L(0, 2) mode into the pipe riser under examination. In practical applications, unidirectional propagation of guided waves can be achieved by using multiple rings of transducers [4]. Hence, in this research, in order to simplify the concerned problem and put the focus on corrosion inspection, one end of the pipe riser with screw thread was cut so that only the guided waves at one direction need to be considered. A ring consisting of a certain number of piezoelectric transducers (PZTs) was bonded at this end to generate and receive guided waves. The PZTs were made of length expander-type piezoelectric material and distributed axisymmetrically, thus ensuring that only the expected longitudinal mode was excited, while the flexural modes were suppressed.

52.3.2 Detection of Corrosion in Pipe Riser

The ability of using high-frequency guided waves to detect the corrosion in pipe riser was firstly investigated. Besides the corroded pipe riser, another normal riser that had the same length, diameter, and thickness with the corroded one was used as the reference. The pictures of these two riser samples are shown in Fig. 52.5. To ensure that experimental data would have a reliable degree of comparability, all measurements were taken with strictly consistent experimental settings.

Figure 52.6 shows the collection of signals received from the two pipe risers, in which dark line corresponds to the wave signal from corroded pipe riser and red line corresponds to the wave signal from normal pipe riser, respectively. It is observed that an echo beginning at 3 ms appears in both signals. Based on the group velocity (5.33 m/ms at 150 kHz) and time information, the echo is proved to be the reflection from the riser ending. The energy of waves attenuates in propagation and such attenuation increases with the deterioration of structural condition of riser, so the amplitude of reflection from the ending of normal riser is much greater than that from the ending of corroded riser. In the wave signal from corroded pipe riser, besides the reflection from ending, another big echo can also

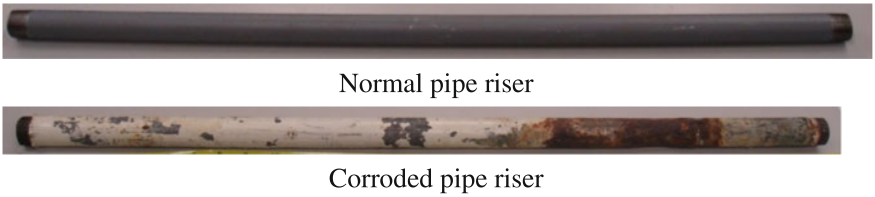


Fig. 52.5 Photographs of the normal and the corroded pipe risers for testing

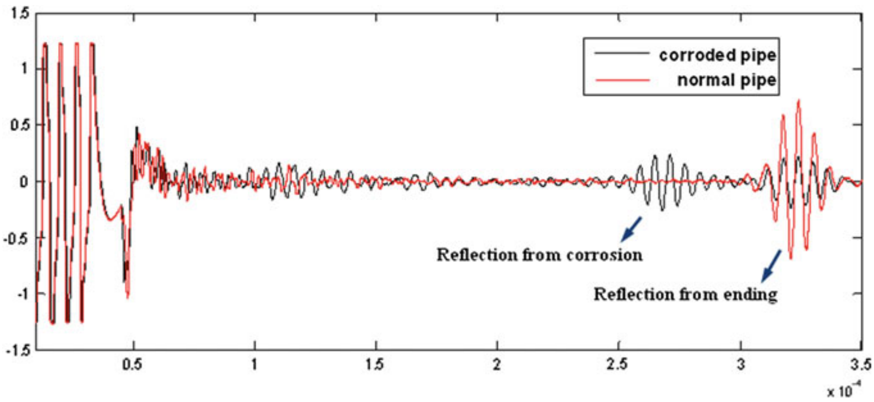


Fig. 52.6 The signals received from the normal pipe riser and the corroded pipe riser

be clearly observed. The main difference between the two signals is this echo. It is thus obvious that this echo is caused by the reflection from the corrosion in the corroded riser. The position calculated based on the group velocity and the time information of this echo is also roughly consistent with the actual position of corrosion. This result indicates that the guided waves can efficiently detect the existence of corrosion in pipe riser.

52.3.3 Identification of Corrosion in Rise

The identification of corrosion is always a challenging task after it has been found because of the complexities presented in the reflection wave signal [5]. The utility of guided waves is thus limited by the lack of the detailed information of detected corrosion. To carry out the planned maintenance and replacement operations on defective pipes accurately and efficiently, the ability of corrosion identification is very important in the practical application of guided waves technique, particularly for the cases in which the corrosion exists in the parts of riser where their accesses are difficult. The scale of difficulty magnifies in real practices as the reflected waves are subjected to the contamination from multiple reflections from multiple components of corrosion itself and a diversity of noise sources. This issue remains as an unsolved research task and is seldom reported in the related publications.

For the aim of identifying complex corrosion, it is important to understand reflection procedure firstly. It has been found by some researchers [6] that the reflection of guided waves from a defect is primarily resulted from the interference between the two reflection components generated at the edges of the defect. We further revealed in our earlier research work that the complexity of the reflection signal was essentially a result of the different features represented by the front-edge and back-edge reflection components. Two-edge reflection signals exhibit similar wave patterns in terms of the number of cycles, frequency, and modulation but have different amplitude and duration of phase shift. Therefore, the decomposition of reflection signal is a very important step for defect identification and subsequent comprehensive defect characterization. For the complex corrosion in pipe riser, each corrosion component that exists in the corroded area of a riser also complies with such reflection feature. Since those corrosion components locate closely, the reflections generated by them are superposed seriously in the received signal, resulting in the difficulty in identifying corrosion information. We newly developed a new method [7] which was based on matching pursuit (MP) decomposition technique with an optimized dictionary to resolve the primary reflection components from defect reflection signal so that the identification of a single pipe defect can be realized. This method can also be used herein to identify the multiple corrosions in corroded area of a pipe riser. The description of this method is presented below.

MP decomposition is a highly adaptive time–frequency signal processing technique introduced by Mallat and Zhang [8]. It could decompose any given signal into a linear expansion of components that belong to a redundant dictionary of waveforms. One of the key issues of MP is the design of the dictionary which is a collection of parameterized time–frequency atom. The atom is the dilation, translation, and modulation of an elementary function and presented by amplitude, phase, frequency, or other parameters. The components are selected from the dictionary to approximate the concerned signal structures so that the interested information embedded in the signal can be obtained. Namely, MP can provide a representation or interpretation of signal structure through decomposition process. In our research, considering the interference structure of reflection signal, the dictionary with the atom composing of two interfering reflection components is designed, which is termed as IRC dictionary. Two-edge reflection waveforms appear in the defined atom g with an interference structure, as defined in Eq. 52.1.

$$g(t) = \sum_{i=1}^{2} a_i S_i(t + \theta_i),\ \|g\|^2 = \int |g(t)|^2 \mathrm{d}t = 1, \tag{52.1}$$

where α and θ are the amplitude and phase index, respectively. The waveform of each signal $S(t)$ reflected from any edge of the defect can be described as the time-delayed and amplitude-scaled replica of tone burst signal in excitation. The IRC-based dictionary makes good use of the sufficient prior knowledge about concerned primary reflection components, so that important inner structure of reflection signal can be obtained. In the MP with IRC dictionary, assuming $\{y_1, y_2 \ldots y_l\}$ is the noisy observation of a target function f at points $\{x_1, x_2, \ldots x_l\}$. Given a complete dictionary $D = \{d_1, d_2, \ldots d_m\}$, which is a redundant collection of unit vectors in a Hilbert space H, the function f can be decomposed into a linear expansion of N atoms selected from D and a residual term $R_N(t)$, as follows:

$$f(t) = \sum_{n=1}^{N} a_n g_n + R_N(t),\ \|g_n\| = 1, \tag{52.2}$$

$$f_N = \sum_{n=1}^{N} a_n g_n, \tag{52.3}$$

where $\{g_1, g_2, \ldots, g_N\} \subset D$ is the basis of this expansion and $\{a_1, a_2, \ldots, a_N\} \subset \mathbf{R}^N$ is the set of corresponding coefficients of the expansion. f_N presents an approximation of f through selecting N proper basis atoms from the IRC dictionary. The algorithm is to select the basis $\{g\}$ and the corresponding coefficient $\{\alpha\}$ so that they can minimize the second-order norm of the residual component $R_N(t)$ for optimal approximation:

$$\|R_N\|^2 = \|y - f_N\|^2 = \sum_{i=}^{l} (y_i - f_N(x_i))^2 = \min, \tag{52.4}$$

In the first step of matching pursuit, the waveform atom that best matches the given signal f is chosen through evaluating the similarity based on inner product. Assuming the signal has been decomposed to $M - 1 \geq 0$ atoms, the M atom decomposition in the consecutive steps is further performed as follows:

1. Compute $|\langle R_{M-1}, g\rangle|$ for all $g \in D$;
2. Select an atom that best matches the residual R_{M-1} from the IRC dictionary;

$$|\langle R_{M-1}, g\rangle| \geq \rho \sup|\langle R_{M-1}, g\rangle|, \quad g \in D, \tag{52.5}$$

where $0 < \rho \leq 1$ is the number independent of M.

3. Compute the new residual as

$$R_M(t) := R_{M-1}(t) - \langle R_{M-1}, g_M\rangle g_M(t). \tag{52.6}$$

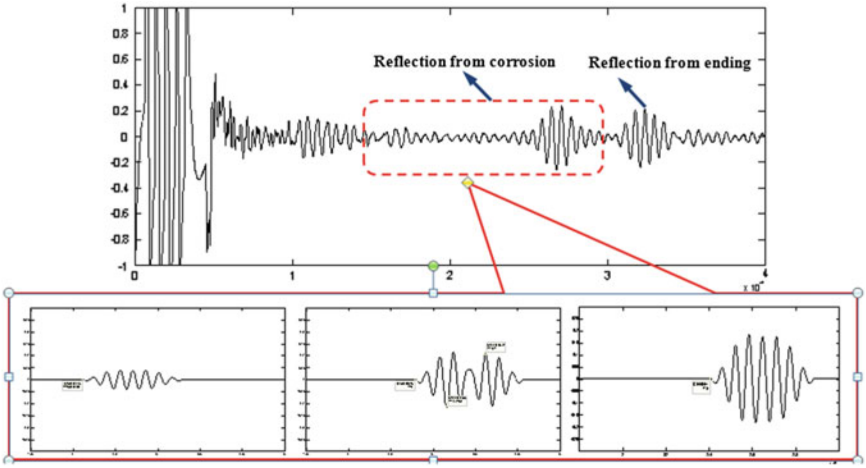


Fig. 52.7 The decomposition of multiple corrosions in the corroded area from the received reflection wave signal

The signal based on M atoms after M−1 decomposition interactions can then be represented as follows:

$$f(t) = \sum_{n=1}^{M} \langle R_n, g_n \rangle g_n(t) + R_M(t), \quad (52.7)$$

In the application concerned in this paper, the relationship between the edge reflection components of each corrosion in the corroded area and their respective features was adopted as prior information for MP decomposition procedure. By using the IRC dictionary to match the received reflection wave signal, the reflection components related to the corrosions in the corroded area can be sufficiently reconstructed. It could also provide a large advantage in high signal-to-noise ratio. The reconstructed result mainly includes the interested corrosion information that can be further used for comprehensive defect characterization. The results of implementing the proposed method to the reflection wave signal received from the pipe riser in this research are presented in Fig. 52.7. It shows that three reflection components are identified, which are proved to be generated by the actual three corrosions in the corroded area through calculating their physical locations based on the information of group velocity and time.

52.3.4 Conclusions

In this paper, the application of advanced ultrasonic guided waves for inspecting corrosion in pipe riser is presented. The problems and difficulties involved in inspecting pipe risers in building are identified and briefly discussed. Experiments on real pipe riser sample were conducted to demonstrate the effectiveness of the designed transduction system and the developed signal processing method. Although the methods introduced here still have rooms for further improvement, it shows good effectiveness and great potential in corrosion inspection and health assessment of pipe risers. With the continuous research efforts and improvement on these advanced techniques, buildings become more comfortable and safer to live in because of less frequent breakdown of services and reduced chance in the occurrence of catastrophe resulted by the sudden failure of ropes and rupture of pipe risers.

Acknowledgments The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 122011), and a grant from City University of Hong Kong (Project No. 7008187).

References

1. Rose JL, Ditric JJ, Pilarski A, Rajana K, Carr F (1994) A guided wave inspection technique for nuclear steam generator tubing. NDT&E Int 27:307–310
2. Wang X, Tse P, Mechefske C, Hua M (2010) Experimental investigation of reflection in guided wave-based inspection for the characterization of pipe defects. NDT&E Int 43:365–374
3. Tse P, Liu X, Liu Z, Wu B, He C, Wang X (2011) An innovative design of using flexible printed coil for magnetostrictive-based longitudinal guided wave sensor in steel strand inspection. Smart Mater Struct 20:055001 (12 pp)
4. http://www3.imperial.ac.uk/pls/portallive/docs/1/55745699.PDF
5. Wang X, Tse P, Djordjevich A (2011) Evaluation of pipe defect's characteristic axial length via model-based parameter estimation in ultrasonic guided-waves based inspection. Meas Sci Technol 22(2):1–13
6. Demma A, Cawley P, Lowe M, Roosenbrand A, Pavlakovic B (2004) The reflection of guided waves from notches in pipes: a guide for interpreting corrosion measurements. NDT&E Int 37:167–180
7. Tse P, Wang X (2012) The evaluation of pipe corrosion through the use of ultrasonic guided wave and novel matching pursuit. J Acoust Soc Am 131(4):3306
8. Mallat S, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Process 41:3397–3415

Chapter 53
A Multivariate Control Chart for Detecting a Possible Outbreak of Disease

Peter W. Tse, Jinfei Hu, A. K. Shrivastava and K. L. Tsui

Abstract Public health data often consist of spatial information, which are correlated between cities. Some researchers point out that monitoring variability is often more critical and sensitive for health surveillance when focusing on the mutually correlated data. In this paper, we present a multivariate control chart, which is based on likelihood ratio tests, to detect the variance–covariance shift caused by a disease outbreak by extending the popular change-point detection scheme for monitoring the mean of the outbreak process. Two patterns of the variance–covariance matrix change frame, namely the ‘step change pattern’ and the ‘persistent change pattern,’ are explored in this study. When monitoring some disease data, the squared exponential function is used to calculate the variance–covariance matrix that describes the information given by the observed disease and the cities. Then, the maximum likelihood method is used to estimate the parameters of the variance–covariance matrix function. The effectiveness of the chart is proven by using two examples. During the study, we observe the performance of the extended model for monitoring the variance–covariance matrix. The results show that the extended chart can detect the variance–covariance shift. With the success in predicting the time that an outbreak of a disease could occur and its possible shift, the management teams of health department and hospitals can prepare all necessary remedy and proper action for such outbreak in advance.

53.1 Introduction

As a main tool in statistical process control (SPC), control charts have been widely applied in monitoring the industrial manufactures to ensure that their processes are in control. A control chart is characterized by a control statistic that is updated

P. W. Tse (✉) · J. Hu · A. K. Shrivastava · K. L. Tsui
SEAM Lab, and CNDT Lab, Department of SEEM, City University of Hong Kong, Kowloon, Hong Kong
e-mail: mepwtse@cityu.edu.hk

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_53,

using current information of the samples at each time point. When the control statistic exceeds some predefined thresholds, the chart signals an alarm, which implies that a change in the underlying process has occurred. Till now, lots of researches have conducted a lot of work on the design and development of the control charts [1–5]. Most of these methods focused on univariate control chart for monitoring the process mean or process variance or both. However, many industrial applications seem much more sophisticated, which concern monitoring several correlated quality characteristics. In recent years, more and more SPC techniques have been applied in the public health surveillance to detect the disease outbreak. Although till now, many spatiotemporal surveillance methods are based on the assumption that the regional observations are independent. In fact, these public health data are often correlated between adjacent regions or time domain due to the contagious characteristic of the diseases. Therefore, the extension of the univariate control charts to the multivariate ones is necessary under such circumstances and it is more efficient to monitor the process by means of the multivariate control charts which capitalize on the correlated information between the variables [6].

Lots of recent works have been focused on developing the control charts for detecting shifts in the mean behavior of the processes [7, 8]. Although monitoring process variability has been always recognized to play second fiddle compared to monitoring the process mean, the increasing awareness of the importance of monitoring process variability has emerged with the requirement of the quality improvement as the variability becomes the major source of the quality problems. To the best of the authors' knowledge, little research effort has been devoted to the problem of monitoring process variance–covariance or the joint of process mean and variance–covariance. To mention a few, Yeh et al. [9] developed a multivariate control chart for monitoring process variability with individual observations under the assumption that the process mean remained unchanged [9]. They examined the variance and covariance components separately, and their simulation results showed that the performance of the proposed method was much better when variables were strongly positively correlated. Huwang et al. [10] proposed two control charts to monitor multivariate process variability, which were based on the traces of the estimated covariance matrices derived from the individual observations [10]. Hawkins and Edgard [11] discussed the major failing of charting the determinant or trace of variance [11]. They pointed out that both kinds of control chart were unable to detect covariance matrix changes in which some aspects of the variability increased, while others decreased. Moreover, Zhang et al. [12] proposed a multivariate charting scheme which integrated the generalized likelihood ratio test and the EWMA procedure [12]. Their single chart could simultaneously monitor the process mean vector and covariance matrix of a multivariate normal process. However, it was difficult to interpret the out-of-control signal due to the limit of their omnibus chart.

Jiang et al. [6] developed a set of multivariate change-point detection surveillance schemes for temporal health surveillance. Their multivariate control charts aimed to detect shift in process mean in the presence of spatial correlations. The simulation results showed that their schemes could provide faster detection of outbreaks of the disease than those methods based on the assumption that the regional observations were independent. However, their correlation structure is assumed to be constant during the whole monitoring time period. It is worth noting that the correlation structure is actually non-constant. In order to further improve the efficiency and accuracy of the detecting behavior, the information from the non-constant covariance matrix should not be ignored. Furthermore, as pointed out by many researchers, temporal surveillance in the presence of spatial correlation has not been thoroughly studied. Hence, these motivate us to further study the multivariate control chart and their application in health surveillance. In our research, we will extend the method of Jiang et al. to monitor the variance–covariance matrix for detecting the outbreaks of disease.

The remainder of the paper is organized as follows. Section 53.2 discusses the problem formulation that develops a likelihood ratio-based approach for temporal surveillance focusing on monitoring the variance–covariance matrix. Section 53.3 presents some simulation results to illustrate the effectiveness of our method. Section 53.4 uses two examples to evaluate our method. Finally, conclusions are provided in Sect. 53.5.

53.2 Problem Formulations

Let $Z_t = [z_{1t}, z_{2t}, \ldots, z_{Nt}]$ be a vector of observed responses at time t at region $s = 1, 2, \ldots, N$. We assume that Z_t follows a multivariate Gaussian distribution $Z_t \sim G(\mu_t, \Sigma_t)$, where the variance–covariance matrix $\Sigma_t = \mathrm{var}(Z_t)$ will be monitored to determine whether the disease outbreaks or not. When the process is in control, $\mu_t = \mu_0, \Sigma_t = \Sigma_0$. In our problem, the in-control parameters will be estimated from a sample of observations during the stable stage. When concerning the spatiotemporal surveillance, we will apply the squared exponential function to calculate the variance–covariance matrix Σ_t to better use the spatial information, that is, $\Sigma_t = \sigma_t^2 * \exp[-\beta_t * h^2]$, where σ_t^2 is the variance of an expected response and β is the correlation parameter vector, $\beta_t = [\beta_{1t}, \beta_{2t}]$ for two spatial dimension problems; h is the distance vector between different regions. In the following sections, we will illustrate two patterns of the variance–covariance matrix change frame: step change and persistent change separately.

53.2.1 Pattern I: Step Change

For simplicity, we usually assume the mean vector to be constant over time, although there is rarely a justified reason to expect shifts only in the mean or only in the variability in multivariate monitoring process [13]. While there is a change point at time $k \geq 0$, the parameters are changed to be $\mu_t = \mu_0, \Sigma_t = \Sigma_1$. The pattern of such process change mode is named as 'step change pattern.' To apply the change-point framework, that is,

$$Z_t \sim \begin{cases} G(\mu_t = \mu_0, \Sigma_t = \Sigma_0), & 0 < t < k, \\ G(\mu_t = \mu_0, \Sigma_t = \Sigma_1), & k \leq t \leq T. \end{cases} \tag{53.1}$$

where k which refers to the change-point time is unknown. T usually refers to the current time. The in-control variance–covariance matrix Σ_0 usually is estimated from some periods of time in which the process is steady. Once there is a change in health surveillance, i.e., an outbreak of disease, the variance–covariance matrix will be changed, denoted by Σ_1, and it will keep the same value for the following stage. The log-likelihood function for this pattern is given by

$$\begin{aligned} L_s(T, k, \Sigma_1 | Z_t) &= \log \frac{\prod_{t=k}^{T} (2\pi)^{N/2} |\Sigma_1|^{-1/2} \exp\left[-\frac{1}{2}(Z_t - \mu_0)/\Sigma_1^{-1}(Z_t - \mu_0)\right]}{\prod_{t=k}^{T} (2\pi)^{N/2} |\Sigma_0|^{-1/2} \exp\left[-\frac{1}{2}(Z_t - \mu_0)/\Sigma_0^{-1}(Z_t - \mu_0)\right]} \\ &= -\frac{1}{2}(T - k + 1) \log \frac{|\Sigma_1|}{|\Sigma_0|} + \sum_{t=k}^{T} \frac{1}{2}(Z_t - \mu_0)/\left(\Sigma_0^{-1} - \Sigma_1^{-1}\right)(Z_t - \mu_0). \end{aligned} \tag{53.2}$$

Then, the statistic is considered as follows:

$$T_s = \max_{1 \leq k \leq T} \max_{\Sigma_1 \in D} L_s(T, k, \Sigma_1 | Z_t). \tag{53.3}$$

The generalized likelihood ratio test T_s is obtained by maximizing $L_s(T, k, \Sigma_1 | Z_t)$ across all possible k values. The permissible range of k is $1 \leq k \leq T$. D is a space which contains all variance–covariance matrices Σ_1 during time $k \leq t \leq T$.

53.2.2 *Pattern II: Persistent Change*

The second pattern for the variance–covariance matrix change mode is named as 'persistent change pattern.' The shape of this pattern looks like waves. When there is a change point at time $k \geq 1$, the parameters are $\mu_t = \mu_0, \Sigma_t = \Sigma_t, k \leq t \leq T$. In this pattern, we also assume that the process mean keeps unchanged over time. The out-of-control variance–covariance matrix Σ_t is not a constant, and it is changing over time from the change point till some realistic strategies are carried out on those disease outbreak regions, such as the administration of vaccination supplies from government. The suitable example for this pattern is the influenza epidemics trend in USA during time period 2003–2011 which has been intensively studied by researches [14]. This change-point framework is given as follows:

$$Z_t \sim \begin{cases} G(\mu_t = \mu_0, \Sigma_t = \Sigma_0), & 0 < t < k, \\ G(\mu_t = \mu_0, \Sigma_t = \Sigma_t), & k \le t \le T. \end{cases} \tag{53.4}$$

The log-likelihood function is given by

$$\begin{aligned} L_p(T, k, \Sigma_t | Z_t) &= \log \frac{\prod_{t=k}^{T} (2\pi)^{N/2} |\Sigma_t|^{-1/2} \exp\left[-\frac{1}{2}(Z_t - \mu_0)/\Sigma_t^{-1}(Z_t - \mu_0)\right]}{\prod_{t=k}^{T} (2\pi)^{N/2} |\Sigma_0|^{-1/2} \exp\left[-\frac{1}{2}(Z_t - \mu_0)/\Sigma_0^{-1}(Z_t - \mu_0)\right]} \\ &= -\frac{1}{2}\sum_{t=k}^{T} \log \frac{|\Sigma_t|}{|\Sigma_0|} + \sum_{t=k}^{T} \frac{1}{2}(Z_t - \mu_0)/\left(\Sigma_0^{-1} - \Sigma_t^{-1}\right)(Z_t - \mu_0). \end{aligned} \tag{53.5}$$

Then, the statistic is considered as follows:

$$T_p = \max_{1 \le k \le T} L_p(T, k, \Sigma_t | Z_t). \tag{53.6}$$

53.3 Simulations

In this section, we will evaluate the ‘step change pattern’ model for monitoring the variance–covariance matrix by using simulation studies. The simulations are executed by plotting the statistics when shifting the parameters σ or β in the variance–covariance matrix function. Because it has the same impact on the variance–covariance matrix by shifting σ or β, we just shift σ for illustrating the performance evaluation. Before the simulation, the in-control mean μ_0 and covariance matrix Σ_0 are estimated from phase I study. The out-of-control Σ_1 is unknown a priori and estimated by using maximum likelihood method. In the simulations, if time T in Eqs. (53.3) or (53.6) is a large value, then searching the k to maximize the statistics remains a very large computational burden. An effective way to reduce the computations without discarding information is retaining only the most recent observations by using a window and restricting the search for k value only to these values, that is, drop the oldest entry and at the same time add the newly updated one. In the following simulations, the in-control state is set as $0 < t \le 10$; the out-of-control state is set as $11 \le t \le 30$. All the data are randomly sampled from ten fixed regions which are also randomly generated. To avoid the possible confusion with the following parameters, the predefined parameters are referred as ‘source σ’ and ‘source β,’ respectively. The in-control parameters are predefined as $[\beta_1, \beta_2] = (1.0, 1.0), \sigma^2 = 0.1$; the out-of-control ‘source σ^2’ is linearly increased with the rate of 0.4, and simultaneously, the out-of-control ‘source β’ keeps unchanged, which are predefined as follows: group (a) $[\beta_1, \beta_2] = (1.0, 1.0), \sigma^2 = 0.5$; group (b) $[\beta_1, \beta_2] = (1.0, 1.0), \sigma^2 = 0.9$; group

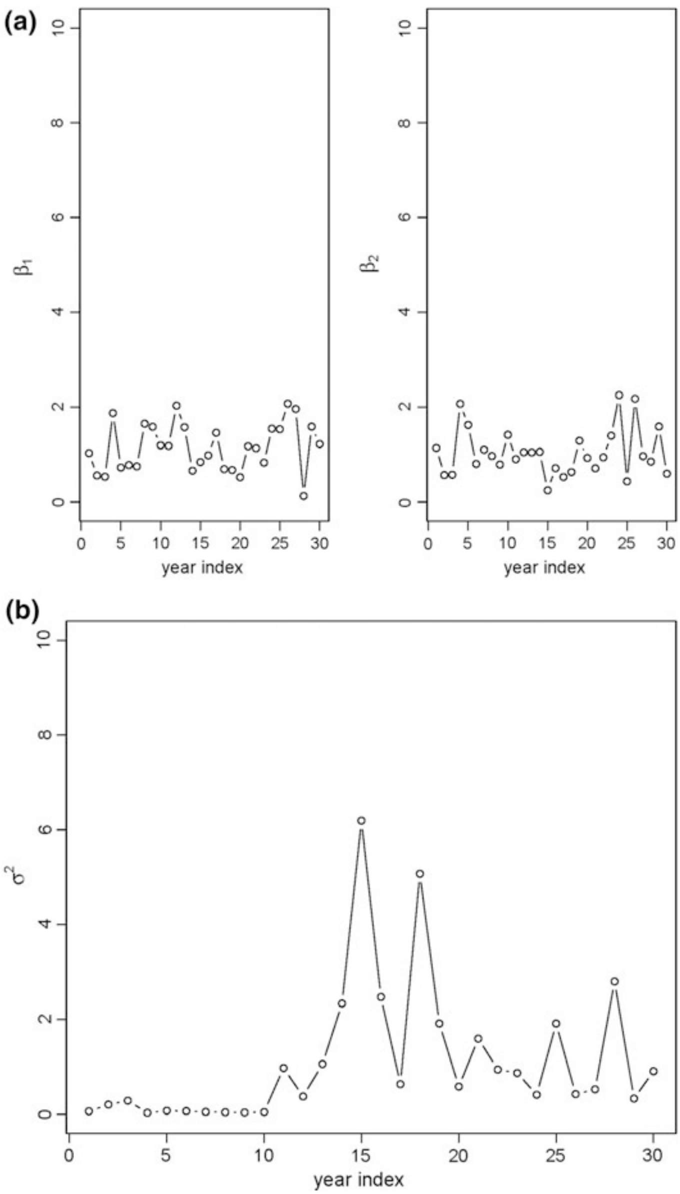


Fig. 53.1 Simulation example: **a** estimated β_1 and β_2 and **b** estimated σ^2

(c) $[\beta_1, \beta_2] = (1.0, 1.0), \sigma^2 = 1.3$; and group (d) $[\beta_1, \beta_2] = (1.0, 1.0), \sigma^2 = 1.7$. Based on such settings, the corresponding samples are generated randomly. Then, the function 'mlegp' in R software is used to estimate the out-of-control parameters σ and β to calculate the variance–covariance matrix function Σ_1 in Eq. 53.3.

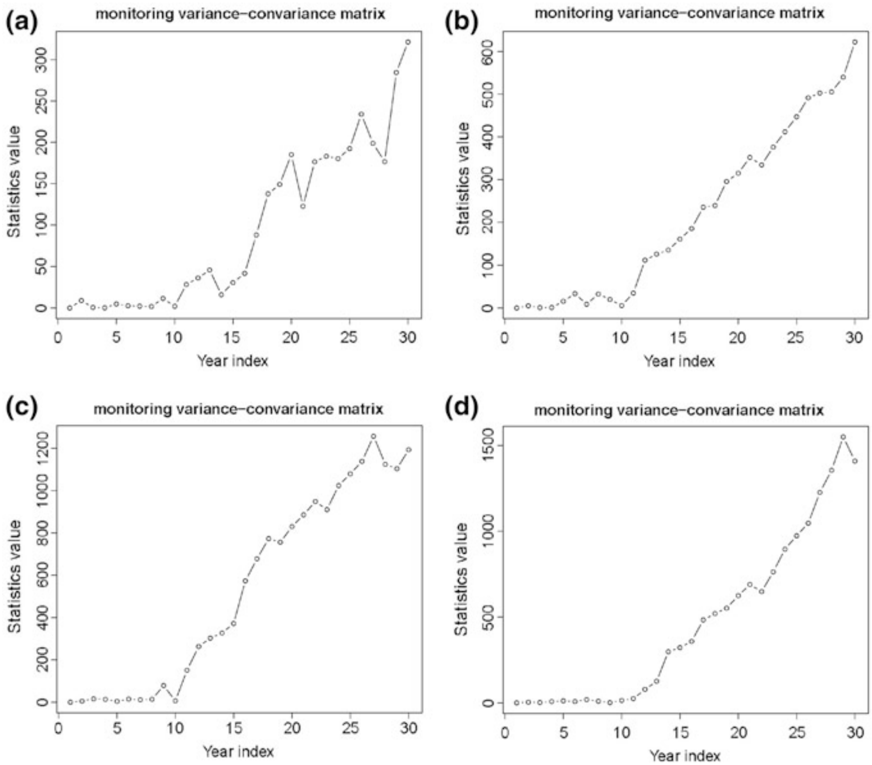


Fig. 53.2 Statistics for monitoring the variance–covariance matrix: in control: $[\beta_1, \beta_2] = (1.0, 1.0)$ **a** out of control: $[\beta_1, \beta_2] = (1.0, 1.0)$, $\sigma^2 = 0.5$; **b** out of control: $[\beta_1, \beta_2] = (1.0, 1.0)$, $\sigma^2 = 0.9$; **c** out of control: $[\beta_1, \beta_2] = (1.0, 1.0)$, $\sigma^2 = 1.3$; and **d** out of control: $[\beta_1, \beta_2] = (1.0, 1.0)$, $\sigma^2 = 1.7$

As the 'step change pattern,' the variance–covariance matrix would show a step change trend. The phenomenon appears from the parameters out of variance–covariance matrix function Σ_1. Taking group (d) as an example, the estimated parameter σ^2 presents an obvious step change since the 11th year and β_1 and β_2 keep approximately unchanged during all 30 years as shown in Fig. 53.1, which is consistent with group (d)'s definition. From Fig. 53.2a–d, it is quite obvious that the statistic is efficient to monitor the variance–covariance matrix which is shifted from the in-control state since the 11th-year index. Moreover, the larger the shifts of the parameter σ^2 from the in-control values are, that is, the larger shifts of the variance–covariance matrix from the in-control ones are, the larger statistic values it shows, which presents that the out-of-control lines are much steeper.

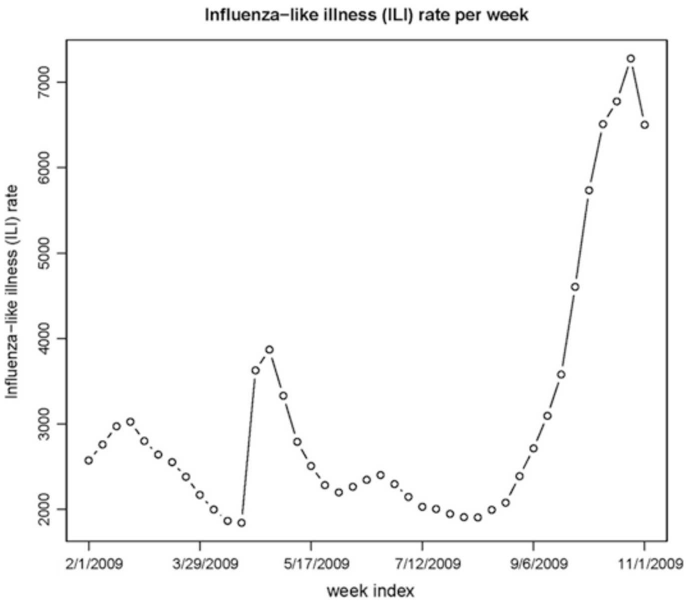


Fig. 53.3 Influenza-like illness rate in California

53.4 Example: Influenza-like Illness in California

In this section, the 'persistent change pattern' model is evaluated for monitoring the variance–covariance matrix by using a real example, which is based on influenza-like illness (ILI) data in California during February 1, 2009–February 21, 2010. The data set is obtained from Google Flu Trends (http://www.google.org/flutrends). The variance–covariance of the ILI rate for 12 cities in California was analyzed. The coordinates of the 12 cities have been mapped into a range of (0, 1). In this example, the original data by using function 'mshapiro.test' in R software were firstly tested to model the ILI rates using a multinormal distribution. The test results show that the cleaned data set approximately follows a multinormal distribution. Figure 53.3 presents the ILI's average counts across the 12 cities. It shows that there are several peak values distributed at some time, such as 'February 22, 2009,' 'April 26, 2009,' and 'October 11, 2009.'

The estimated parameters σ^2 and β_1 and β_2, as shown in Fig. 53.4, well present the persistent changes. Comparing the ILI's average counts in Fig. 53.3 with the estimated parameter σ^2 in Fig. 53.4b, it is shown that both of the ILI's average counts and estimated parameter σ^2 present the same persistent change route, especially appearing the same peak values on the same week index when the flu season is happening, which illustrates the effectiveness for detecting the outbreak of the influenza by monitoring the variance–covariance matrix. From Fig. 53.5, it seems quite obvious that there is an outbreak of the influenza since October 4,

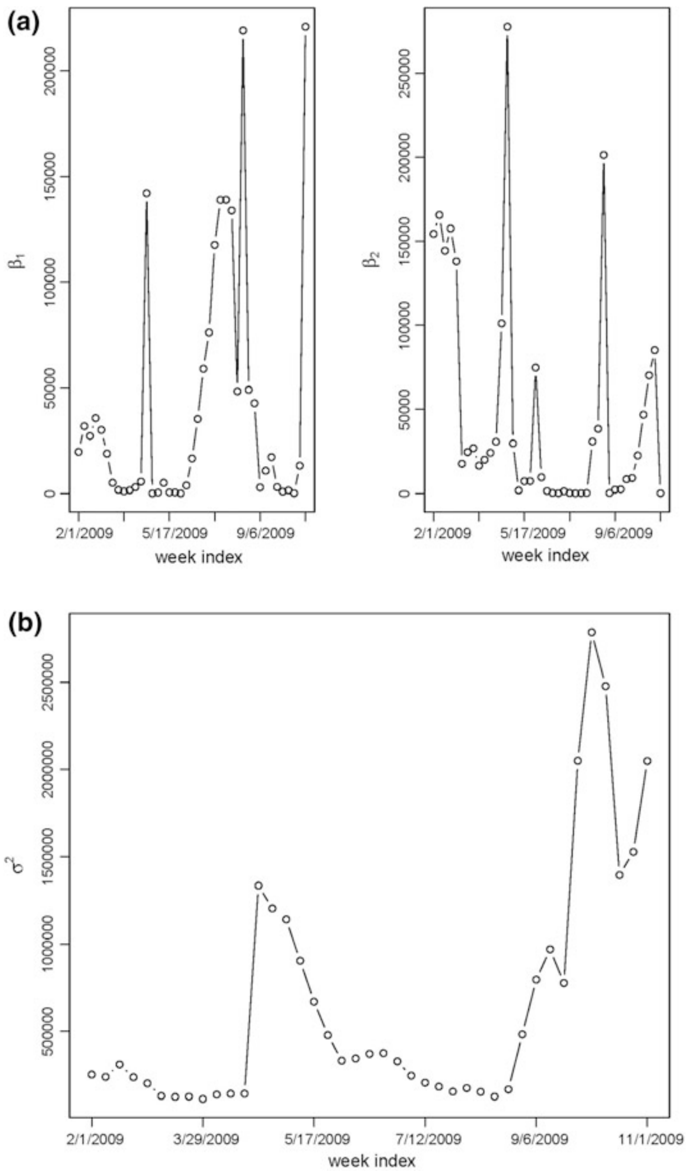


Fig. 53.4 Influenza-like illness in California: **a** estimated β_1 and β_2 and **b** estimated σ^2

2009 as the statistic starts to increase sharply. The result is consistent with the one from the Google Flu Trends (http://www.google.org/flutrends), where presents a big outbreak since October 4, 2009. Moreover, the result from the Google Flu

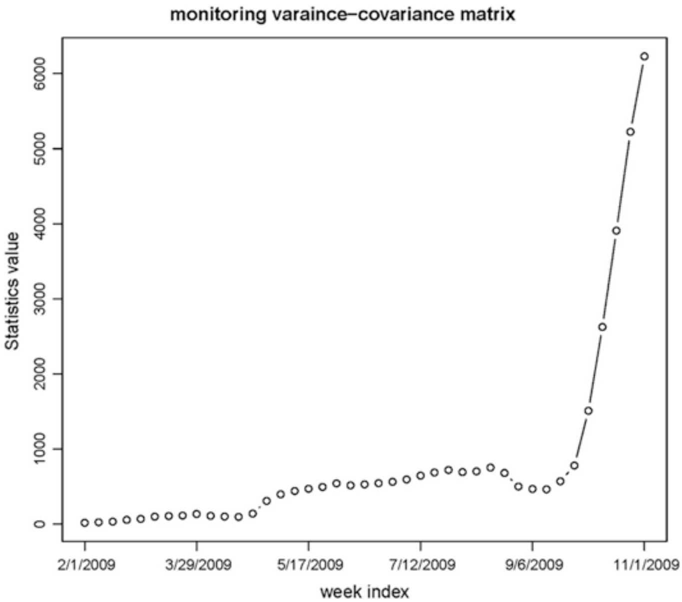


Fig. 53.5 Surveillance statistic for ILI rate in California

Trends is generated from more cities in California. In a word, it can be easily seen from the results that our proposed model could predict the outbreak of the influenza very timely, which is more important in practice.

53.5 Conclusions

In this paper, the multivariate control chart of W. Jiang et al. is extended to monitor the process variance–covariance matrix for detecting the disease. The model is based on the likelihood ratio tests. Two patterns of the variance–covariance matrix change frame were explored in this study, including the 'step change pattern' and the 'persistent change pattern.' The appearances of both the change patterns present in the estimated parameters σ^2 and β out of the variance–covariance matrix function. To estimate the parameters of the variance–covariance matrix function, the popular maximum likelihood method is used in our paper. The effectiveness of the multivariate control chart for monitoring the variance–covariance shift is illustrated by using both simulation studies and a real data example. The results show that the proposed method is effectiveness, and both change pattern models present good performance for detecting the variance–covariance shift. With the success in predicting the outbreak time of a disease and its possible shift, the management teams of health department and hospitals can prepare all necessary remedy and proper action for such outbreak in advance. However, there is still one problem unsolved, that is, the place where the outbreak occurs cannot be determined. Therefore, in the future work, we will further improve the model and investigate the diagnosis of the outbreak regions(s). Furthermore, in previous research, we assume that the mean vector keeps unchanged when monitoring the variance–covariance shift. However, in practical application, such assumption may need further discussion. Therefore, another issue would be analyzed to investigate the influence of the mean shift over the variability shift.

Acknowledgments The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 122011), and a grant from City University of Hong Kong (Project No. 7008187). This work was also supported by a grant from City University of Hong Kong (Project No. 7002601).

References

1. Siegmund D, Venkatraman ES (1995) Using the generalized likelihood ratio statistic for sequential detection of a change point. Ann Stat 23:255–271
2. Rogerson P, Yamada I (2004) Monitoring change in spatial patterns of disease: comparing univariate and multivariate cumulative sum approaches. Stat Med 23:2195–2214
3. Capizzi G, Masarotto G (2003) An adaptive exponentially weighted moving average control chart. Technometrics 45:199–207
4. Linderman K, Love E (2000) Economic and economic statistical designs for MEWMA control charts. J Qual Technol 32:410–417
5. Shu L, Jiang W (2008) A new EWMA chart for monitoring process dispersion. J Qual Technol 40:319–331
6. Jiang W, Han W, Tsui L, Woodall W (2011) Spatiotemporal surveillance methods in the presence of spatial correlation. Stat Med 30:569–583
7. Ngai HM, Zhang J (2001) Multivariate cumulative sum control charts based on projection pursuit. Stat Sinica 11:747–766
8. Hawkins M (1993) Regression adjustment for variables in multivariate quality control. J Qual Technol 25:170–182
9. Yeh A, Huwang L, Wu CW (2005) A multivariate EWMA control chart for monitoring process variability with individual observations. IIE Trans 37:1023–1035
10. Huwang LC, Yeh A, Wu CW (2007) Monitoring multivariate process variability for individual observations. J Qual Technol 39:258–278
11. Hawkins D, Edgard MT (2008) Multivariate exponentially weighted moving covariance matrix. Technometrics 50:155–166
12. Zhang JJ, Li ZH, Wang ZJ (2010) A multivariate control chart for simultaneously monitoring process mean and variability. Comput Stat Data Anal 54:2244–2252
13. Reynolds M, Cho GY (2006) Multivariate control charts for monitoring the mean vector and covariance matrix. J Qual Technol 38:230–253
14. Ginsberg J, Mohebbi M, Patel R, Brammer L, Smolinski M, Brilliant L (2009) Detecting influenza epidemics using search engine query data. Nature 457:1012–1015

Chapter 54
Value-Focused Thinking: An Approach to Structure Company Values for Asset Risk Management

T. E. van der Lei and A. Ligtvoet

Abstract For public and semipublic organizations, the use of risk assessment matrices is becoming a standard approach for asset management decisions. The risk matrix allows companies to assess the chance and effects of different risks with respect to proposed investment decisions. When the values in the risk matrix reflect the company values, the risk matrix allows for investment decisions to be made in line with these values and even for further strategic decision making. In this paper, we describe how a hierarchy of values can be used to uncover the core values of a company. These values are often shared by the people in a company and are institutionalized in core documents, such as the strategic vision or annual reports. We describe an approach in which we used value-focused thinking [1] for the systematic elicitation of company values to determine the aggregation level of the values represented in the asset risk matrix of the Port Authority of Rotterdam (PoR).

54.1 Introduction

Many public and semipublic organizations in the Netherlands are developing their risk-based asset management approaches. For risk-based asset management, risk assessment matrices are used as a decision support tool. The risk matrix allows companies to assess different investment decisions.

A risk assessment matrix states the chance that an event will occur and the effect of the event. The resulting risk is the amount of harm that is incurred for the given event. There are many types of risk assessment matrices that follow this line of reasoning. Differences lie in the scales that are used for the likelihood of the event, effect of the event, and severity of the risk. Most often, risk assessment

T. E. van der Lei (✉) · A. Ligtvoet
Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands
e-mail: t.e.vanderlei@tudelft.nl

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_54,

matrices use a single effect scale describing the effects. The risk assessment matrix that was developed in the project is a risk matrix with multiple values with corresponding effects scales. This setup of the risk matrix allows for decision making with respect to multiple values.

The values in the risk matrix are a reflection of the core values of the organization that uses the matrix. Many large companies use this type of risk matrices in their daily practice, e.g., Shell. Public and semipublic asset management organizations differ from other types of companies as their assets are most often located in the public domain and may also be used by the public. These companies are in need of tools to deal with large risks and uncertainties [e.g., 2, 3], and also, the values of these companies differ from private companies that tend to have a strong focus on shareholder value. In the Netherlands, however, public and semipublic organizations have been introducing this type of multi-value risk-based asset management approach. While there is sufficient literature on the use and misuse of risk assessment matrices [e.g., 4, 5], an approach for the elicitation of the values for the risk assessment matrix is not described [6]. In this paper, we describe how we used a three-step approach to elicit the company values of the Port Authority of Rotterdam (PoR). The PoR manages a wide range of different assets: roads, railroads, public spaces and greenery, buildings, quays, and the waterways (in particular the depth of these waterways). In our approach, we used value-focused thinking [e.g., 1] combined with document analysis and interviews to elicit the values for the risk assessment matrix of the company involved.

The PoR has a vision to become a global hub and Europe's industrial cluster and wants to mature in its asset management practices by better embedding the asset management philosophy into the organization. This means additional effort regarding the optimization of business processes across different levels of the company. A seamless integrated process of asset management from strategic to operational level can be seen as a high level of "maturity" of the organization [7, 8]. The PoR vision of world-class asset management for the company is aimed at this high level of maturity in the business. Risk management has been identified as a central pillar of asset management that is not sufficiently supported in the current business.

54.2 Analysis of Values

Value elicitation differs with respect to the purpose of the elicitation and the scientific field. Fischhoff [8] shows that different research paradigms have different assumptions with respect to values. He describes three different paradigms. The first paradigm assumes that people know what they want and is called the philosophy of articulated values. The second paradigm assumes that people have stable perspectives but that they are incoherent—the philosophy of partial perspectives. The third paradigm assumes that people lack articulated values on specific topics but have pertinent basic values that are stable. This final paradigm is called the philosophy of basic values. Value elicitation of stable values Fischhoff [8] argues differs from value elicitation in the two other paradigms. The risk assessment matrix assumes that the values are (relatively) stable and that choices can be made with respect to these values. This corresponds with the third paradigm.

The matrices that are in use in the energy sector in the Netherlands today are adaptations of the risk matrix approach of Shell that had the values of people, safety, environment, and reputation. The reason to adopt the matrix was that a tried and tested tool does not need to be changed substantially. Since the original implementation of the risk matrix, over the years, most companies have altered the values in their matrices. For example, for one organization, the value "sustainability" (in terms of an ecological footprint) was added, while for another the value "legality" was dropped arguing that all operations should comply with legal requirements. Overall, the values in the matrices of the sector have remained relatively stable.

Values arise from a process in which shared values are formed. These values are expressed, communicated, and recorded in documents. Reasoning from this perspective of shared value formation, business values are not thought to be decided, but they are the result of individual thought processes and communication of the employees (and thus can be elicited by researchers). Of course, some individuals, such as a CEO, can have more influence on the values of an organization [9].

When values are recognized by the organization, they are recorded and most often this is done by management. When values are seen as formal objectives, this is a further step in translating values into business practice. These types of values are periodically established in strategic documents, which in turn influence the thinking about values.

To elicit corporate values, official documents can be studied and its people interviewed. As for documents, a translation takes place into text, and the documents are merely an official representation of the values actually underwritten by the employees. These official documents are important, as they are formalized and institutionalized. To illustrate, most companies today have mission statements and corporate strategic documents that reflect their values [10]. Depending on the type of strategic document (e.g., annual reports, operational plans, or visionary statements), the values are expressed in different ways. Most often, the values represented in these documents are not centrally collected and systematically structured.

In order to gain an understanding of the corporate values, careful argumentative analysis is needed to uncover values from text. A text may read as follows:

> We continuously improve our company to become the most secure, efficient and sustainable in the world. For our customers we create value by developing value chains, networks and clusters, both in Europe and in emerging markets worldwide. Our company is an entrepreneur and developer of world-class partnerships for customers in the petrochemical, energy, and transportation and logistics sector. In this way we also strengthen the competitiveness of the Netherlands.

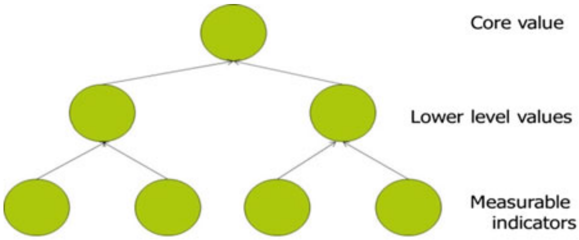


Fig. 54.1 Hierarchy of values

Keeny has provided a method that allows for the structuring of values. The method of Keeny falls into the paradigm of stable values as he describes values as principals used for evaluation [1, p. 6]. Values may be structured using a hierarchy of values or objectives hierarchy (see Fig. 54.1) [1, 11–14]. At the top of the hierarchy, the central value is provided, and this value is defined by the lower values in the hierarchy. The lowest-level values may be operationalized into key performance indicators, making the values measurable and/or quantifiable. The principle of value trees for a company is simple: The top is formed by the most important value in the mission statement of the organization. The layer below contains the values that define the mission. These values can be further specific by lower-level values as needed until measurable indicators can be assigned.

54.3 Approach

Setting up a risk matrix based on corporate values is not a standard procedure. Public and semipublic organizations in the Netherlands that we interviewed have approached the construction of the matrix very differently. Some choose to organize company-wide discussion sessions to set up the values in the matrix, and others frame the values in a limited circle—such as the executive team. A structured approach like the one presented in this paper was not used.

For this research project, we developed a three-staged approach in order to elicit the corporate values of the PoR (see Fig. 54.2).

Analysis of strategic documents: In this stage, the corporate strategic documents were analyzed for values, objectives, criteria, and means. This was done by the researchers. To supplement the document analysis, an automated word count of multiple strategic documents was done to investigate what type of nouns were used most often in the strategic documents and whether these nouns represent values (e.g., when in a strategic implementation plan the word "environment" is mentioned most, we assume that this is an important focal point).

Application of value-focused thinking: In this stage, the collected goals, objectives, and criteria were clustered and used for the construction of the value tree. The clustering and the construction of the value tree were done by the researchers. In addition, two strategic documents were translated separately into

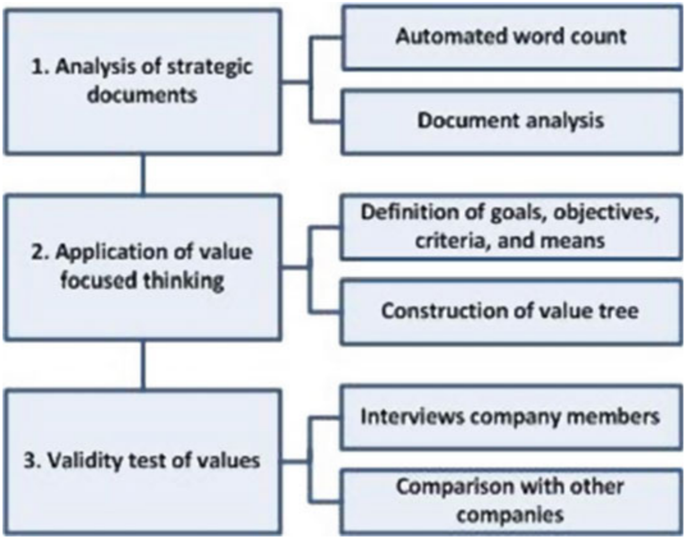


Fig. 54.2 Adopted value elicitation approach

value trees. Based on the analysis of the company documents, a synthesis tree was built that combined elements of both trees.

Validity test of values: Validity was measured as: the extent to which the uncovered values were shared by the company. In order to receive this feedback regarding the elicited values, the tree was used in eleven interviews with company employees. Nine heads of department were interviewed; a program manager and an advisor at corporate strategy level were asked to check the values and assumptions that were made on lower levels with respect to the constructed trees.

54.4 Results

For the PoR, three value trees were constructed. The first two value trees were based on two important strategic documents that were assumed to be known throughout the whole company. The first document was a medium-term document, a 5-year organizational plan, with key performance indicators but also broader goals such as internationalization and cooperation. The second document was a longer-term document, a 20-year Port Vision, a document with an integrated vision for the region that goes beyond the borders of control of the PoR as it also encompasses regional development. The translation of the values in these core documents into value trees allowed us to explain the concept of value trees, which was new to the organization. The third tree contained the values we collected from a larger sample of documents using the above-explained approach (see Fig. 54.3).

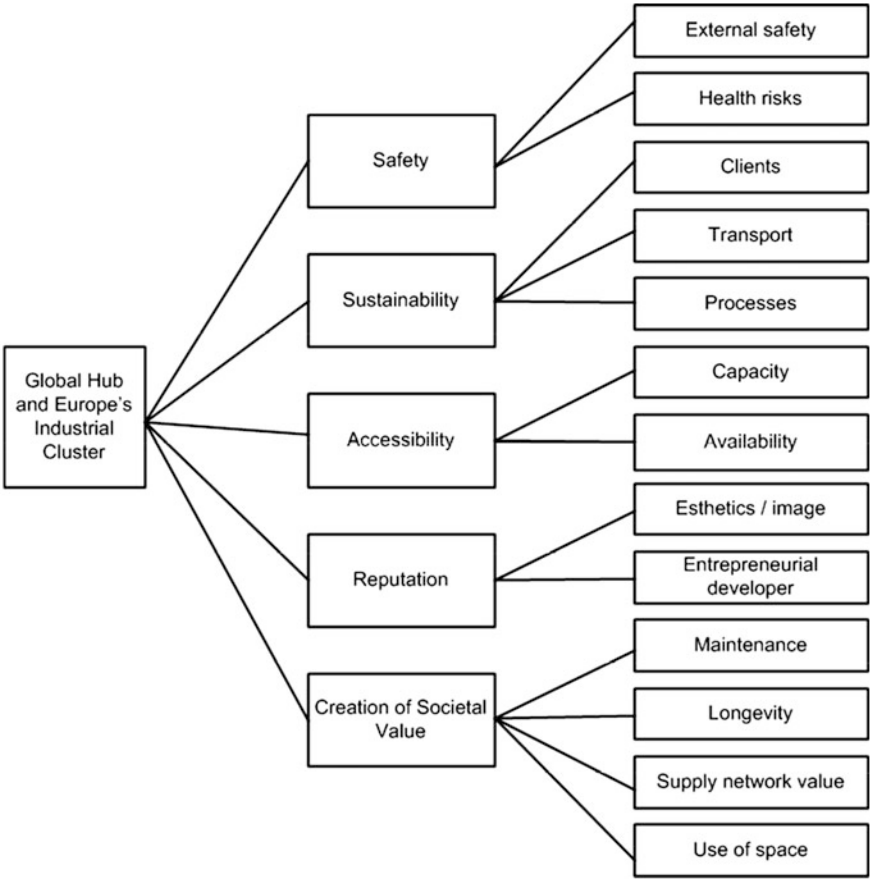


Fig. 54.3 Adopted value elicitation approach

This value tree was larger and contained more levels than the first and second tree that were based solely on the corporate documents.

As a result of the second step in the process, we arrived at five corporate values that were further subdivided into lower-level values. In order to be a global hub and Europe's industrial cluster, the following values were defined:

Safety: Safety was generally recognized as an important value. In all documents, various aspects such as environmental safety (emissions and industrial accidents), nautical safety, structural safety (such as collapsing structures), and road accidents were mentioned. Being a maritime organization, the documents mainly focused on nautical safety. However, we assumed (and were confirmed in our assumption) that *health risks* of people both inside and outside the company (number of injuries and/or deaths) were a general concern, not bound to specific modes of transportation as was *external safety* (which is safety to people and organizations outside the company).

Environmental Sustainability: This encompasses environmental impacts (emissions, discharges, etc.) and could be seen as another word for "clean." Sustainability is to be achieved both in the company's own *operations*, in related *transport* activities, as well as the activities of its *clients*. The strategic documents as well as some insights delivered by the risk department clearly indicated reducing emissions and decreasing the use of raw materials and energy as the focus of sustainability. However, sustainability appeared to be a difficult concept as it meant different things to different people: Some of the asset managers concerned with structures also called sustainability the longevity of their assets. Although assets of higher quality generally lead to less time and energy spent on maintenance and replacement, we eventually moved the longevity aspect to societal value creation.

Accessibility: This includes turnaround time of ships and accessibility via roads and rails. In short, this value is about the flow of goods. At first, the synthesis tree reads efficiency, as efficiency was recognized in the business. The debate was whether this was the best word: Efficiency is nothing more than a goal at an acceptable cost. Access is ultimately what was meant by efficiency: The four modalities shipping, rail, road, and pipelines need to have sufficient *capacity* for transport to and from clients, and these modalities need to have sufficient capacity and need to be *available* as much as possible.

Reputation: In this context means a combination of *aesthetics and a positive image* among (potential) customers, shareholders, and the press. More specifically, the organization wants to be known and active as an *entrepreneurial developer*. Reputation provided most debate as some of our interviewees argued that reputation is the result of a job done well. Furthermore, the term did not appear in the strategic documents we analyzed. However, the Board's reasoning is that reputation is a separate point of attention: It ensures the "license to operate" and especially the "license to grow."

Creating (societal) value: This includes not only generating sufficient funds for the organization itself for continued operation (such as return on investment, profitability) but also a broader goal to create value for clients, the surrounding region, and other stakeholders. Ultimately, this value is expressed in Euros. When we made the translation from the five values to the risk assessment matrix (not discussed in this paper), the less lofty term "finances" was chosen, as this was also widely recognized. The lower-level goals related to asset management were cost-effective *maintenance* and *longevity* of assets. Other concerns are facilitating the *supply networks* and the *use of space*.

In arriving at these five values, some decisions had to be made as to what should be considered a main value and what a supporting value: A need-to-have versus a nice-to-have. The medium-term strategic document, for example, mentioned the ambition to expand beyond the borders of the Netherlands. Although this was a prominent item in the document, we argued that international presence is not a core value of the company. In other words, the organization would still be itself without the international presence, whereas loosing safety or sustainability as a value would constitute a major change for the entire organization. Eventually, all interviewees agreed on the values. Most discussion took place regarding the lowest-level values.

54.5 Summary and Conclusion

In large organizations that are not strictly hierarchically controlled, similar activities take place at different departments. This may be because the organization is “ripe” for its next organizational changes. Following a meeting with the risk department, it turned out that the higher corporate levels made use of five values unknown to the asset management department, these were: safety, assets/finance, environment, reputation (public, stakeholders/customers), and core business. The similarity to our identified values was very large, with the only exception that what the risk department had called “core business,” we had labeled “accessibility.” The shared focus on risk was seen as an opportunity for further risk (based asset management) thinking throughout the whole organization.

Although this value discovery exercise was set up with the aim of constructing a risk matrix, knowing ones values (as an organization) has several purposes. The values provide a guideline for many business decisions: They allow for prioritizing projects, for investments decisions, and act as guidelines for annual reports. Risk assessment is becoming an integral part of total company practice.

In the Netherlands where the semipublic energy network companies are most advanced in their risk-based asset management practices, we see this trend of integration. Three of these companies we interviewed with respect to their risk-based asset management practices. All organizations visited used their corporate values in their risk assessment matrices. One of the companies is very advanced with respect to risk-based management as the entire company uses the same values for decision making—from the strategic level down to the operational level the same risk matrix is used. In the second interviewed company, the risk matrix was used solely in the department of asset management, but the tools used at the risk management department and the department of asset management communicate with each other (there are schemes in which the two systems are coupled together). Finally, at the third energy company, the part of the company responsible for asset management uses a risk matrix based on the company values, while the parent company uses a different approach. However, in this company, there is an active discussion on what level the risks should be established.

Such a scenario is also conceivable for the Port of Rotterdam: The risk matrix developed for the departments involved in becoming world class in asset management does not have to be rolled out company-wide immediately. A transition phase with multiple risk assessment tools in place is possible. However, as one of the objectives of asset management is the proper integration of asset life cycle management throughout the organization, it is advisable to aim for company-wide integration of the risk assessment tool used.

Finally, in practice, it is difficult to draw up a consistent value tree. The requirement of a good tree is that each value is defined by more than one underlying value and that it is comprehensive (no lacking items) and non-overlapping (no double items). The question whether a value is stated at the appropriate level in the hierarchy is a matter of logic, but its importance is also subject to the evaluation of the actors from the organization. As such, there is no "right" value tree.

Although value trees can be created using a bottom-up process, or by using document analysis with selected interviews as we did, the corporate values that are selected to for risk-based asset management should be supported by the board of directors. Then, only does the tree turn into the "right" tree. It should be noted that such a tree will require regular updates due to (1) new insights by the service providers and asset managers and (2) new strategic insights and directions at the asset owner level.

Acknowledgments This work was for the project "Asset Life Cycle Management for the Port Authority of Rotterdam" and was supported by the Next Generation Infrastructure Foundation.

References

1. Keeny RL (1992) Value focused thinking: a path to creative decision making. Harvard University Press, Cambridge
2. Marshall HE (1988) Techniques for treating uncertainty and risk in the economic evaluation of building investments. U.S. Department of Commerce
3. Phillips S, Martin J, Dainty A, Price A (2007) Uncertainty in best value decision making. J Financ Manage Property Constr 12:63–72
4. Cox LA Jr (2008) What's wrong with risk matrices? Risk Anal 28(2):497–512
5. Wijna Y (2010) Asset risk management: issues in the design of a risk matrix
6. Mills GR, Austin SA, Thomson DS, Devine-Wright H (2009) Applying a universal content and structure of values in construction management. J Bus Ethics 90:473–501
7. Van der Lei TE, Ligtvoet A, Volker L, Herder P (2011) Evaluating asset management maturity in the Netherlands: a compact benchmark of eight different asset management organizations. In: Proceedings of the world congress on engineering asset management 2011, Cincinnati
8. Volker L, Van der Lei TE, Ligtvoet A (2011) Developing a maturity model for infrastructural asset management systems. In: Conference on applied infrastructure research, Berlin, 7–8 Oct 2011
9. Fischhoff B (1991) Value elicitation: is there anything in there? Am Psychol 46(8):835–847
10. Collins JC, Porras JI (1996) Building your company's vision. Harvard Bus Rev 74:65 (96501)
11. David FR (1989) How companies define their mission. Long Range Plan 22:90–97
12. Keeney RL (1988) Building models of values. Eur J Oper Res 37:149–157
13. Keeney RL (1994) Using values in operations research. Oper Res 42:793–813
14. Kirkwood CW (1997) Strategic decision making: multiobjective decision analysis with spreadsheets. Duxbury Press, Pacific Grove

Chapter 55
Information Use in Dutch Sewer Asset Management

Wouter van Riel, Jeroen Langeveld, Paulien Herder and François Clemens

Abstract Sewer systems are capital-intensive infrastructures, requiring appropriate asset management to safeguard serviceability. In the Netherlands, effective sewer asset management is described by De Leeuw's control paradigm. Reliable data and information are key elements in decision making for rehabilitation and renewal of the assets. Rehabilitation is often based on limited asset condition information. Although various sources described procedures for guiding decision making for rehabilitation, it remains unclear which and how information sources are used in this process, and to what extent value trade-offs influence decisions. In order to improve current sewer asset management, this study assesses the availability and use of information in decision making for sewer system renewal in the Netherlands. Eighteen interviews were conducted at seven municipalities, combined with an analysis of their municipal sewerage plans. The interviewees described the decision making process and the information sources they use in this process. Decisions for sewer system renewal are often based on intuition in implicit risk analyses, where risk is defined as 'feeling times consequence'. Sewer asset management is ineffective, because it relies mostly on intuition, hampering justification, accountability, and repetition of decisions and preventing evaluation. Evaluation procedures and a critical attitude toward relevancy and quality of information are recommended.

W. van Riel (✉) · J. Langeveld · F. Clemens
Section Sanitary Engineering, Delft University of Technology, Stevinweg 1,
2628 CN Delft, The Netherlands
e-mail: w.a.p.vanriel@tudelft.nl

P. Herder
Section Energy and Industry, Delft University of Technology, Jafflaan 5,
2628 BX Delft, The Netherlands

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_55,

55.1 Introduction

Sewer systems are capital-intensive infrastructures, with a design service life between 50 and 90 years, requiring appropriate asset management to maintain system serviceability and minimize costs for rehabilitation. Asset management is described, for example by PAS 55 [1], which is set up aiming at maximizing profits of Anglo-Saxon businesses. Brown and Humphrey [2] describe asset management as balancing performance, costs, and risk.

In the Netherlands, sewer asset management has developed differently compared to other countries, because of two reasons. First, Dutch municipalities are responsible for collection and transportation of wastewater, including operation and maintenance of sewer systems. Municipalities are asset owner and asset manager. The treatment of wastewater is a responsibility of waterboards. The costs for municipal sewerage are fully covered by issuing taxes to households and companies. Because a municipality's goal is not to maximize profit, PAS 55 seems unsuitable to apply for managing a public infrastructure within a governmental body. Second, in the 1980s, the Netherlands was one of the first countries where the importance of sewer asset management increased considerably, because of an emphasis shift from expansion to maintenance of infrastructure serviceability [3]. This shift was initiated by the fact that in the 1980s, a sewer connectivity of 90 % was reached and further network expansion was no longer driving the sector. By the work of Oomens [4] in the early 1990s, Dutch sewer asset management adopted principles of De Leeuw's [5] control paradigm, by reformulating sewer asset management as a control problem. This study adopted De Leeuw's control paradigm, in order to evaluate the development of sewer asset management since the 1990s. Another approach to assess organizational structures is by Mintzberg's organizational configurations. For the scope of this study, applying the control paradigm is appropriate. Shown in Fig. 55.1, the paradigm states that a controller controls a system, both of which interact with their environment. Control is defined as any form of directed influence. For sewer asset management, the controller is the department responsible for sewerage and the controlled system is the sewer system.

The controller's ability to successfully control its system depends on five preconditions for effective control. De Leeuw describes effective as 'achieving the desired effect'. Figure 55.2 depicts that as efficacy. Dutch sewer asset management focuses on cost-efficiency, which is defined as achieving a desired outcome at lowest costs.

1. The controller has an objective and an evaluation mechanism to check whether the goals are met.
2. The controller has a model of the controlled system to predict the effect of potential control actions.
3. The controller has information about the environment and the controlled system.

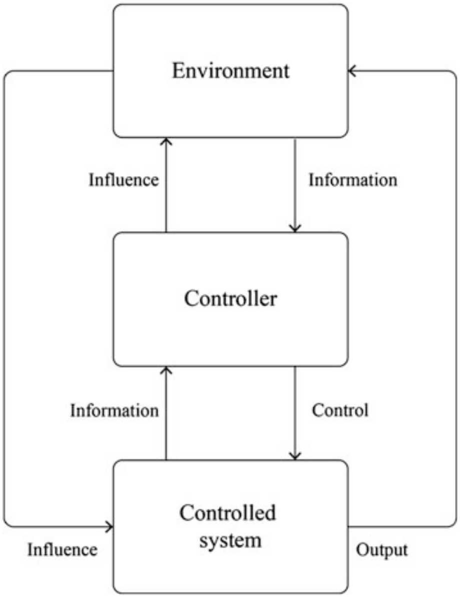


Fig. 55.1 De Leeuw's control paradigm [5]

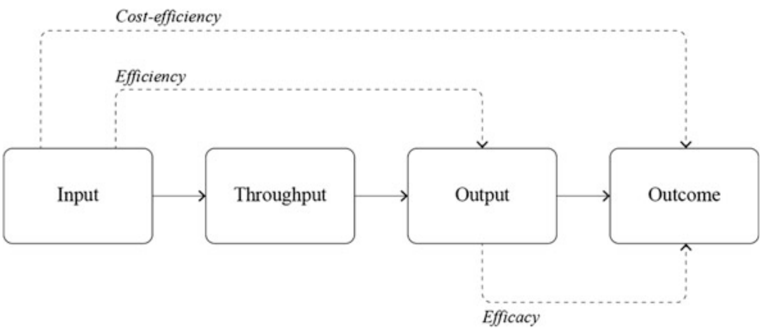


Fig. 55.2 Evaluating organizations [6]

4. The controller has sufficient control actions to cope with the variability of the system.
5. The controller has sufficient information processing capacity to transform incoming information into effective control actions that are in line with the objectives.

Current guidelines for sewer asset management are described in the European Standard EN 752 [7] and, for the Netherlands, in Urban Drainage Guidelines. Dutch national legislation obligates municipalities to set up a sewerage plan every 5 years. These plans describe policy objectives and costs for managing the sewer system. Adoption of the full paradigm for sewer asset management has not been completed, affecting the procedures in EN 752, Dutch Urban Drainage Guidelines and the municipal sewerage plans. For example, an evaluation mechanism for rehabilitation measures is lacking, preventing sewer system managers from improving asset management practices.

Reliable data and information are key elements in guiding the complicated decision making for rehabilitation of the physical sewer assets [8–10]. Asset managers are confronted with decreasing public acceptance, tighter available budgets, stricter legislation and increasing performance requirements [11]. More specifically for sewer asset management, the relation between sewer system works and other urban infrastructure works, such as road works, complicates planning in space and time [3, 12]. Decisions for rehabilitation of sewer systems are often based on limited asset condition information, such as pipe age and status, where camera inspections determine the status [12–16]. Anecdotal evidence suggests that municipalities in the Netherlands also use other types of information in the decision making process, including hydraulic performance, citizens' complaints, urban renewal plans, available budget or working capacity.

Although the EN 752 and municipal sewerage plan describe the process of information use for rehabilitation decisions, the practical application is different. It is unclear which and how information sources are used in deciding upon sewer system rehabilitation, and to what extent rehabilitation decisions are influenced by value trade-offs. This observation is not different from any other decision making process, but in order to improve current sewer asset management, this study assesses the availability and use of information in decision making for sewer system rehabilitation projects in the Netherlands.

55.2 Methods and Materials

The decision making process and current use of information were assessed by interviewing sewer system managers at Dutch municipalities and analyzing their sewerage plans. Each municipal sewerage plan contains a specific section about sewer system renewal, including motivation and projected efforts. Information sources were extracted from these sections.

A total of twenty-one interviews were conducted at eight Dutch municipalities, ranging in population size from less than 10,000 inhabitants to over 750,000 inhabitants. The interviewed municipalities constitute approximately 15 % of total population in the Netherlands. Table 55.1 shows several characteristics of the municipalities included in this study. Employees at three organizational levels were interviewed (strategic, tactical and operational), reflecting potential differences in information use, decision making and perspective toward asset management. Strategic management mainly addresses 'why' things are done. It is the management of aspects that are essential for the existence of an organization in its environment. This level is mainly concerned with defining and assessing middle-

Table 55.1 Characteristics of interviewed municipalities

Municipality	No. of inhabitants at 01-01-2011 [18] (–)	Population density [19] (inh./km^2 of land)	Sewer length[a] (km)	Available budget for 2012[a] (M Euro)	Available budget per inhabitant (Euro/inh.)	Available budget per km sewer pipe (K Euro/km)
Almere	190,655	1.469	1,100	8.7	45.6	7.9
Amsterdam	779,808	4.700	3,811	64.9	83.2	17.0
Barneveld	52,490	298	624	9.1	173.4	14.6
Breda	174,599	1.379	1,050	13.5	77.3	12.9
Ede	108,285	340	986	9.6	88.7	9.7
Rotterdam	610,386	2.987	2,906	51.2	83.9	17.6
The Hague	495,083	6.046	1,439	33.3	67.3	23.1

[a] Data are extracted from the municipal sewerage plan per municipality

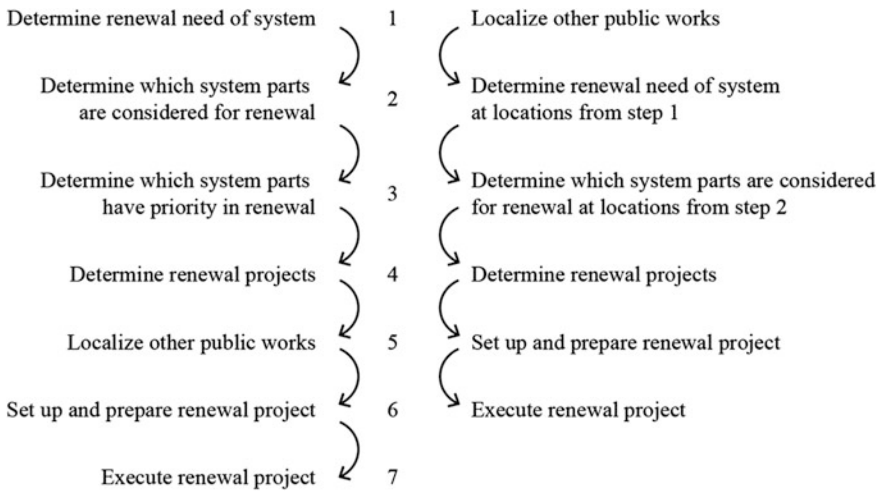


Fig. 55.3 Flow charts used during interviews

and long-term objectives, strongly based on influence from the environment the organization is in. Tactical management is organizing and structuring tasks, necessary from strategic management. Operational management is governing that tasks, necessary from strategic and tactical management, are carried out most appropriately [17].

A semi-structured interview schedule was used. During the interviews, two flowcharts were presented that visually assisted the interviewees during discussing the decision making process for sewer renewal projects. Figure 55.3 shows these flowcharts. Four exploratory interviews were conducted to prepare the flowcharts. The starting point of each flowchart differs, related to information flow in the control paradigm. One flowchart starts with system information (system condition) and the other flowchart starts with environment information (other public works).

Through the flowcharts, the following topics were addressed in chronological order: justification and completeness of flowcharts, information sources per step, budget allocation and identification of organizational levels. The interviewees described these topics based on their knowledge and experience.

55.3 Results and Discussion

The interviewees indicated that the decision making process for sewer system renewal consists projects of five steps, shown in Fig. 55.4.

Figure 55.4 shows that the starting point is the five-year budget allocation. According to the municipal sewerage plans, this budget allocation for sewer system renewal is based on pipe age and camera inspections. The yearly budget is allocated as approximately one-fifth of the five-year budget. After the budget is allocated, individual renewal projects are formulated, based on the renewal need and potential cooperation with other public works.

The interviewees mentioned twenty-two information sources they consider in deciding upon sewer system renewal. Table 55.2 shows twelve of the twenty-one identified information sources. These are the sources mentioned more than once. The information sources are categorized on the basis of the control paradigm, which distinguishes three categories: system, environment and organization information. This categorization is applied for this study.

Figure 55.4 shows discrepancy between strategic and operational decision making. The budget is allocated before and on other grounds than the operational activities. On a strategic level, decision making for renewal is based on pipe age and status. For the operational activities, sewer system managers try to cooperate with other public works as much as possible, usually with road works. This shift in controlled system complicates decision making for sewer system managers.

Deciding upon the need for sewer system renewal is mostly based on pipe condition assessment by pipe age and camera inspection images. Experience and intuition form the basis for interpreting pipe age and camera inspections and converting these to renewal actions. Camera inspection is the only information source for which decision making is normalized (European EN 13508-2 [20]), resulting in the intense use of it. Apart from the fact that camera inspections are unsuitable to judge about system performance, assessment of camera inspections images introduces significant uncertainty in the overall condition assessment [21]. Hydraulic modeling is used to check whether the hydraulic performance of the sewer system meets its objectives. Korving [22] showed that significant uncertainty for decision making is introduced by hydraulic modeling. For municipalities located on soft soil, settlement is also used for determining the need for renewal. Dirksen et al. [23] described the relation between settlement and sewer system performance. Yet, no reference model for settlement is available to assist in

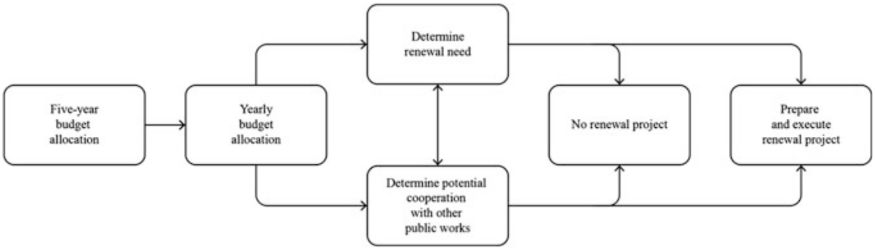


Fig. 55.4 Decision making process for sewer system renewal

Table 55.2 Identified information sources and identification frequency of eighteen interviews

Information category	Information source	Relative frequency (n = 18)
System	Camera inspection images	1.0
	Pipe age	0.9
	Hydraulic modeling	0.4
	Maintenance reports	0.4
	Function of sewer branch	0.2
Environment	Planning of public works	0.9
	Settlement rate	0.4
	Planning of urban development	0.4
	Citizens' complaints	0.3
	Traffic density of road above sewer	0.2
Organization	Experience and intuition	0.5
	Municipal water management strategy	0.2

decision making. In short, deciding upon the need for sewer system renewal is based on an implicit risk analysis of two aspects: pipe collapse and insufficient hydraulic performance.

Instead of a product of probability and consequence, the risk is determined by 'feeling times consequence'. A lack of insight into sewer deterioration processes causes sewer managers to avoid risks by renewing sewer pipes without knowing the remaining service life. Renewal decisions are ill-founded, because sewer system managers have insufficient insight into sewer deterioration processes and lack necessary information on sewer system condition and functioning [24]. In dealing with this issue, sewer renewal works integrate with other public works. Because of the risk aversive attitude, availability of sewer services is good compared to other utility services. The customer minutes lost in the Netherlands for sewer services is 0.2 min/customer/year, while for gas 0.4, drinking water 14.0, and electricity 28.9 min/customer/year [25]. The interviewees indicated they would like to be able to make predictions about the system's performance and structural condition to have more control over their system and planning of renewal works.

Integrating sewer works with other public works is also based on an implicit risk analysis considering three aspects: extra nuisance for citizens and related imago of the municipality due to extra excavation works, higher costs due to extra excavation works and road reconstruction, and extra traffic disruption due to extra excavation works.

The implicit risk analyses, both for renewal need and potential cooperation, are usually made by one or two sewer system managers within a municipality. The risk judgment can differ per person judging the risk and over time. This is caused by the absence of quantitative variables and a reference model for these risks.

Apart from pipe age, pipe status and cooperation with other public works, many other information sources are used for deciding upon renewal as indicated by the interviewees and described in the municipal sewerage plans. Examples are municipal water management strategy, citizens' complaints and maintenance reports. Although these information sources are relevant to base renewal decisions on, a uniform and structured procedure for this is absent.

The identified information sources from the interviews do not differ per organizational level. The interviewees often indicated that in practice, working activities per organizational level overlap to a large part. In fact, what is seen as the strategic level of sewer asset management is in practice a tactical task.

55.4 Conclusions and Recommendations

The goal of this study was to assess the availability and use of information in decision making for sewer system renewal in the Netherlands. The control paradigm of De Leeuw was chosen as a reference model to evaluate sewer asset management, because this model formed the basis for developing it in the Netherlands. Based on the preconditions for effective control, it is concluded that current sewer asset management in the Netherlands is ineffective, although service availability is good.

Precondition one: an evaluation mechanism for control actions or fulfillment of goals is absent. Next to that, the influence of intuition hampers justification, accountability and repetition of decisions, preventing evaluation. Precondition two: a model of the controlled system for predicting the effect of control actions is absent. Although hydraulic modeling is often applied, no mechanism is available to predict the effect of various control actions. Precondition three: information of the system and environment is collected, but a structured and uniform procedure to transform information into control actions is absent. The main question is what information is needed for sewer asset management.

For further organizational analysis of sewer asset management, it is recommended to compare De Leeuw's control paradigm with other theories of organizational configuration. It is recommended to implement evaluation mechanisms in order to evaluate control actions, fulfillment of goals and cost-efficiency of control. Next to that, sewer system managers need a critical attitude to judge the relevancy and quality of information.

Acknowledgments The authors would like to acknowledge the funding by (in alphabetical order) ARCADIS, DHV, Gemeente Almere, Gemeente Breda, Gemeente 's-Gravenhage, Gemeentewerken Rotterdam, GMB Rioleringstechnieken, Grontmij, KWR Watercycle Research Institute, Royal Haskoning, Stichting RIONED, STOWA, Tauw,Vandervalk & De Groot, Waterboard De Dommel, Waterboard Vallei & Eem, Waternet and Witteveen + Bos as part of the Urban Drainage Research program.

References

1. IAM, BSI (2008) Pas 55-2: 2008 Asset management, part 2: guidelines for the application of PAS 55-1. Institute for Asset Management and British Standard Institution, Bristol
2. Brown RE, Humphrey BG (2005) Asset management for transmission and distribution. IEEE Power Energy Mag 39–45
3. Thissen WAH, Oomens AJ (1991) Systems engineering for sewer management. In: IEEE international conference on systems, man, and cybernetics, decision aiding for complex systems
4. Oomens A (1992) Rioleringsbeheer: het structureren van het beheerproces aan de hand van de voorwaarden voor effectieve besturing, (Sewer management: structuring the management process on the basis of conditions for effective control). Delft University of Technology, Delft
5. De Leeuw ACJ (1974) Systeemleer en organisatiekunde. Een onderzoek naar mogelijke bijdragen van de systeemleer tot een integrale organisatiekunde, (Systems theory and organizations theory, a study about possible contributions of systems theory to an integral theory of organizations). Eindhoven University of Technology, Eindhoven. H.E. Stenfert Kroese b.v., Leiden, The Netherlands
6. Katz D, Kahn RL (1978) The social psychology of organizations, 2nd edn. Wiley, New York
7. European Committee for Standardization (CEN) (2008) EN 752 Drain and sewer systems outside buildings, Brussels, Belgium
8. Osman H, Atef A, Moselhi O (2012) Optimizing inspection policies for buried municipal pipe infrastructure. J Perform Constructed Facil 26(3):345–352
9. Sægrov S, Hafskjold S, Köning A, Røstum J, Selseth I, Sjøvold F (2006) Wastewater network challenges and solutions. Integr Urban Water Resour Manag 10:147–158
10. Scheidegger A, Hug T, Rieckermann J, Maurer M (2011) Network condition simulator for benchmarking sewer deterioration models. Water Res 45(16):4983–4994
11. Wijnia YC, Herder PM (2009) The state of asset management in The Netherlands. In: Proceedings of the 4th world congress on engineering asset management and intelligent maintenance systems (WCEAM 2009), Athens, Greece. 28–30 Sept 2009
12. Arthur S, Crow H, Pedezert L, Karikas N (2009) The holistic prioritisation of proactive sewer maintenance. Water Sci Technol 59(7):1385–1396
13. Chughtai F, Zayed T (2008) Infrastructure condition prediction models for sustainable sewer pipelines. J Perform Constructed Facil 22(5):333–341
14. Ana EV, Bauwens W, Pessemier M, Thoeye C, Smolders S, Boonen I, De Gueldre G (2009) An investigation of the factors influencing sewer structural deterioration. Urban Water J 6(4):303–312
15. Arthur S, Burkhard R (2010) Prioritising sewerage maintenance using inferred sewer age: a case study for Edinburgh. Water Sci Technol 61(9):2417–2424
16. Yang MD, Su TC (2006) An automation model of sewerage rehabilitation planning. Water Sci Technol 54(11–12):225–232
17. De Leeuw ACJ (2002) Bedrijfskundig management. Primair proces, strategie en organisatie. (Corporate management. Primary process, strategy and organization). 2nd edn. Assen, The Netherlands, Koninklijke Van Gorcum BV

18. Statistics Netherlands (2012) Statline. [Online database]. Available from: http://statline.cbs.nl/
19. Statistics Netherlands (2009) Demografische kerncijfers per gemeente 2009. (Demographics key figures per municipality 2009). The Hague and Heerlen, The Netherlands, p 68
20. European Committee for Standardization (CEN) (2003) EN 13508-2 Conditions of drain and sewer systems outside buildings—part 2: visual inspection coding system
21. Dirksen J, Clemens FHLR, Korving H, Cherqui F, Le Gauffre P, Ertl T, Plihal H, Müller K, Snaterse CTM (2011) The consistency of visual sewer inspection data. Struct Infrastruct Eng 9:1–15
22. Korving JL (2004) Probabilistic assessment of the performance of combined sewer systems. Delft University of Technology, Delft
23. Dirksen J, Baars EJ, Langeveld JG, Clemens FHLR (2011) Settlement as a driver for sewer rehabilitation. In: 12th international conference on urban drainage, Porto Alegre, Brazil. 11–16 Sept 2011
24. Dirksen J, Ten Veldhuis JAE, Clemens FHLR, Baars EJ (2010) Sensible sewer system rehabilitation using information on sewer system settlement. In: Novatech, Lyon, France. 27 June-1 July 2010
25. RIONED Foundation (2010) Riolering in beeld. Benchmark rioleringszorg 2010. (Benchmark sewer system management 2010). Bennekom, The Netherlands

Chapter 56
Managing Knowledge for Asset Management: Shifting from Process to Relational Frames

Anna Wiewiora, Kerry Brown, Subas P. Dhakal and Muhammad Nateque Mahmood

Abstract The purpose of this paper is to review existing knowledge management (KM) practices within the field of asset management, identify gaps, and propose a new approach to managing knowledge for asset management. Existing approaches to KM in the field of asset management are incomplete with the focus primarily on the application of data and information systems, for example the use of an asset register. It is contended these approaches provide access to explicit knowledge and overlook the importance of tacit knowledge acquisition, sharing, and application. In doing so, current KM approaches within asset management tend to neglect the significance of relational factors; whereas studies in the KM field have showed that relational modes such as social capital is imperative for effective KM outcomes. In this paper, we argue that incorporating a relational approach to KM is more likely to contribute to the exchange of ideas and the development of creative responses necessary to improve decision making in asset management. This conceptual paper uses extant literature to explain KM antecedents and explore its outcomes in the context of asset management. KM is a component in the new integrated strategic asset management (ISAM) framework developed in conjunction with asset management industry associations (AAMCoG 2012) that improves asset management performance. In this paper, we use Nahapiet and Ghoshal's [24] model to explain antecedents of relational approach to KM. Further, we develop an argument that relational KM is likely to contribute to the improvement of the ISAM framework components, such as organizational strategic management, service planning, and delivery. The main contribution of the paper is a novel and robust approach to managing knowledge that leads to the improvement of asset management outcomes.

A. Wiewiora (✉) · K. Brown · S. P. Dhakal · M. N. Mahmood
Southern Cross University, Gold Coast Campus, Southern Cross Drive,
Bilinga, QLD 4225, Australia
e-mail: anna.wiewiora@scu.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_56,

56.1 Introduction

Thus far, in managing engineering assets, there appears to be little cognizant when adopting integrated asset management approach that reflects the processes and interrelations between economics, engineering, information technology, sustainability and human elements of the organization. Only recent research has begun to draw attention to the importance of integrated asset management approaches and calls for introducing human aspect into the management of organizational assets [1–3]. For instance, Schuman and Brent [3] emphasized that an early involvement of multiskilled people from the operating, production, and maintenance instills a sense of ownership in the asset project, suggesting that addressing concerns and viewpoints from multiple stakeholders involved in the project influence better decisions [3]. In these recent studies, asset management has been viewed as a socioeconomic system including social components (culture, trust, social networks, leadership, and culture, knowledge) and technical components (machinery, plant, etc.) [2]. Interactions between these components determine the direction of system development, thus both components, social and technical, should be given equal relevance to ensure integrated asset management outcomes [2]. This means that the importance of the human issues, so far overlooked in asset management, has to be taken into consideration to ensure improved asset performance [1].

Integrated Strategic Asset Management Guide (ISAM Guide) developed in conjunction with asset management industry associations Australian Asset Management Collaborative Group (AAMCoG) [4] provides a contemporary outline to assist those responsible for delivering and managing built assets to meet community and service delivery needs. ISAM Guide focuses on an integrated approach to managing assets by bringing together economics, engineering, information technology, sustainability, and human elements and recognizing interrelationships and interdependencies between these elements.

This paper aims to explain how the integration of these elements can be achieved through effective knowledge management (KM) efforts. According to AAMCoG [4], KM involves information systems and effective KM processes and underpins the capacity to develop new ways of thinking and creative responses that are necessary to improve decision making and increase productivity. Nevertheless, KM in asset management domain is still immature. Many firms do not consider asset management as a business approach that integrates business activities. Therefore, knowledge related to asset management is often based on organizational information sources, which have been designed for other than asset management purpose [5]. Furthermore, existing approaches to managing knowledge for asset management focus primarily on the application of data and information management systems, i.e., asset register. Although these systems provide quality and timely data for decision makers, they primarily contribute to management of explicit knowledge and overlook the importance of tacit knowledge. The lack of relational approaches for managing tacit knowledge means that KM for asset management is only fragmentary. Such incomplete KM practices are likely to result in suboptimal asset management performance.

56.2 Knowledge Management

In recent years, companies function in a rapidly changing and a knowledge intensive environment. In these conditions, firms need to be highly competitive to achieve continuous growth in the industry. To accomplish this, companies need to ensure the best use of their organizational knowledge. This can be achieved through KM that enables effective organization of knowledge in a company providing specified process for acquiring, organizing, and sharing both tacit and explicit knowledge [6, 7]. The overall purpose of KM is therefore to maximize the enterprise's knowledge-related effectiveness and returns from its knowledge assets and to achieve competitive advantage [8].

In asset management, the importance of KM has also been acknowledged by existing guidelines and standards [e.g., 9, 10]. However, these standards focus mostly on asset data and information management, giving limited attention to KM, and overlooking a significant distinction between knowledge, data, and information.

56.2.1 Knowledge—Information—Data

Knowledge is a multifaceted concept with multilayered meanings [11]. It has been described as a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information [12 p. 137]. Knowledge originates in the minds of knowledge holders and can be transferred into documents, organizational routines, processes, practices, and norms. It is necessary to distinguish knowledge from two other dimensions: data and information. In asset management, these terms are sometimes used interchangeably to knowledge; however, their scopes differ significantly.

Data are a set of discrete, objective facts about events. There is no meaning in data. Data provide no judgement or interpretation or basis of action [12]. Information is a message, usually in the form of a document or an audible or visible communication. It has a sender and a receiver and moves around organization through hard and soft networks. Unlike data, information has a meaning. Data becomes information when its creator adds meaning, for example by contextualizing, condensing, or categorizing it [12]. Once the information is used and becomes actionable, it is transformed into knowledge [12]. In other words, knowledge is created and organized by the flow of information attached to the commitment and beliefs of its holder [11]. When knowledge is learned and shared among individuals and adapted in organizational processes, it becomes a valuable, intangible asset.

In asset management, all three, data, information and knowledge, are necessary. At several stages of the asset life cycle, information is required on the condition of the assets. Knowing what to measure, how to measure it, and what to do with the information becomes very important. Often information must be maintained for many years in order to identify long-term trends [5]. There is a range of asset information systems available that provide access to different types of information captured in documents, drawings, and photographs. These systems allow to record work activities related to an asset, forecast asset demand, capture data related to asset performance, serviceability, asset location, and monitoring of asset condition; providing all sorts of asset attributes, e.g., make, model, serial number, age, capacity, and subjective information about the asset [5, 9]. Although maintaining accurate and quality asset data and information is imperative, the ultimate purpose for collecting data and information is often to make an informed decision. This can be done only by making a meaning out of data and information and translating it into knowledge, which combines experience, values, information in context, and insight, thus forming a basis for decision making [9].

56.2.2 Relational Knowledge Management

Current approaches to managing knowledge in asset domain focus primarily on the application of data and information systems (i.e., asset register). It is contended these approaches provide access to explicit knowledge, but overlook the importance of tacit knowledge acquisition, sharing, and application. Furthermore, as highlighted by Laue et al. [1], these information systems deal primarily with categorizing and providing asset information, but they do not overcome barriers of KM related to education and communication, trust-building and team-enabling activities, or establishing a climate of continuous learning [13].

According to The Institute of Asset Management [9], the quality of knowledge as derived from the experience, values, information in context, and insight will affect the reliability and quality of decision making. So not just a good data, but good knowledge is required to forecasts future behaviors and assist in asset management decision making. Therefore, incorporating a relational approach to KM is more likely to improve KM behaviors while contributing to the exchange of ideas, thus improving decision making in asset management. We define relational KM in the context of asset management as:

> The ability to capture, share, apply and integrate experiences, values, contextual information, and expert insights through formal and informal channels in order (1) to improve the capacity needed to develop new ways of thinking and (2) to generate creative responses necessary to improve asset management decision-*making*.

Existing research found that social networks such as informal meetings, coffee breaks, and workshops are essential for effective knowledge sharing [14–18]. Databases have grown to large proportions, but are often underutilized as employees are much more likely to turn to peers and colleagues than to impersonal sources for necessary knowledge [18]. Additionally, Mintzberg [19] indicated that people prefer to turn to other people rather than documents for information. More recently, the same tendency has been found even for people with ready access to the Internet and their firm's IT-based knowledge repository [20]. Newell et al. [21] recognized that social networks and informal dialogue are more effective than IT techniques and that IT should only complement social networks in knowledge transfer activities.

56.2.3 Knowledge Management from the Lens of Social Capital

Social capital is a set of social resources embedded in relationships [22]. Social capital includes many aspects of a social context, such as social ties, trusting relations, and value systems that define relationships, thus facilitate actions of individuals located within that context [23].

In this paper, we present the view on relational KM from the perspective of social capital, based on the model outlined by Naphiet and Ghoshal [24]. The model refers to three dimensions of social capital: structural, cognitive, and relational. *The structural dimension* includes network ties, which represent connections among members participating in social exchange. Network ties provide access to resources based on the principle that 'who you know' affects 'what you know' [24]. Maintaining strong ties, demonstrated by frequent and close social interactions, allows actors to know one another, to share important information, and to create a common point of view [23]. Furthermore, specific network configurations, such as density, connectivity, and hierarchy, influence the way how knowledge is shared between members. Therefore, social relations based on maintaining healthy network systems are likely to perform more effectively than sophisticated information systems. This is because social relations reduce the amount of time and investment required to gather information improving the ability of personal contacts to provide information sooner than it becomes available to people without such contacts [24]. Early access to knowledge may be especially important in time-constrained project-driven environment, where asset management operates. *The cognitive dimension* of social capital refers to shared language, codes, stories, and metaphors and provides powerful means in communities for creating, exchanging, and preserving knowledge. Tsai and Ghoshal [23] argue that when organization members share the same vision, they can avoid possible misunderstandings in their communications, thus exchange knowledge more effectively. *The relational dimension* is characterized by a high level of trust, shared norms, obligation, and identification [24]. For example, existing research identified trust as an important factor for successful knowledge sharing [25–27]. Koskinen et al. [28] noted that the greater the level of trust, the greater the level of people accessibility, and the greater the chance knowledge is shared in the team. Furthermore, Inkpen and Tsang [25] agreed that an atmosphere of trust contributes to the free exchange of knowledge, because people do not feel they have to protect themselves from others' opportunistic behaviors.

Overall, existing research has shown that building social capital improves relational KM. Tsai and Ghoshal [23] found that the structural and relational dimensions of social capital were significantly related to the extent of resource exchange, which in turn led to product innovation. Furthermore, Levin and Cross [26] revealed that relational dimension of social capital, in particular the trust, improves transfer of knowledge between teams. From these past researches, it is therefore apparent that trust and strong ties influence the formation of social capital and improve KM. In this research, we explain how relational KM, influenced by social capital, improves asset management outcomes. Next section focuses on explaining how relational KM contributes to ISAM framework components, namely *organizational strategic management, service delivery and community needs, and expectations.*

56.3 Improving Asset Management through Relational KM

56.3.1 Organizational Strategic Management

According to ISAM Guide, organizational strategic management involves an understanding of governance, corporate policy, objectives, and corporate strategy. It gives effect to whole-of-government policy through service delivery helping to determine how the delivery of asset and service should occur and what is required [4]. Consequently, organizational strategic management involves the ability to access and leverage existing social and organizational relationships and utilization of organizational capabilities including skills, expertise, and knowledge.

Turoff [29] advises that strategic planning, such as policy and asset procedures development, should normally take place during sessions with a group of experts across organizational departments and levels including company board, asset managers as well as operation, maintenance, and engineering representatives. This is because input from a range of actors will ensure that the decision maker is making optimum choice on the policy development or adjustment [29]. Accordingly, the better the knowledge base upon which policies are built, the more likely they are to succeed [30, 31]. In particular, good asset management policies, goals, and plans are likely to emerge when there is a good amount of knowledge shared between policy makers and other stakeholders involved in asset management. Exchanging valuable expertise and insights between participants during the policy development process, while taking into account knowledge from other stakeholders involved in asset management, is then likely to ensure that all possible options have been put on the table for consideration. So the policies, asset management strategy, plans, and goals are attainable, relevant, and understood across the organization and beyond its boundaries.

Prior research identified that utilization of informal networks and personal knowledge supports the knowledge base for policy makers [32]. According to

Riege and Lindsay [31] when creating public policy, stakeholders may include a range of people or organizations whose interest may be positively or negatively affected, such as government or private organizations, local authorities, the general community, and other interested parties. Riege and Lindsay [31] argue that enhanced partnerships with those stakeholders provides a cost-effective way of obtaining good quality knowledge. Based on that, knowledge input from stakeholders affected by and engaged in asset management practices is likely to improve knowledge creation and enhance knowledge database for asset management. Corporate policy and strategy developed through ongoing knowledge sharing between decision makers and other stakeholders engaged in asset management activities, and followed by ongoing review, may then lead to improved organizational strategic management outcomes.

56.3.2 Service Delivery

In the context of asset management, service delivery is a transaction aimed at meeting the needs and expectations of clients [33]. Appropriate planning on strategic, tactical, and operational levels is therefore imperative to achieve desired service delivery outcomes [4] ensuring that asset in fact delivers required service on time, within the budget and to the right quality. Effective use of skills and knowledge of all the parties engaged in the asset management can assist in the service delivery planning process. For instance, expert involvement can provide useful knowledge and expertise in the development of acquisition, maintenance, operation, and disposal plans [4]. Accordingly, appropriate solutions for service delivery depend on a range of stakeholders, with diverse interest and influence. These stakeholders may include other government agencies, asset users, and the broader community [4].

Dawson [34] argues that to achieve high-value service delivery outcomes, it is important to partner up with other services providers. For this reason, firms often form temporary ventures—project-based organizations [35], which by blending internal and external skills, expertise, and knowledge, can meet a variety of client needs in a rapid and effective manner [36]. This approach is especially common in the public sector where government resolves its financial and resources constraints in the provision of services by using skills, knowledge, and resources of private organizations to increase the efficiency, effectiveness, and quality of facilities and services delivery [37].

A number of different parties engaged in project-based organizations, including project team members, contractors, subcontractors, clients, and other stakeholders, means that these entities have strong knowledge capabilities, but it also suggests that to make the best use of it, effective knowledge coordination in project-based organizations is essential. However, despite the clear benefit to involve a range of skills and capabilities in service delivery activities, government agencies face a challenge to integrate these skills and to achieve consistency and transparency when delivering projects. Dawson and Horenkamp [36] advises that building knowledge-based relationships between the agencies as well as between agency and contractors and other stakeholders will improve service delivery outcomes and enhance knowledge capabilities, thus facilitate intra-and interfirms knowledge sharing. To be effective, these relationships need to be based on mutual trust and willingness to disclose information while being open to new ideas. According to Dawson and Horenkamp [36], clients in such a knowledge-based relationships are less affected by price because they realize the role of open communication, collaborative teams, and complementary expertise in creating value.

56.3.3 Community Needs and Expectations

According to ISAM Guide, understanding community needs and expectations is vital when delivering services [4]. Building strong and trusting relationships with communities and other stakeholders engaged in asset management endeavors is likely to be a driving force for understanding needs and expectations necessary to identify service and asset demands. Nevertheless, government agencies only inform communities and do not actively engage them into decision-making process, whereas Riege and Lindsay [31] suggest that government needs to focus on two-way KM processes that is not only obtaining knowledge from communities, but also transfer certain knowledge back to them. It has to be also taken into account that different communities may have diversity views and varying capacities to interact effectively. Riege and Lindsay [31] brought to attention that some groups will be more than others capable at representing themselves and therefore capable of engaging effectively in service delivery with government. According to Friis [38], one of the main reasons for government failure in this area seems to be too much emphasis on technology, rather than management processes that encourage people to interact. Based on that, community participation, critical to understanding service demands, can be enhanced through relational approach to KM. This can be achieved by creating opportunities for face-to-face interactions that in a longer term build stronger relationships and enhance mutual trust between government agency and community groups. Accordingly, the relational KM approach will allow for more effective management of various and often conflicting expectations. In a view of what has been said, relational KM is expected to assist organization in achieving greater community participation, knowledge sharing, and demonstrate accountability.

56.4 How to Improve KM for Asset Management

This paper argued that focus solely on data and information management systems is insufficient for asset management to improve its outcomes and that shifting toward relational frames based on building strong and trusting relationships is more likely to facilitate the access to both tacit and explicit knowledge. Leveraging social and organizational relationships will allow a fully utilization of organizational capabilities including skills, expertise, and knowledge. This in turn will contribute to organizational strategic management, service delivery outcomes, and meeting community needs and expectations—components of ISAM.

There are number of ways to improve relational KM. Building social networks through the creation of formal and informal channels such as scheduled and unscheduled meetings, informal seminars, training sessions, plant tours has been found to improve the distribution of highly context-specific tacit knowledge [6]. In the context of asset management, this means that well-established social networks are more likely to create opportunities for novel ideas to emerge and improve decision making though better informed assessment of a problem at hand. One way to do this is to organize frequent meetings with stakeholders and community members during which asset and service delivery demands and expectations are openly discussed. Also, early involvement of parties engaged in later stages of asset management life cycle, including operation and maintenance representatives, is expected to improve decisions related to planning and design of the asset. Furthermore, when developing organizational asset management policies and strategy, it is recommended to facilitate sessions with a group of experts across organizational departments and levels including company board, asset managers as well as operation, maintenance, and engineering representatives. Using techniques such as focus groups or policy Delphi can assist in getting experts together and creating opportunities for knowledge sharing between them [39], thus assist in more desirable outcomes for strategic organizational management.

Another way to improve relational KM is to create environment for trust building. To intentionally create trust or manage another party's propensity to trust is a difficult task. However, it is possible to enhance conditions for trust building. Trust is a predictor of positive working relationships between stakeholders, increasing the willingness of various project stakeholders to cooperate [40]. Organization can create trust-building environment by reviewing organizational norms and practices that encourage or discourage the high frequency of interaction and collaboration, supporting and recognizing KS initiatives, endorsing and maintaining a friendly and non-competitive atmosphere at work while creating an atmosphere for learning and not blaming [18, 41]. Trusting relationships established among parties involved in the asset management planning processes are then likely to create environment of confidence and openness, thus positively contribute to problem solving.

Finally, it is important to recognize that relaying solely on relational KM may not be sufficient. As Cooper [42] noted, during face-to-face interactions, issues are often raised and forgotten because attention is diverted elsewhere. Decisions are made based on sketchy information that is not revisited. Opportunities are lost because no one is accounted to follow up on them [42]. Therefore, incorporating relational approach to KM into existing technology based mechanisms appears to be the most optimal way to bring the best outcomes when managing knowledge for asset management.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Laue M, Brown KA, Keast R (2011) Strategic and human issues in asset management models: 6th world congress on engineering asset management, Cincinnati, Ohio, USA
2. McGeoch M, Brunetto Y, Brown K (2011) Current issues in strategic asset management: world congress on engineering asset management (WCAEM), Cincinnati, USA
3. Schuman CA, Brent AC (2005) Asset life cycle management: towards improving physical asset performance in the process industry. Int J Oper Prod Manage 25:566–579
4. AAMCoG (2011) Guide to integrated strategic asset management: Australian Asset Management Collaborative Group (AAMCoG), Brisbane
5. Koronios A, Lin S, Gao J (2005) A data quality model for asset management in engineering organisations: In: Conference on information quality Cambridge, MA, USA
6. Alavi M, Leidner D (2001) Review: knowledge management and knowledge management systems: conceptual foundations and research issues. MIS Q 25:107–136
7. Love P, Fong P, Irani Z (2005) Management of knowledge in project environments. Elsevier/Butterworth-Heinemann, Oxford
8. Wiig K (1997) Knowledge management: where did it come from and where will it go? Expert Syst Appl 13:1–14
9. The Institute of Asset Management (2011) Asset management—an anatomy, vol 1. The Institute of Asset Management, Bristol, UK
10. New South Wales Treasury (2004) Asset information guideline
11. Nonaka I (1994) A dynamic theory of organizational knowledge creation. Organ Sci 5:14–37
12. Davenport TH, Prusak L (1998) Working knowledge. Harvard Business School Press, Harvard
13. James P (2005) Knowledge asset management: the strategic management and knowledge management nexus. In: DBA thesis, Southern Cross University, Lismore, NSW, p 239
14. Cook S, Brown J (1999) Bridging epistemologies: the generative dance between organizational knowledge and organizational knowing. Organ Sci 10:381–400
15. Foos T, Schum G, Rothenberg S (2006) Tacit knowledge transfer and the knowledge disconnect. J Knowl Manage 10:6–18
16. Liebowitz J (2005) Conceptualizing and implementing knowledge management: management of knowledge in project environments In: Love P, Fong P, Irani Z (eds) Elsevier/Butterworth-Heinemann, Oxford, pp 1–18
17. Newell S, Bresnen M, Edelman L, Scarbrough H, Swan J (2006) Sharing knowledge across projects: limits to ICT-led project review practices. Manage Learn 37:167–185
18. Abrams LC, Cross R, Lesser E, Levin DZ (2003) Nurturing interpersonal trust in knowledge-sharing networks. Academy of Management Executive 17:64–77
19. Mintzberg H (1973) The nature of managerial work. Harper & Row, New York
20. Cross R, Parker A, Borgatti S (2000) A bird's-eye view: using social network analysis to improve knowledge creation and sharing. Knowl Dir 2:48–61
21. Newell S, Goussevskaia A, Swan J, Bresnen M, Obembe A (2008) Interdependencies in complex project ecologies: the case of biomedical innovation. Long Range Plan 41:33–54
22. Burt RS (1992) Structural holes: the social structure of competition. Harvard University Press, Cambridge
23. Tsai W, Ghoshal S (1998) Social capital and value creation: the role of intra firm networks. Acad Manag J 41:464–476

24. Nahapiet J, Ghoshal S (1998) Social capital, intellectual capital, and the organizational advantage. Acad Manag Rev 23:242–266
25. Inkpen AC, Tsang EWK (2005) Social capital, networks, and knowledge transfer. Acad Manag Rev 30:146–165
26. Levin DZ, Cross R (2004) The strength of weak ties you can trust: the mediating role of trust in effective knowledge transfer. Manage Sci 50:1477–1490
27. Ding Z, Ng F, Cai Q (2007) Personal constructs affecting interpersonal trust and willingness to share knowledge between architects in project design teams. Constr Manag Econ 25:937–950
28. Koskinen KU, Pihlanto P, Vanharanta H (2003) Tacit knowledge acquisition and sharing in a project work context. Int J Project Manage 21:281–290
29. Turoff M (2002) The policy Delphi: updated version of Turoff M. In: The design of a policy Delphi, Technological Forecasting and Social Change, vol 2(2) (1970) (ed)
30. Althaus C, Bridgman P, Davis G (2007) The Australian policy handbook. Allen & Unwin, Sydney
31. Riege A, Lindsay N (2006) Knowledge management in the public sector: stakeholder partnerships in the public policy development. J Knowl Manage 10:24–39
32. Huysman M, de Wit D (2004) Practices of managing knowledge sharing: towards a second wave of knowledge management. Knowl Process Manage 11:81–92
33. Smith S (1998) How to create a plan to deliver great customer service: best practices in customer service. In: Zemke R, Woods JA (eds) AMACOM, New York
34. Dawson R (2007) Service delivery innovation, creating client value and enhancing profitability: SAP thought leadership SAP for professional services (ed)
35. Project Management Institute (2008) A guide to the project management body of knowledge: (PMBOK Guide), 4 edn. Project Management Institute Inc, USA
36. Dawson R, Horenkamp M (2007) Service delivery innovation, creating client value and enhancing profitability. In: SAP thought leadership SAP for professional services, SAP America Inc, Germany
37. Treasury HM (2000) Public private partnerships—the Government’s approach. HM Treasury, London
38. Friis CS (2002) Knowledge in public administration. In: Linz (ed) The 3rd international workshop by international federation for information processing Universitätsverlag Rudolf Trauner Copenhagen, pp 3–12
39. Turoff M (2002) The policy Delphi. Updated version of Turoff M. In: The design of a policy Delphi, technological forecasting and social change vol 2(2) (1970). http://is.njit.edu/pubs/delphibook/ch3b1.pdf
40. Pinto JK, Slevin DP, English B (2009) Trust in projects: an empirical assessment of owner/contractor relationships. Int J Project Manage 27:638–648
41. Wiewiora A (2011) The role of organisational culture, trust and mechanisms in inter-project knowledge sharing. In: PhD thesis: faculty of built environment and engineering, Queensland University of Technology, Brisbane, Australia
42. Cooper LP (2003) A research agenda to reduce risk in new product development through knowledge management: a practitioner perspective. J Eng Tech Manage 20:117–140

Chapter 57
Designing an Asset Management Guideline for the Dutch Wastewater Industry

Ype Wijnia and John de Croon

Abstract In the Netherlands, the collection and processing of wastewater is executed by 415 municipalities (collection) and 25 water boards (processing). Those organizations have signed an agreement to reduce costs. Implementation of asset management is supposed to play a vital role in this cost reduction. In order to facilitate the implementation of asset management, the representative body of the industry [Stichting RIONED and STOWA (Dutch acronym for the Foundation for Applied Water Research)] envisioned an asset management guideline tailored to the specific needs of the wastewater industry. However, due to the diversity in size and scope of the asset managers to be and the assets in scope, it was not certain whether a single guideline would be possible nor if it would be cost-effective. To get a clearer view on those issues before committing large sums of money, the process of creating the guideline was split up in two phases. The first phase would result in a road map: design of the guideline, high-level planning of its realization, and the contours of a business case for implementation of asset management by means of the guideline. The second phase would then consist of actually realizing the guideline. In this paper, we describe the process we followed to arrive at the road map for the asset management guideline. The following can be concluded. Although the technical content of asset management for wastewater collection and wastewater treatment is very different, the process of management assets does not have to be different. By focusing on the asset management process, there was no need to distinguish wastewater treatment from wastewater collection. By making the guideline a collection of relative independent elements, it would do no harm to the concept of the guideline if in future elements would be added or removed.

Y. Wijnia (✉) · J. de Croon
AssetResolutions B.V., P.O. Box 30113, 8003 CC Zwolle, The Netherlands
e-mail: ype.wijnia@assetresolutions.nl

Y. Wijnia
Faculty of Technology, Policy and Management, Delft University of Technology, Delft
The Netherlands

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_57,

57.1 Introduction

In the covenant 2011–2015 [1] (In Dutch: Bestuursakkoord 2011–2015) between central government, municipalities, and water boards, it has been agreed upon to reduce the cost of the administration. One of the focus areas is water, which has a separate arrangement, the Covenant Water [2] (In Dutch: Bestuurakkoord Water). Key agreement of this pact is to reduce the cost of the water chain by €380 million per year, on the total annual expenditure of about 3,000 million euro. This number is based on a fact-finding report on Efficient Water Management of 2010 [3]. In this report, the implementation of asset management is mentioned as a key success factor. The cost reduction of 380 million euro per year is "only" 13 % of total expenditure, and given the state of asset management in the industry [4] and experiences in other sectors, this should not be an unrealistic target. However, a complicating factor is that the wastewater industry is not a single entity.

The wastewater chain is managed by 25 water boards and 415 municipalities, which are all independent public authorities with separate democratic legitimacy. The (physical) cooperation between the municipalities and water boards is organized by means of some 60 wastewater regions, the so-called treatment circles. In general, treatment circles consist of one water board and multiple municipalities. However, the large municipalities often have more than one treatment circle, though the water board is not necessarily the same for all treatment circles within one municipality (borders between water boards and municipalities do not necessarily correspond). The role of the municipalities is to collect the wastewater (in essence a reversed distribution), whereas the water boards take care of the transportation of the collected effluent to the sewage treatment plants and the actual treatment. The treated water is discharged to surface waters. Within the industry, there are significant differences between involved parties. The municipalities range in size from 1,000 to 800,000 inhabitants [5]. Some parties (like the water boards and large municipalities) have a separate department for asset management; in smaller municipalities, asset management is generally one of the tasks of one or two employees. In one organization (Waternet in Amsterdam), collection, transportation, and treatment are executed in one hand.

Another complicating factor is that in the wastewater chain, two very distinct types of asset management are needed [6, 7]. Since the wastewater collection is a reversed distribution infrastructure, it possesses all characteristics that are typical for infrastructures [8]. The asset management to be applied should address the problems for infrastructure. The main cost in infrastructures is capital expenditure, about 60 % of the total [3]. The wastewater plants, on the other hand, are not infrastructures. They are much more like factories: The assets are not passive but operated, and operational costs are the major part of the total expenditure, about 55 % [3].

As a response to the covenants, a representation of the wastewater industry filed a proposal with STOWA (Dutch acronym for the Foundation for Applied Water Research) to prepare a guideline for the implementation of asset management. This guideline should help all organizations involved in the management of the wastewater chain with the design and implementation of their asset management practice.

There were two key challenges in the design of the guideline. The first challenge was with respect to the content. As mentioned, asset management for the wastewater collection is something different than for wastewater treatment and the guideline should respect that. On the other hand, given that both forms of asset management interact by means of the treatment circles (in an "n to n relationship" of water boards and municipalities), a standardized model would be very welcome.

The second challenge was purely the process. The representation that filed the proposal was just a working group, not a true representative body. Getting all involved parties to accept and implement the guideline would require more than just a good guideline. Choices would have to be documented carefully, and attention should be given to the "marketing" of the guideline and allow for future adaptations.

When the proposal was filed, it was not yet clear what such a guideline should look like. There seemed to be some implicit agreement on the form: the guideline should be a binder with separate sections, for example like the Infrastructure Management Manual [9]. However, it was suspected that creating such a guideline would require a significant effort. Therefore, the proposal splits the development of the guideline into two phases. Phase 1 was the development of a road map for the development of the guideline, and phase 2 would be the actual development of the guideline itself. In this paper, we describe the approach to arrive at the road map.

57.2 Approach and Framing

The approach that was followed in the development of the road map was that of divergence and convergence. First ideas on what could be in the guideline would be collected (the divergence), followed by clustering and ranking ideas to arrive at the elements that should be part of the guideline. Clustering, however, is more an art than a science, as Morgan showed in his paper on clustering risks [10]. The best clustering depends on the goal of clustering. As the end goal of implementing asset management is a (significant) contribution to the cost reduction of €380 million per year, the choice was made to structure the ideas along their "style" of cost reduction. This made it possible (at least in theory) to construct a business case for each of the deliverables, though in practice, it was hard to go beyond a qualitative assessment of what contribution a deliverable would make.

57.2.1 Styles of Cost Reduction

According to Porter [11], cost leadership is one of the three strategies that can be chosen in order to be successful. Given its importance in the business context, it is no surprise there is a wealth of literature on this topic. Earliest attempts to apply scientific principles to the design of production processes date back to scientific management of Taylor [12]. However, it was not the purpose of the project to identify all options for cost reduction in a top–down approach. A structure that would help in assessing the value of ideas would suffice. The structure that "emerged" during the project is shown in Fig. 57.1.

The first style of cost reduction is improved decision making. If an organization is solving the wrong problem, only thinks of the standard solutions and does not optimize on the overall value, the projects the organization executes will be a waste of money, no matter how well they are executed. This style is thus about doing the right things.

But doing the right things is not the only style of cost reduction, there is in general also room for improvement in the way things are done, i.e., doing the things right. This style is about improving the efficiency of production and has to do with the organization of the work and the coordination between people. Borrowing form "Structures in five" [13], there are three applicable coordination mechanisms: standardization of the work processes, standardization of the output, and standardization of competences. In short, standardization of work processes prescribes how things should be done and is in terms of efficiency improvement preferred, culminating into fully automated production plants. But this form of standardization is not always possible; sometimes local circumstances dictate that other methods should be used in order to arrive at the same output. In that case, it is better to let the operator decide how to do things, and just prescribe the output that should be delivered. This is called standardization of output. However, in some cases, it is not even possible to prescribe what should be made, as that is again determined by local circumstances. The operator then has to decide both on what and how. The only coordination mechanism available for those circumstances is standardization of competences, i.e., making certain that the operator has the right knowledge and skills to make the right decision. In practice, organizations tend to use all forms of coordination simultaneously.

The third style of cost reduction is synergy. Some organizations active in the wastewater industry are relatively small, and combining activities may help spread some fixed overhead over a larger number of activities, thus lowering the average cost. This synergy can be achieved in two directions. Either the asset management scope is extended (e.g., water boards also manage flood protection and waterways; municipalities manage roads, buildings, and so on), or the organizational scope is enlarged, for example by mergers between water boards and between municipalities. The past years have shown a trend for merging. In 2005, the Netherlands contained 467 municipalities, per 1 January 2012 the number is down to 415 [14]. With regard to the water boards, in 1950, there were 2,600, in 2012 only 25 [15].

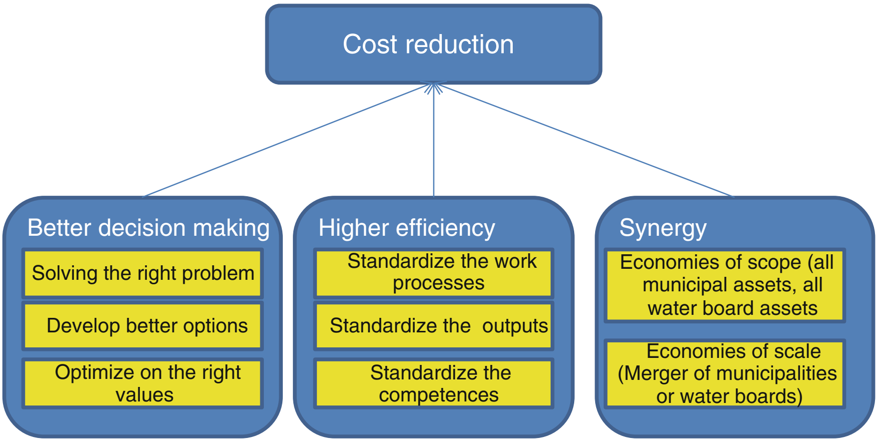


Fig. 57.1 Styles for cost reduction

57.2.2 The Road Map and the Guideline

Between the idea of implementing asset management and the actual realization of the savings is a number of steps. The general scheme for the realization of intended effects is given below. The process starts with an idea to do something new or different. Based on that idea, a project is developed, culminating (according to PRINCE2) into a Project Initiation Document (PID). This PID is a detailed plan of action for producing the project deliverables. Once the PID has been approved, the realization starts. The project is finished if all the deliverables are completed. But this is not the realization of the intended effect. Those are only captured once the organization starts using the project deliverables (Fig. 57.2).

If this scheme is translated to the context of the asset management guideline, the road map can be regarded as the project design, and the guideline itself as the deliverable. However, where a project is normally indivisible, this is not necessarily the case for the guideline. Some elements will have value on their own and will not depend on other parts. A better maintenance strategy will provide benefits even if investment decisions are still made the old-fashioned way. It does not mean the value of the whole cannot be higher than the sum of the values of the parts (like moving from optimized investment and maintenance decisions to optimization of the total cost of ownership), but there is value in the parts as well. The guideline therefore is more like a program (a collection of projects) than a single project. As a result, the road map is not a detailed project design, but the specification of the subprojects that make up the guideline, including their dependencies. The plans for the subprojects are out of scope. On the other hand, many of the subprojects will be relatively small (10s of days) which makes it hard to justify a full PID for every single one of them.

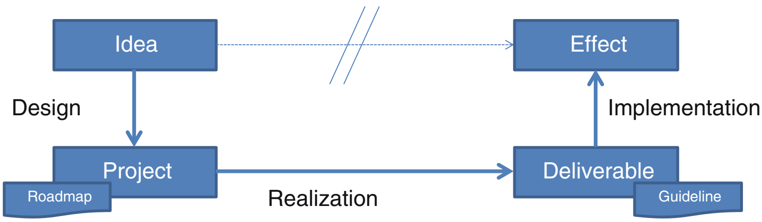


Fig. 57.2 From idea to effect

57.3 Results

The first step of the process was to capture ideas that should be addressed in the guideline. However, the brainstorm was not a typical brainstorm with participants writing down term on post-its and putting them on the wall. Instead, they were asked to write down (in a prepared template) the idea with a detailed description and a motivation. Only the idea was put on a post-it, but the filled in template was used in processing the session. This way much more of the rationale behind the idea could be captured. The brainstorm resulted in some 180 ideas, which were grouped into 34 clusters in the brainstorm session itself. Not all (clusters of) ideas could be translated directly into deliverables. Some ideas were more like questions than solutions. As the road map needed to be a plan for the realization of deliverables, some offline work was required. This resulted in 53 different potential deliverables. Each of the inputs of the brainstorm was linked to one of those deliverables.

However, not all potential deliverables would necessarily have to be part of the guideline. Selection of the potential deliverables was the subject of a separate session. In preparation for this convergence session, the list of potential deliverables was sent to the participants. The request was to grade the importance of the deliverable for the guideline on a scale from 0 (insignificant) to 10 (condition sine qua non). About 80 % of the participants fulfilled the request. The scores are plotted in the diagram below (Fig. 57.3). The red line indicates the average grade, green is the high score (average plus one standard deviation), blue is the low score (average minus one standard deviation), and the purple line was the mark. The scores were split into 4 groups. Group one consisted of the elements for whom both the average and the low score were above the mark. There could be a discussion on the deliverable, but essentially everybody agreed that it should be included. The second group consisted of deliverables with an average above the mark, but a low score below the mark. Those deliverables were nominated to be included, but apparently some participants disagreed, so a discussion was needed. In this discussion, focus would be on finding arguments for not including it. The third group was the opposite: it was nominated not to be included, but some participants saw the deliverable as important. The fourth group, finally, consisted of deliverables that should not be in the road map, as both the average and the high

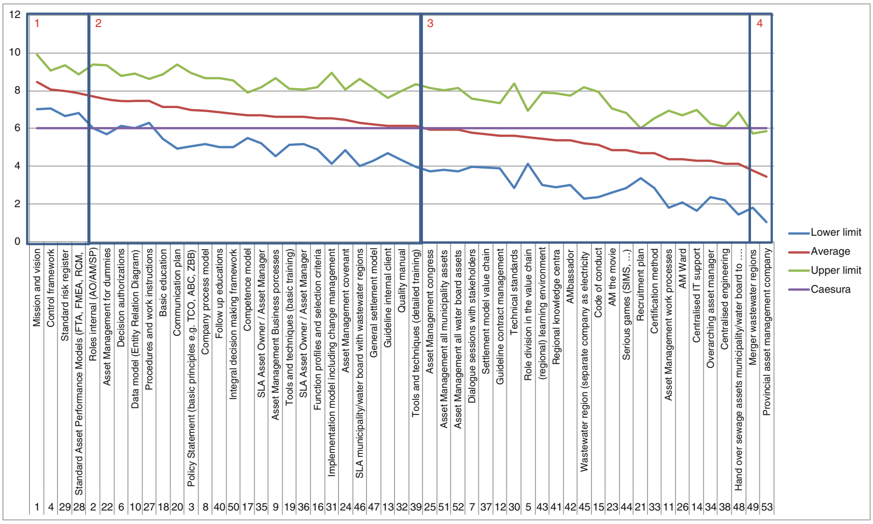


Fig. 57.3 Results of prioritization

score were below the mark. The arguments that were exchanged during the discussion were recorded as comments in the product description. The end result was a selection of 24 deliverables to be included in the road map.

Based on this prioritization, a proposal for the road map was made. The results are given in Table 57.1. The scope of the road map is a little bit wider than only the guideline itself, some elements for marketing the guideline and a proof of concept of the guideline are included as well. These are called “enabler” (see column “Type”).

The products match many requirements of the asset management standard PAS55 [16]. According to this standard, it is required to have an asset management policy and a related strategy, objectives, and plans. The products such as mission and vision, the policy statement, and the control framework are an answer to these requirements. The process models, procedures, work instructions, authorizations, and approvals cover quite a lot of sections of PAS55. Think of requirements on information management, asset management enablers and controls, structure, authority and responsibilities. Besides, the standard captures explicit requirements on tools. Specific products which meet these requirements include the standard risk register and standard performance models. Outsourcing of asset management activities is both covered in the process models as in the guideline internal client. The implementation of all products is supported in the implementation model and the (regional) learning environment. PAS55 (version 2008, p. V) suggests that the awareness of the employees, competencies, and cross-functional coordination are essential for the implementation of good asset management. In order to (at least partially) meet these requirements, attention is given to this with the competence model, the basic asset management education, the learning environment, and

Table 57.1 Elements of the guideline

No.	Product	Aim	Type
1	Mission and vision	Ensure that the organization knows the asset management direction	Content
2	Roles internal (AO/AM/SP)	Ensure that in the organization, it is clear what the roles, responsibilities of the roles asset owner/asset manager/service provider (AO/AM/SP) are	Content
3	Policy statement (basic principles)	Ensure that the organization knows what the basic principles are as well as the way of work	Content
4	Control framework	Ensure that members of the organization know the decision-making framework	Content
5	Authorizations and approvals	Make sure that the decision powers of the AO/AM/SP are clear	Content
6	SLA asset owner/asset manager	Ensure that it is clear what results the asset manager is expected to provide to the asset owner	Content
7	SLA asset manager/service provider	Ensure that it is clear what results the service provider is expected to provide to the asset manager	Content
8	Tools and techniques (basic principles)	Providing tools for selecting the right asset performance models and the selection of continuing education	Enabler
9	(Regional) learning environment	Ensure that the people who work in the sector in asset management learn in a structured way from each other supported by a content expert	Enabler
10	Company process model	Make sure that in the organization is clear what processes exist	Content
11	Asset Management processes	Make certain that in the organization is clear how the processes are linked to each other	Content
12	Asset management detailed processes	Ensure that the organization is known in detail what the outputs, what responsibilities go with it and which information needs are related to the processes	Content
13	Standard risk register	Ensure that wherever possible the same risks are assessed so that the risk levels and mitigation measures can be compared	Content
14	Implementation model	Ensure that in implementation of asset management, the organization is not reinventing the wheel	Enablers
15	Standard asset performance models	Make sure that common decisions are taken the same way	Content
16	Procedures and work instructions	Ensure that people who work within asset management know how they should perform relevant activities	Content
17	Data model	A description which documents and organizes the business data for communication between team members including the entities and the relations	Content
18	Guideline internal client	Providing tools for designing and implementing the interfaces between asset manager and service provider	Enablers

(continued)

Table 57.1 (continued)

No.	Product	Aim	Type
19	Basic education	To provide a quick insight into asset management to all those who work or will work within asset management	Enablers
20	Competence model	Providing a tool in selecting people for certain asset management functions and assist in the preparation of the personal development plans	Enablers
21	Asset management for dummies	Informing stakeholders (e.g., directors) on the operation and the importance of asset management	Enablers
22	Communication plan	Ensure that all relevant stakeholders in the right way at the right time can be informed	Enablers
23	Asset management congress	Giving the kick off for the trial run of asset management	Enablers
24	Asset management covenant	Giving respect and status to the introduction of asset management in the water chain	Enablers

education on tools and techniques. Asset management for dummies supports information requirements of relevant stakeholders. These need to be taken into account according to the standard. Their demands will be captured in the control framework as well. On page VI of the standard, important interfaces such as motivation, communication, leadership, and teamwork are mentioned. The products communication plan, asset management congress, and asset management covenant fill in these requirements. The implementation plan will also deal with change management. When the products are defined and implemented, then the relative companies have taken a major step toward PAS55 certification. It, however, should be noted that certification is not the ambition.

57.4 Conclusion

Based upon this project, two conclusions can be drawn. First, even though the technical content of asset management for wastewater collection and wastewater treatment is very different, the group arrived at an understanding that the process of managing assets does not have to be different. This is further emphasized by a (more or less) shared value system between municipalities and infrastructures as they both represent the public interest. That meant that models for optimization could be similar, even though for municipalities focus would be on investment decisions and for water boards on maintenance and operations decisions. Summarizing this, by focusing on the asset management process, there was no need to distinguish wastewater treatment from wastewater collection.

The second conclusion is that approach allowed for future flexibility. By making the guideline, a collection of relative independent elements, it would do no harm to the guideline if in future elements would be added or removed. Given the explicit linking of the elements (both selected and non-selected) to the ideas of the initial brainstorm session, such a future discussion could be easily facilitated. Combined with the attention for marketing the guideline, this provided the group with enough trust that actually making the guideline would be valuable effort.

Acknowledgments The authors would like to thank all participants of the working group for their valuable contributions and discussions. Without them, the result would not have been possible.

References

1. van Nederlandse Gemeenten V (2011) Rijksoverheid. Interprovinciaal Overleg, Waterschappen Uv, Bestuursakkoord 2011–2015
2. van Nederlandse Gemeenten V, (2011) Rijksoverheid. Interprovinciaal Overleg, Waterschappen Uv, VEWIN, Bestuursakkoord Water Den Haag
3. Nederlandse Gemeenten V, Van Waterschappen U, Rijk (2012) Commissie feitenonderzoek. Doelmatig beheer waterketen, VEWIN
4. Wijnia YC, Herder PM (2009) The state of asset management in the netherlands, In: World conference on engineering asset management. Springer, Athens
5. Metatopos (2012) Alle 415 gemeenten in 2012. (cited 2012 september 11); Available from: http://www.metatopos.eu/gemprov.html
6. Wijnia YC, de Croon JAW, Lyanage JP (2011) Towards an asset management reference model: basis for a unified approach. In: The 6th world conference on engineering asset management Cincinnati
7. Wijnia YC, de Croon JAW, Lyanage JP (2011) Application of a unified reference model across asset types: comparative cases. In: The 6th world conference on engineering asset management Cincinnati
8. Herder PM, Wijnia Y (2011) A systems view on infrastructure asset management, in asset management: the state of art in Europe from a life cycle perspective, In: van der Lei T, Herder P, Wijnia Y, (eds) Springer, Dordrecht
9. IAM (2002) International infrastructure management manual. V 2.0 UK 2002 ed: UK Institute of Asset Management
10. Morgan MG, Florig HK, Dekay ML, Fischbeck P (2000) Categorizing risk for risk ranking. Risk Anal 20(1):49–58
11. Porter ME (1980) Competitive strategy. Free Press, New York
12. Taylor FW (1911) Principles of scientific management. Harper & brothers, New York
13. Mintzberg M (1983) Structures in five: designing effective organisations. Prentice Hall, Englewood cliffs
14. Wikipedia (2012) Lijst van Nederlandse gemeenten. Available from: http://nl.wikipedia.org/wiki/Lijst_van_Nederlandse_gemeenten. Accessed 11 Sept 2012
15. Van Waterschappen U (2011) Factsheet kerngegevens waterschappen 2011
16. Institute of Asset Management (2008) Asset management part 1: Specification for the optimized management of physical assets, in PAS 55-1, BSI, Editor

Chapter 58
Empirical Hazard Function Using Continuous-Time Failure Data

Gang Xie, Fengfeng Li, Yong Sun and Lin Ma

Abstract A theoretically sound and accurate empirical hazard function may be used directly for analysis of lifetime distribution of the continuous-time failure data or can be used as a basis for further parametric modeling analysis in asset management. For the sake of bridging the gaps between probability theory and data analysis practice, this paper starts from clarifying the relationship between the concepts of hazard function and failure rate. Then, two often-used continuous-time data empirical hazard function formulas are derived directly from discrediting their theoretic definitions of the hazard function. The properties of these two different formulas are investigated, and their estimation performances against the true hazard function values are compared using simulation samples from exponential and Weibull distributions. Under the specified assumption conditions, both theoretic calculation and simulation results show that the formula that calculates the average failure rates (AFR) gives less biased estimation than the other one in all cases we have examined. We also showed that under practical situations, the relative difference of the calculated empirical hazards between these two formulas is less than 6 %. Based on the result of this study, we proposed a rule of thumb for applications of these two most often-used empirical hazard function formulas in data analysis practice.

58.1 Introduction

Hazard function plays an essential role in the application of probability theory in engineering reliability study. For example, the mean time to failure (MTTF) is calculated as the inverse of hazard rate if we assume that the asset system lifetime

G. Xie (✉) · F. Li · Y. Sun · L. Ma
School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: john.xie@qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_58,

distribution follows an exponential distribution. In the data analysis stage for asset management, however, the term failure rate is more often used when we try to work out the MTTF. In a sense, there is a gap between probability theory and data analysis when we talk about *hazard function* and *failure rate*. Because people can be confused with the questions like, are these two terms interchangeable? If the answer is yes, why not just use one of the terms; if they are different, what are the differences? A short answer is hazard or hazard rate $h_i \equiv h(t_i)$ is the instantaneous failure rate (for non-repairable asset systems) at a time instant $t_i\, i = 1, 2, \cdots$. However, when we talk about failure rate in data analysis, it is more often a short term for *average failure rate* (AFR) over a time period $t_2 - t_1$ (assuming $0 \le t_1 < t_2$). We know that AFR can be calculated using the formula [1]:

$$\mathrm{AFR} = \frac{\int_{t_1}^{t_2} h(u)du}{t_2 - t_1} \tag{58.1}$$

Equation 58.1 is nothing but the average hazard function formula which is considered as the most typical estimation of the true hazard function values [2]. Therefore, we need an empirical hazard function formula so that we can estimate the hazard function $h(t)$ based on observed sample data.

We may treat sample failure-time data as discrete data, i.e., we consider the observed sample failure times as the events that occur at preassigned times, $0 \le t_1 < t_2 < \cdots$, and that under a parametric model of interest, the hazard function at t_i is $h_i = h(t_i|\theta)$. Let us consider a set of intervals $I_i = [t_i, t_{i+1})$ covering $[0, \infty)$ for an engineering asset system with N functional components at $t_1 = 0$. Let us also denote $d_i = N(t_i) - N(t_{i+1})$ where $N(t_i)$ and $N(t_{i+1})$ are the number of components which are functional at time t_i and time t_{i+1}, respectively. Then, the quantity d_i is the number of failures in interval I_i and $r_i \equiv N(t_i)$ is the number of components at risk (i.e., having the potential to fail) at t_i. It can be shown that the maximum likelihood estimator (MLE) is

$$\hat{h}_i = \frac{d_i}{r_i}, \tag{58.2}$$

from which the well-known *Kaplan–Meier* estimator for the reliability function

$$\hat{R}(y) = \prod_{i:t_i < y} (1 - \hat{h}_i) = \prod_{i:t_i < y} \left(1 - \frac{d_i}{r_i}\right),$$

is derived. Equation 58.2 is valid under independent right censoring [3] (pp. 193–197) and [4] (pp. 268–270). Note that *Kaplan–Meier* estimator is also valid for randomly censored data. For the randomly censored data, the formula for the calculation of d_i should be modified as

$$d_i = N(t_i) - N(t_{i+1}) - N_{c(i)}, \quad (58.3)$$

where $N_{c(i)}$ is the number of components being censored in interval I_i.

In data analysis practice, we may be interested in treating the sample failure-time data as continuous-time data as shown in Eq. 58.1. Two often-used empirical hazard function formulas for treating the continuous-time data are

$$\hat{h}_i = \frac{N(t_i) - N(t_i + \Delta t)}{\Delta t \cdot N(t_i)} = \frac{1}{\Delta t}\frac{d_i}{r_i} \equiv \widehat{h1}_i, \quad (58.4)$$

and

$$\widehat{h_i} = -\frac{1}{\Delta t}\log\left[1 - \frac{N(t_i) - N(t_i + \Delta t)}{N(t_i)}\right] = -\frac{1}{\Delta t}\log\left(1 - \frac{d_i}{r_i}\right) \equiv \widehat{h2}_i, \quad (58.5)$$

where log represents the natural logarithm operation. The notation $\Delta t \equiv t_{i+1} - t_i$ ψ is used to emphasize that failures can happen at any time instants, not necessarily at t_i $i = 1, 2, \cdots$ under the continuous-time data setting. The same cares need to be taken in applying Eqs. 58.4 and 58.5 when calculating the empirical hazards for the censored data. Equation 58.3 needs to apply in calculating d_i.

At a glance, Eqs. 58.4 and 58.5 are very different. When people need to make a decision in choosing one of the above two formulas for the calculation of empirical hazard function, questions like "which one should I use and why" are naturally asked. In addition, industry people may have a good chance of not knowing; hence, they want to know how Eqs. 58.4 and 58.5 relate to Eq. 58.1. These questions are necessary to be answered for correctly estimating the true hazard function values using sample failure-time data in asset management practice. These questions are not theoretically difficult, but it seems that they have been ignored so far in publication. This paper is aiming at filling this gap.

The rest of the paper is organized as follows. In Sect. 58.2, we derive Eqs. 58.4 and 58.5 directly from discretizing their theoretic definitions of the hazard function followed by a detailed discussion on the properties of these two formulas in terms of estimation of the true hazard function values. In Sect. 58.3, we verify our theoretic results by calculating the empirical hazards based on two simulation samples—one generated from an exponential distribution and the second one generated from a Weibull distribution. Section 58.4 concludes this paper with a proposed rule of thumb for applications of Eqs. 58.4 and 58.5 in engineering reliability data analysis practice.

Throughout this paper, the open source statistical package R [5] is used for data analysis.

58.2 Empirical Hazard Function Derivation and Discussion

The empirical hazard function formulas can be derived in various ways. For example, Eq. 58.4 was given in [6] and [7]; Eq. 58.5 was derived from the discussion of the probability of failure in the period $[t_i, t_{i+1})$ given survival to t_i in [3]. We will derive Eqs. 58.4 and 58.5 directly from the definition of the hazard function.

As can be found in any standard textbook on failure-time data analysis, we have the following definition and relationship equations for the hazard function. Assuming the time to failure T is a random variable which can take any value in the interval $[0, \infty)$, the hazard function of T is defined as

$$h(t) = \frac{f(t)}{1 - F(t)} = \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t \cdot (1 - F(t))}, \tag{58.6}$$

where $f(t)$ and $F(t)$ are the probability density function (pdf) and the cumulative distribution function (cdf) of T, respectively.

Since $f(t) = \mathrm{d}F(t)/\mathrm{d}t$, after some algebra, we get another form of the definition for the hazard function as

$$h(t) = -\frac{\mathrm{d}[\log(1 - F(t))]}{\mathrm{d}t} = \lim_{\Delta t \to 0} -\frac{\log(1 - F(t + \Delta t)) - \log(1 - F(t))}{\Delta t}. \tag{58.7}$$

By discretizing Eqs. 58.6 and 58.7 respectively, we get

$$\hat{h}(t) = \frac{F(t + \Delta t) - F(t)}{\Delta t \cdot (1 - F(t))}, \tag{58.8}$$

and

$$\hat{h}(t) = -\frac{\log(1 - F(t + \Delta t)) - \log(1 - F(t))}{\Delta t} = -\frac{1}{\Delta t} \log\left[\frac{1 - F(t + \Delta t)}{1 - F(t)}\right]. \tag{58.9}$$

Given our early defined notations N, $N(t_i)$, $\Delta t \equiv t_{i+1} - t_i$, and $h_i \equiv h(t_i)$, using the relative frequency as the estimator for $F(t_i)$, we have

$$F(t_i) \approx \frac{N - N(t_i)}{N} = 1 - \frac{N(t_i)}{N}. \tag{58.10}$$

By applying Eq. 58.10 to Eqs. 58.8 and 58.9 accordingly, Eqs. 58.4 and 58.5 fall out after some trivial and tedious algebra.

Up to this point, it is clear that both Eqs. 58.4 and 58.5 converge to the true values of h_i as Δt approaches zero. Note that this asymptotic property of convergence still holds after the introduction of Eq. 58.10 in the derivation process due to the law of large numbers [8]. We now investigate their theoretic properties when $\Delta t > 0$. First, let us rewrite Eq. 58.8 as

$$\hat{h}(t) = \frac{\int_t^{t+\Delta t} f(u)du}{\Delta t} \frac{1}{1 - F(t)}. \tag{58.11}$$

Equation 58.11 implies that Eq. 58.4 estimates the true hazard function values by dividing the average density $\left(\frac{\int_t^{t+\Delta t} f(u)du}{\Delta t}\right)$ over $1 - F(t)$, the system reliability value at time t. This implies that Eq. 58.4 will underestimate the true hazard function values if the true density function (pdf) is decreasing over the interval Δt and overestimate if the true pdf is increasing. Another way to show that Eq. 58.4 may underestimate the true h_i values is to consider Δt as a unit time interval, e.g., one hour, one day, or one year. Then, without the loss of generality, we have

$$\hat{h}_i = \frac{N(t_i) - N(t_i + \Delta t)}{N(t_i)} \equiv \widehat{h1}_i,$$

which implies that the empirical hazard values will never be greater than 1 per unit time.

Now let us rewrite Eq. 58.9 as

$$\hat{h}(t) = \frac{H(t + \Delta t) - H(t)}{\Delta t}, \tag{58.12}$$

where $H(t) = \int_0^t h(u)\mathrm{d}u = -\log(1 - F(t))$ is the cumulative hazard function. Equation 58.12 implies that Eq. 58.5 calculates the average values of the true hazard function. Therefore, we should expect to see that Eq. 58.5 will underestimate the true hazard function during its decreasing stage and overestimate it during the true hazard function's increasing stage. If the true hazard function is constant, Eq. 58.5 will give an unbiased estimation.

These theoretic properties of Eqs. 58.4 and 58.5 are verified by numeric calculation results as shown in Fig. 58.1 from which we can further examine to what extent these two empirical hazards formulas bias. In Fig. 58.1, plots on the left column are the densities of the specified distributions (i.e., exponential and Weibull); plots on the right column are the corresponding hazard function values calculated based on the specified parameters. For exponential distribution, the true hazards are calculated as $h(t)$ = rate; for Weibull distribution, the true hazards are $h(t)$ = (shape/scale) * (t/scale)$^{(\text{shape}-1)}$. The $\widehat{h1}_i$ values are calculated using Eq. 58.8; the $\widehat{h2}_i$ values are calculated using Eq. 58.12.

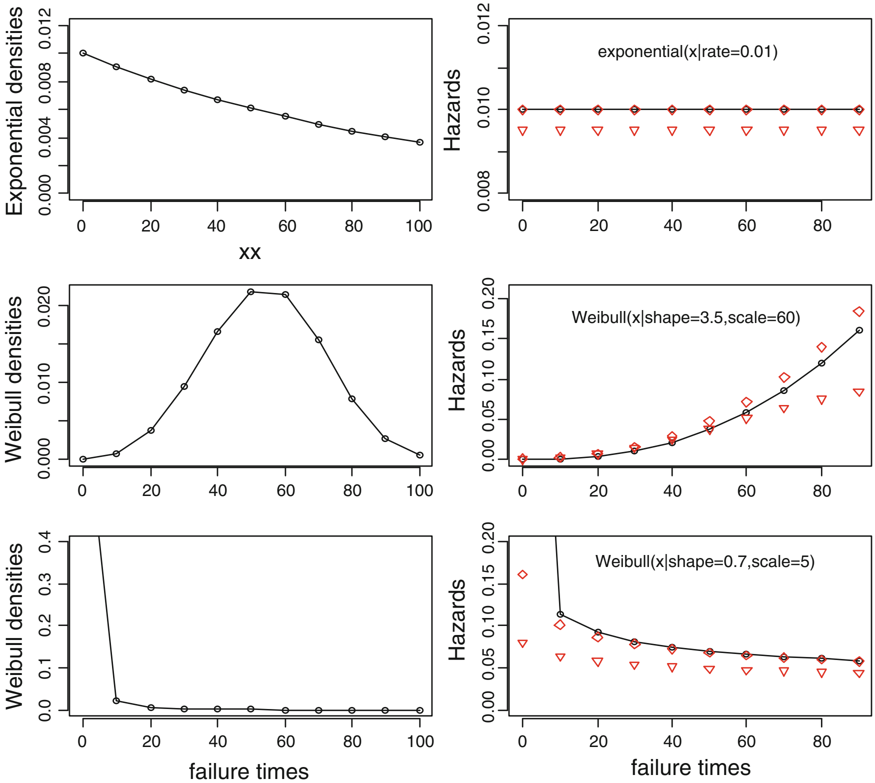


Fig. 58.1 Investigation of the bias effects of the empirical hazard function values calculated using $\widehat{h1}_i$ (the *top–down small triangle* points) and $\widehat{h2}_i$ (the *small diamond* points): *Circle* points are the true hazard function values connected by a *fine solid line*

Figure 58.1 shows that Eq. 58.5 gives much less biased estimation of the true hazard function than Eq. 58.4. In particular, Eq. 58.4 underestimates the true hazard function values in most cases and the underestimation is substantial. On the other hand, the bias created by Eq. 58.5 is minor or none if the fitted model is an exponential distribution. Note that the extremely large underestimation of the very first point in the bottom plots of Fig. 58.1 is because the true hazard value is positive infinity at $t = 0$ (in the case of a Weibull distribution with shape parameter less than one).

If we denote that $t + \Delta t \equiv t_2$ and $t \equiv t_1$ and hence $\Delta t = t_2 - t_1$, we realize that Eq. 58.1 and Eq. 58.12 are identical. This is how Eq. 58.5 related to AFR, but Eq. 58.4 does not have this direct connection.

As from Eq. 58.6, the hazard function $h(t)$, also referred to as hazard rate at time t, is defined as a conditional density function, i.e., the ratio of probability density $f(t)$ over the reliability $1 - F(t)$ (a probability), which is not as intuitive to interpret as the concept of failure rate used in data analysis. The direct connection of Eq. 58.5 with the AFR fills the mental gap between the probability theory and data analysis.

Theoretically, the difference between formulas (4) and (5) is significant. However, in data analysis practice, the numeric calculation results from both formulas can be very close. We now examine how different the estimation results can be between Eqs. 58.4 and 58.5. As a standard mathematical result [8] (pp. 251), it is known that if $|\mathrm{x}| \leq 2/3$, then

$$\log(1+x) = x - \frac{x^2}{2} + \theta(x),$$

where $|\theta(\mathrm{x})| \leq |\mathrm{x}|^3$. Therefore, it is straight forward to show that if $0 < x \leq 0.1$, then the relative difference between $-\log(1-x)$ and x (i.e., $[-\log(1-x) - x]/-\log(1-x)$) is less than 6 %.

In the next sections, using the simulation failure-time data samples, we will compare the estimation performances of Eqs. 58.4 and 58.5 to verify the theoretic results we have obtained above.

58.3 Verification of Theoretic Results Using Simulation Samples

A random sample of an exponential distribution of sample size n = 10,000 is generated with the parameter specification rate = 0.1 (using random seed 101 for exact repeatability of the analysis results). A second random sample of a Weibull distribution of sample size n = 10,000 is generated with the parameter specification: shape = 1.8 and scale = 30 (random seed = 101). Based on these two simulation random samples, the empirical hazard values $\widehat{h1}_i$ of Eq. 58.4 and $\widehat{h2}_i$ of Eq. 58.5 are calculated and compared with the true hazard function values to verify the theoretic results obtained from Sect. 58.2.

Figure 58.2 presents the simulation results of comparing the empirical hazard values $\widehat{h1}_i$ and $\widehat{h2}_i$ (in vertical bars) against the true hazard function values (in circles connected by a fine solid line) based on the exponential distribution random sample. In calculating $\widehat{h1}_i$ and $\widehat{h2}_i$, the most important setting is to specify the number of intervals over the full sample data range. The specification of the number of intervals is equivalent to specify the length of Δt. Therefore, we would expect to see that the larger the number of intervals, the better the approximation of the $\widehat{h1}_i$ and $\widehat{h2}_i$ values to the true hazard values. In Fig. 58.2, the empirical hazards in the top two panel plots are calculated using 20 intervals, and in the bottom two panel plots, the number of intervals is 50. As concluded in Sect. 58.2, we would expect to see $\widehat{h2}_i$ to be an unbiased estimator of the true hazard function,

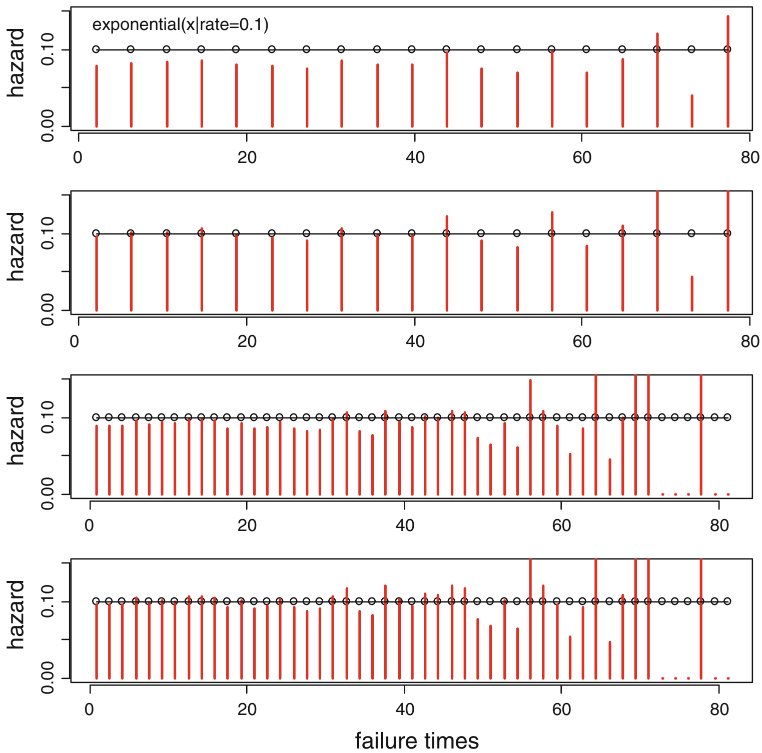


Fig. 58.2 Empirical hazard function values calculated using $\widehat{h1}_i$ (the *top* and third panel *plots*) and $\widehat{h2}_i$ (the second and *bottom* panel *plots*): *Circle* points are the true hazard function values connected by a fine solid line; vertical bars are the empirical hazard function values

and $\widehat{h1}_i$ will underestimate. The graph shows us that $\widehat{h2}_i$always performs better than $\widehat{h1}_i$, which is consistently underestimating the true hazards. The difference is much significant when the number of intervals is small. We also notice that it is $\widehat{h1}_i$, which is much more sensitive to the number of intervals specification, while $\widehat{h2}_i$'s estimation results are very robust (i.e., almost not affected by the change of the number of intervals specification).

With this particular exponential distribution sample, the 99 % quantile value is about 45 time units which only spread over less than 60 % of the full sample data range. Note that for both $\widehat{h1}_i$ and $\widehat{h2}_i$, the estimates fluctuate widely after the 99 % quantile point because of the sparseness of observations over the upper part of range interval. Actually, $\widehat{h2}_i$ will alway have an infinite large hazard value for the last interval because surely all components must fail in the end. On the other hand, $\widehat{h1}_i$ always equals to $1/\Delta t$ for the last interval. Therefore, empirical values of the very

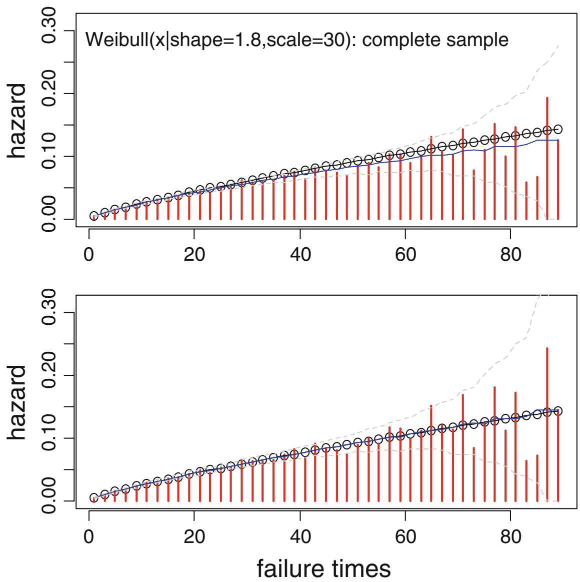


Fig. 58.3 Empirical hazard function values calculated using $\widehat{h1}_i$ (*top* panel plot) and $\widehat{h2}_i$ (*bottom* panel plot): *Circle* points are the true hazard function values connected by a *fine solid line*; vertical bars are the empirical hazard function values; the *thick blue line* is the medians of the empirical hazard function values calculated from 500 bootstrap samples (each of sample size $n* = 10{,}000$); the two *grey dashed lines* are the approximate 95 % confidence band

last interval should not be included. We, therefore, propose using only the estimates calculated from those sample observations which are up to 99 % quantile point.

Figure 58.3 examines the simulation results of comparing the empirical hazard values $\widehat{h1}_i$ (top panel) and $\widehat{h2}_i$ (bottom panel) against the true hazard function values based on a Weibull distribution random sample. Figure 58.3 follows the same drawing convention as in Fig. 58.2, i.e., the empirical hazard values $\widehat{h1}_i$ and $\widehat{h2}_i$ are represented in vertical bars against the true hazard function values (in circles connected by a fine solid line). The number of intervals is chosen to be 45, i.e., $\Delta t = 2$ time units. In addition, the approximate 95 % confidence bands for $\widehat{h1}_i$ and $\widehat{h2}_i$ values are constructed using the parametric bootstrap method [9]. Based on the Weibull distribution specification, 500 bootstrap samples (each of $n* = 10{,}000$) are generated and $\widehat{h1}_i$ and $\widehat{h2}_i$ are calculated for each of these bootstrap samples. The medians of empirical hazards are superimposed using a thick (in blue color) solid line with the dashed lines (in gray color) for the lower and upper limits, respectively.

Based on the theoretic results obtained in Sect. 58.2, we know that $\widehat{h1}_i$ will overestimate when failure times are small and underestimate when failure times become larger; $\widehat{h2}_i$ will overestimate slightly the true hazards. In Fig. 58.3, we see that the overestimation effect of $\widehat{h1}_i$ and the overestimation effect of $\widehat{h2}_i$ are visually unidentifiable. In contrast, the underestimation effect of $\widehat{h1}_i$ is substantial. In addition, in this particular Weibull distribution sample, the 99 % quantile point is at about 70 time units. In Fig. 58.3, the superimposed confidence bands show us visually how much the sampling variation can be over the upper part of the sample data range.

58.4 Conclusion and Future Work

Our research results have shown that $\widehat{h2}_i$ (defined in Eq. 58.5) is nothing but a finite approximation of AFR, whereas $\widehat{h1}_i$ (defined in Eq. 58.4) is a finite approximation of the instantaneous hazard rates. However, in their limiting forms, both $\widehat{h1}_i$ and $\widehat{h2}_i$ converge to the true hazard function h_i.

For data analysis purpose, a rule of thumb for calculating empirical hazard function of continuous-time failure data may be summarized as follows: If the maximum failure rate over the time interval periods of our concern is less than 0.1, then both $\widehat{h1}_i$ and $\widehat{h2}_i$ are good estimators of the true hazard function values. Most asset management reliability study cases should fall into this category. Otherwise, $\widehat{h2}_i$ should be used for calculating the empirical hazard function.

Note that both formulas are valid for randomly censored continuous-time failure data. In this paper, we have concentrated on discussing the calculation of the complete failure-time data using simulation samples. As for future research topics, more studies are needed to investigate the calculation of empirical hazards with different types of censored data and with real-life failure-time data sets.

The proposed rule of thumb should fill the gap between the probability theory and the data analysis practice in applications of the hazard function.

References

1. Ramakumar R (1993) Engineering reliability: fundamentals and applications. Prentice Hall, New Jersey
2. Meeker WQ, Escobar LA (1998) Statistical method for reliability data. Wiley, New Jersey
3. Davison AC (2003) Statistical models. Cambridge University Press, Cambridge
4. Venables WN, Ripley BD (1996) Modern applied statistics with S-Plus (corrected fourth printing). Springer, New York
5. Team RDC (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
6. Elsayed EA (1996) Reliability engineering. Addison Wesley Longman, Inc, Reading
7. Rai B, Singh N (2003) Hazard rate estimation from incomplete and unclean warranty data. Reliab Eng Syst Saf 81:79–92
8. Chung KL, AitSahlia F (2003) Elementary probability theory with stochastic process and an introduction to mathematical finance, 4th edn. Springer, New York
9. Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall/CRC, Boca Raton

Chapter 59
Simulation Study for Analysis of Online Dielectric Measurement Problems on Transformer Insulation System

Xu Yang, Gerard Ledwich and Shawn Nielsen

Abstract Determining the condition as well as the remaining life of an insulation system is essential for the reliable operation of large oil-filled power transformers. Frequency-domain spectroscopy (FDS) is one of the diagnostic techniques used to identify the dielectric status of a transformer. Currently, this technique can only be implemented on a de-energized transformer. This paper presents an initial investigation into a novel online monitoring method based on FDS dielectric measurements for transformers. The proposed technique specifically aims to address the real operational constraints of online testing. This is achieved by designing an online testing model extending the basic "extended Debye" linear dielectric model and taking unique noise issues only experienced during online measurements into account via simulations. Approaches to signal denoising and potential problems expected to be encountered during online measurements will also be discussed. Using fixed-frequency sinusoidal excitation waveforms will result in a long measurement times. The use of alternatives such as a chirp has been investigated using simulations. The results presented in the paper predict that reliable measurements should be possible during online testing.

59.1 Introduction

In the context of a liberalized energy market, condition-based maintenance [1] is necessary for power companies to cut costs of maintenance, which requires the condition monitoring of equipment in real time. This is only possible with online testing. Considering the transformer, its insulation aging status should be monitored in an active way since most of the failures of transformers occur due to

X. Yang (✉) · G. Ledwich · S. Nielsen
Queensland University of Technology, 2 George Street, Brisbane, QLD 4001, Australia
e-mail: xu.yang@student.qut.edu.au

W. B. Lee et al. (eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)*, Lecture Notes in Mechanical Engineering,
DOI: 10.1007/978-3-319-06966-1_59,

insulation deterioration [2]. Dielectric measurements have been recently widely used to determine aging status of this kind of insulation system. However, almost all the dielectric measurements used so far are applied in the offline mode [3], which requires transformer shutdown. The consequence of shutdown may make the system more vulnerable to a local failure of the power supply. Moreover, results from offline measurements will not reflect the real situation of insulation system due to the different thermal and electric stresses on the insulation between an operational transformer and an offline transformer. It is known that these stresses can have a big influence on the results [4].

The above disadvantages of offline testing can be overcome by online dielectric measurement. The online technique presented in this paper is based on Frequency-domain spectroscopy (FDS) and was evaluated via simulations. FDS can reflect the low-frequency responses (0.1 mHz – 1 Hz), with the advantage of inferring the moisture content in paper/pressboard [1, 5]. Other measurements such as polarization and depolarization current (PDC) [6] and return voltage measurement (RVM) [7] require a set of short-circuited and open-circuited test measurements. Those requirements make them unsuitable for online measurement.

Little research work has been carried out in the application of online dielectric measurement, especially in the application of FDS. Therefore, this paper presents a proposed method to prove that the online dielectric measurement can be implemented by carefully designing the testing circuit. It also includes a unique noise issue only experienced during online test and provides signal processing. In addition, this paper evaluates the effects of a chirp excitation as an alternative excitation signal with the purpose of quickening the measuring time. MATLAB simulation analysis is used to evaluate those approaches, which provides good preparation for online experiment in the future.

59.2 Debye Dielectric Model and Theory of FDS

59.2.1 Extended Debye Model

Figure 59.1 shows the extended Debye model used to interpret dielectric system responses [1]. This model consists of $R_i - C_i$ $(i = 1 \sim n)$ series circuit in a parallel arrangement. Those RC branches represent the non-instantaneous component of polarization of different insulation parts and corresponding time constant given by $\tau_i = R_i C_i$. C_0 and R_0 represent geometric capacitance and insulation resistance of insulation system [8].

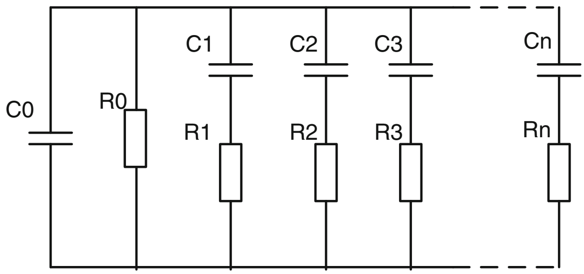


Fig. 59.1 A schematic illustration of the data acquisition setup in the experiment

59.2.2 FDS

FDS is a measurement of the dissipation factor and complex capacitance of insulation system within frequency band from 0.1 mHz to 10 kHz [1]. This can be done by injecting sinusoidal voltage with different frequencies onto insulation system and measuring the current response. Dielectric response of moisture content in cellulosic insulation part corresponds to low-frequency band 0.1 mHz – 1 Hz. Therefore, for the online test, this paper focuses on this critical frequency band.

When a dielectric medium is subjected to a sinusoidal voltage $U(\omega)$, there is a resultant current $I(\omega)$ flowing through it [9]. The current responses can be written as in Eq. 59.1:

$$\begin{aligned} I(\omega) &= i\omega C_0\left\{\varepsilon'(\omega) - i\left[\varepsilon''(\omega) + \frac{\sigma_0}{\varepsilon_0\omega}\right]\right\}U(\omega) \\ &= i\omega[C'(\omega) - iC''(\omega)]U(\omega) \\ &= i\omega C^*(\omega)U(\omega) \end{aligned} \tag{59.1}$$

where $\varepsilon'(\omega)$ is its real part which represents stored energy in dielectric medium; $\varepsilon''(\omega)$ is its imaginary part which represents dissipation of energy within dielectric medium. $C^*(\omega)$ is complex capacitance of the insulation; the real part $C'(\omega)$ represents the capacitive component of the insulation material, and the imaginary part $C''(\omega)$ represents the losses.

The dissipation factor can be expressed as [1]:

$$\tan\delta = \frac{C''(\omega)}{C'(\omega)} \tag{59.2}$$

where $C^*(\omega) = C_0 + \frac{1}{i\omega R_0} + \sum_{i=1}^{n} \frac{C_i}{1+i\omega R_i C_i}$ can be derived from its equivalent circuit in Fig. 59.1.

59.3 Online Measurement Model

In view of online FDS, the signal (illustrated in Fig. 59.2 as Vs.) is designed to be applied through neutral line on the primary side of the transformer under test, while the corresponding response current is measured through neutral line on the secondary side. During this procedure, noise in the neutral connection should be taken into account. In order to precisely measure the low-frequency responses, it is necessary to design an analog filter to divide the total responses into low-frequency part and high-frequency part and then collect low-frequency results. Furthermore, since the neutral noise mainly comes from unbalanced loads on the system, which is typically 50 Hz, the filter needs to block 50 Hz and associated harmonics from entering the measurements. The filter was designed as in Fig. 59.3. The Bode diagram of this filter is illustrated in Fig. 59.4.

In Fig. 59.3, low-frequency path consists of inductor in series with a resistor and passes low-frequency current while removes current with frequencies higher than 1 Hz. In comparison, high-frequency path contains a capacitor in parallel with damping resistor and passes current with frequencies higher than 1 Hz, while low-frequency signal is removed.

59.4 Analysis of Online Measurement Issues

59.4.1 Noise Analysis

In order to extract the current responses within 0.1 mHz – 1 Hz frequency band, it is critical to analyze the noise characteristics in the neutral connections. In the neutral line, the noise will be made up of power frequency current variation due to unbalanced load variation in the system as well as low-frequency noise induced into the power system [10]. Under normal operating conditions, the unbalanced load variations will depend on customer switching, which happen randomly. In this case, the switching pattern in load side can be seen as 1/f noise. Although the switch action is random, it is modulating load at 50 Hz since the loads are working at this frequency. Therefore, the random switch load current in neutral line can be modeled as below:

$$\mathrm{Inv} = \left(1 + \frac{1}{f}\mathrm{noise}\right) \times \mathrm{In}_{50\mathrm{Hz}} \tag{59.3}$$

where Inv represents the variation of neutral current and $\mathrm{In}_{50\mathrm{Hz}}$ is the neutral current without variation. The equation can be seen as 1/f noise amplitude modulation. And its power spectrum is illustrated in Fig. 59.5, and it can be seen that the noise energy is very small within 0.1 mHz – 1 Hz band with most of the power centralized around 50 Hz. This implies that the filtering strategy will be easier to implement if the noise is low enough in the frequency band of interest.

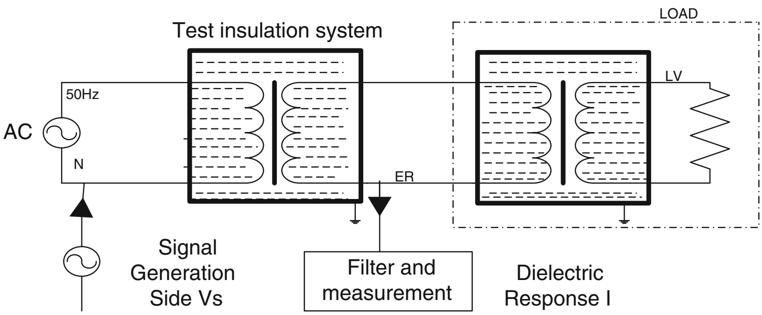


Fig. 59.2 Layout of online dielectric measurement

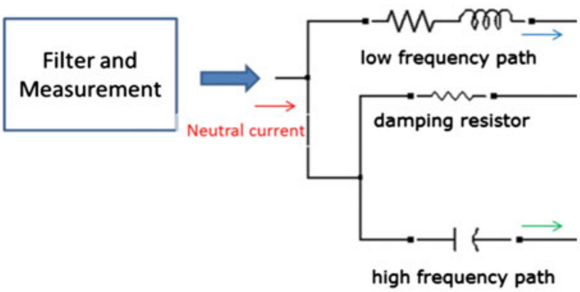


Fig. 59.3 Analog filter

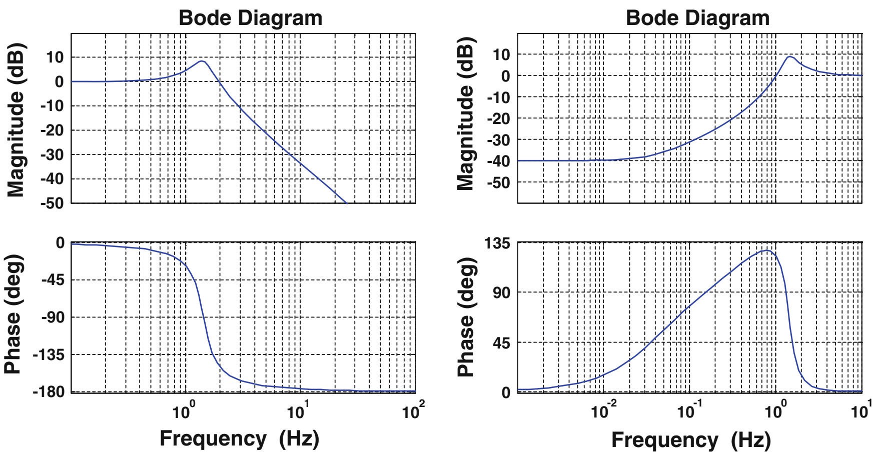


Fig. 59.4 Bode diagrams low-frequency path (*left*) and high-frequency path (*right*)

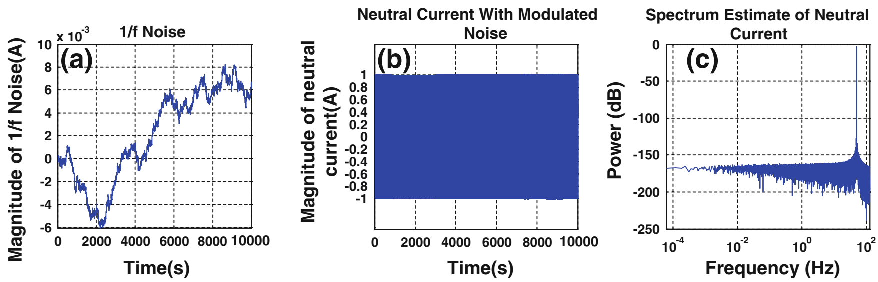


Fig. 59.5 **a** 1/f noise; **b** normalized neutral current with noise; and **c** spectrum estimation of neutral current with modulated noise

59.4.2 Chirp Excitation

For low-frequency band, the time consumed during the online procedure will be long and should be reduced. Usually, the FDS measurement is repeated at several important frequencies using discrete sine waves and the whole measurement could take over 3 h [5]. In this paper, a chirp is used as an alternative excitation in order to reduce the testing time. A low-frequency band, 0.1 mHz – 1 Hz, is used. In the case of the chirp signal, more energy should be put in the low-frequency range, where the information about moisture is strongest. This can be achieved by using a logarithm swept chirp signal implemented in MATLAB.

59.5 Simulation Results

The proposed online method was simulated in Simulink, as shown in Fig. 59.6. The input chirp signal had an amplitude of 200 V, and the dielectric Debye model parameters were found in the literature [11].

The chirp voltage (Hamming windowed) was applied on the primary side of the transformer. The dielectric current is then passed through the insulation model. On the secondary side, the neutral current with modulated noise was added onto the neutral line. The noisy current then was forced to pass through filter part whose structure is shown in Fig. 59.3. The input voltage and output current without neutral noise waveforms are shown in Fig. 59.7. It is noticed that the magnitude of pure dielectric current is very small and the time for the test is only 5000 s (less than 2 h). This small response current is easily swamped by noise.

Figure 59.8 below shows the filtering procedure in the secondary side. Before application of the analog filter, the signal was totally covered by neutral noise in Fig. 59.8a. Its spectrum estimate is illustrated in Fig. 59.10. And the signal-to-noise ratio (SNR) at this stage was −138 dB. In Fig. 59.8b, there was still small amount of 50 Hz noise, but the SNR was reduced to −18 dB. After the analog

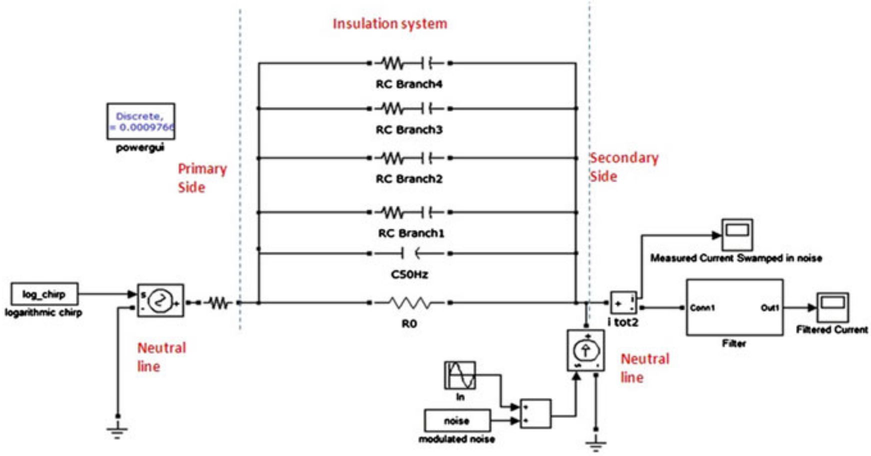


Fig. 59.6 Simulation layout in Simulink for online dielectric measurement

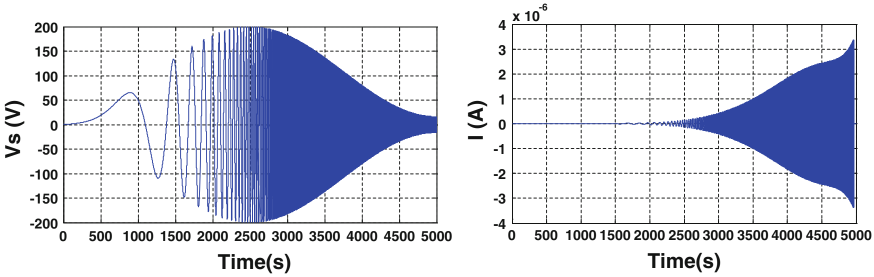


Fig. 59.7 Log-chirp input voltage (*left*) and its pure current output (*right*) with frequency range (10^{-4}–10^{1}Hz)

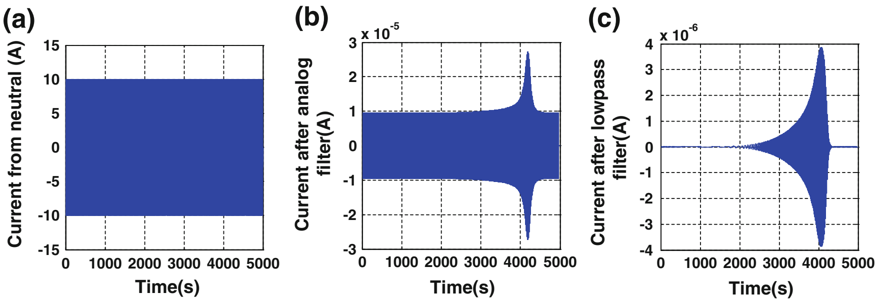


Fig. 59.8 Filtered procedure: **a** unfiltered neutral current; **b** filtered current by analog filter; and **c** further filtered current by Butterworth low-pass filter

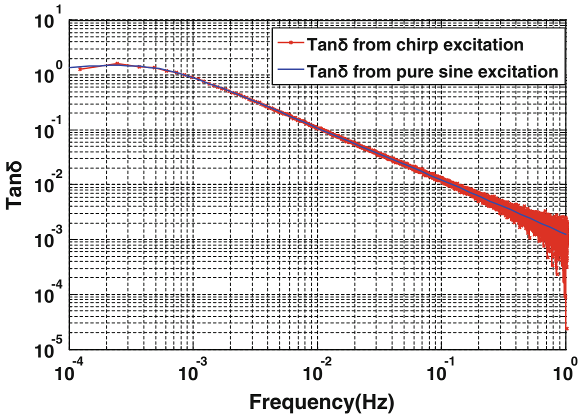


Fig. 59.9 Dissipation factor of the insulation system

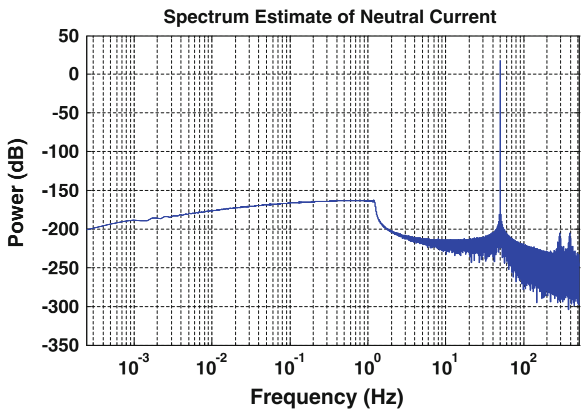


Fig. 59.10 Power spectrum of neutral current

filter, a low-pass Butterworth filter was used to remove all the noise with frequencies above 1 Hz. Figure 59.8c shows that the dielectric current response can be attained after those filter methods. The SNR is now better at 66 dB. Compared with Fig. 59.7b, the frequency between 1 Hz and 10 Hz was filtered. As shown in Fig. 59.9, dissipation factor obtained by using log-chirp excitation was compared to the results when the excitation signal was a discrete sine wave.

59.6 Conclusion and Future Work

An online testing method based on FDS has been presented in this paper for large oil-filled power transformers. The online test has been proved possible by injecting the test signal and extracting its responses through neutral line of a transformer. The method was tested by means of simulation. A log-chirp signal was used to reduce the testing time of traditional FDS test. The low-frequency dielectric response can be enhanced by using filters. The filtering method was applied based on the analysis of neutral noise, which was modulated at 50 Hz.

Experiments should be completed to validate the simulation results in this paper. Further filter strategies are needed when noise is much larger than the dielectric current in low frequency. Online dielectric measurement is a difficult task considering the denoising issue and online dielectric response interpretation part. Future work will focus on the experiment testing.

Acknowledgments This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centres Programme. The authors gratefully acknowledge the financial support provided by the CRC.

References

1. Zaengl WS (2003) Dielectric spectroscopy in time and frequency domain for HV power equipment. I, theoretical considerations. IEEE Electr Insul Mag 19:5–19
2. Xiang Z, Gockenbach E (2008) Asset-management of transformers based on condition monitoring and standard diagnosis. IEEE Electr Insul Mag 24:26–40
3. Saha TK (2003) Review of modern diagnostic techniques for assessing insulation condition in aged transformers. IEEE Trans Dielectr Electr Insul 10:903–917
4. Fofana I, Hemmatjou H, Meghnefi F (2010) On the frequency domain dielectric response of oil-paper insulation at low temperatures. IEEE Trans Dielectr Electr Insul 17:799–807
5. Koch M, Raetzke S, Krueger M (2010) Moisture diagnostics of power transformers by a fast and reliable dielectric response method. In: Conference record of the 2010 IEEE international symposium on electrical insulation (ISEI), San Diego, CA, U.S.A 1–5 June 2010
6. Muhamad NA, Phung BT, Blackburn TR (2009) Polarization and depolarization current (PDC) tests on biodegradable and mineral transformer oils at different moisture levels. In: Australasian universities power engineering conference (AUPEC), Adelaide, Australia, 1–6 September 2009
7. Dwivedi CK, Daigavane MB (2009) Evaluation of moisture content in paper-oil of aged power transformer using RVM. In: 2nd International conference on emerging trends in engineering and technology (ICETET), Nagpur, India, pp 470–475 December 2009
8. Raju GG (2003) Dielectric loss and relaxation—II. In Dielectrics in electric fields. (ed), CRC Press
9. Jonscher AK (1983) Dielectric relaxation in solids. Chelsea Dielectrics Press, London
10. Masoum M, Moses PS, Masoum AM (2008) Derating of asymmetric three-phase transformers serving unbalanced nonlinear loads. IEEE Trans Power Delivery 23:2033–2041
11. Roongroj C, Pattanadech N, Praisuwanna P (2008) Equivalent circuit approximation of transformer insulation by using PDC measurement. In: 5th International conference on electrical engineering/electronics, computer, telecommunications and information technology, Krabi, Thailand, pp 925–928 May 2008