Lecture Notes in Mechanical Engineering

Joseph Mathew · C. W. Lim · Lin Ma
Don Sands · Michael E. Cholette
Pietro Borghesani *Editors*

Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies

Proceedings of the 12th World Congress on Engineering Asset Management and the 13th International Conference on Vibration Engineering and Technology of Machinery

Springer

Lecture Notes in Mechanical Engineering

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. 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

To submit a proposal or request further information, please contact: Dr. Leontina Di Cecco Leontina.dicecco@springer.com or Li Shen Li.shen@springer.com.

Please check the Springer Tracts in Mechanical Engineering at http://www.springer.com/series/11693 if you are interested in monographs, textbooks or edited books. To submit a proposal, please contact Leontina.dicecco@springer.com and Li.shen@springer.com.

More information about this series at http://www.springer.com/series/11236

Joseph Mathew · C. W. Lim
Lin Ma · Don Sands · Michael E. Cholette
Pietro Borghesani
Editors

Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies

Proceedings of the 12th World Congress on Engineering Asset Management and the 13th International Conference on Vibration Engineering and Technology of Machinery

Springer

Editors
Joseph Mathew
The Asset Institute
Queensland University of Technology
Brisbane, QLD, Australia

C. W. Lim
Department of Architecture and Civil Engineering
City University of Hong Kong
Hong Kong SAR, China

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

Don Sands
Synengco Pty Ltd
Highgate Hill, QLD, Australia

Michael E. Cholette
School of Chemistry, Physics, Mechanical Engineering
Queensland University of Technology
Brisbane, QLD, Australia

Pietro Borghesani
School of Mechanical and Manufacturing Engineering
UNSW Sydney
Sydney, NSW, Australia

ISSN 2195-4356 ISSN 2195-4364 (electronic)
Lecture Notes in Mechanical Engineering
ISBN 978-3-319-95710-4 ISBN 978-3-319-95711-1 (eBook)
https://doi.org/10.1007/978-3-319-95711-1

Library of Congress Control Number: 2018951213

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

WCEAM VETOMAC 2017—from 2 to 4 August 2017, Brisbane, Australia

The 12th World Congress on Engineering Asset Management was held jointly with the 13th Vibration Engineering and Technology of Machinery on 2–4 August 2017 at the Brisbane Convention and Exhibition Centre. It was organised by the Asset Institute and hosted annually by the International Society of Engineering Asset Management (ISEAM).

The event shared two themes: "Asset Intelligence through Integration and Interoperability: From Research to Industry" and "Creative and Novel Research for Contemporary Vibration Engineering Technologies" and attracted 244 delegates from 28 countries.

The offerings were rich and diverse. The Opening Ceremony hosted the following speakers:

- Adjunct Professor Joseph Mathew, Asset Institute CEO and Congress Chair, Queensland University of Technology, Australia.
- Professor Joe Amadi-Echendu, ISEAM Chair, University of Pretoria, South Africa.
- Professor C. W. Lim, City University of Hong Kong, China.
- Professor John Bell on behalf of Prof. Arun Sharma, Deputy Vice Chancellor, Queensland University of Technology.
- The Opening Address by the Lord Mayor Graham Quirk, Lord Mayor of Brisbane, Australia.
- The Opening Address by the Platinum Sponsor, Mr. Darren Covington of Mainpac who spoke on, "Delivering Operational Effectiveness in Asset-Intensive Industries through Asset Intelligence".

The event hosted seven keynotes, a dinner speaker, 14 special sessions on several application and discipline streams, three minicourses, three workshops and three-panel discussions.

Our keynotes this year were:

1. Alan Johnston, MIMOSA and Standards Leadership Council, USA; Jess B. Kozman, Professional Petroleum Data Management (PPDM) Association, Singapore. “Intelligent Integration and Interoperability of Critical Infrastructure and Assets”.
2. Professor Kerry Brown, Edith Cowan University, Australia, “Engineering Asset Management: Understanding the Management Element”.
3. Associate Professor Marco Macchi, Politecnico di Milano, Italy, “The 4th Industry Revolution: Reflecting on the Opportunities, Barriers and Risks for Asset Management”.
4. Professor Romuald Rzadkowski, Airforce Institute of Technology, Poland, “Asset Management Through Life Estimation”.
5. Professor Tang Loon Ching, National University of Singapore, “Systems Resilience: a Unifying Framework and Associated Measures”.
6. Professor Klaus Blache, University of Tennessee, USA, “Technologies and Asset Management: What is really going on in Industry”.
7. Professor J. P. Liyanage, Cluster for Industrial Asset Management, University of Stavanger, Norway, “High Risk Assets Under Uncertain Conditions: Strategic Imperatives and New Initiatives Towards Defensive Solutions in a Rapidly Changing Environment”.

The dinner speaker was Dr. Paul Simshauser, Director General, Department of Energy and Water Supply, Queensland Government, who spoke on “Energy industry—challenges ahead”.

One hundred and fifty management and technical presentations reported on outputs of research and development activities as well as the application of knowledge to industry and government in the practical aspects of the various themes that comprises Engineering Asset Management and Vibration Engineering. These papers and presentations were made available for the personal reference of delegates only via an eProceedings Dropbox repository at the Congress, and not for distribution.

Full papers were further peer-reviewed by two members of the Scientific Committee and this e-book contains the final selection of these papers.

The Congress was augmented with a series of three interactive workshops.

The first of these workshops was entitled “**The Long Future Sustainability for Asset Managers**” in which David Hood and Guy Lane encouraged participants to think about and discuss their individual mindsets and roles as asset managers in influencing and creating a sustainable future for our planet. At the workshop, David remarked that the principle of sustainability is often confined to the financial sense rather than encompassing a more holistic and long-term sense by including environmental sustainability.

Jess Kozman, a Regional Representative of the Professional Petroleum Data Management Association and independent data management practitioner, lead a workshop on the “**Evaluation of data management maturity in relation to**

engineering assets". The workshop took participants through a short exercise in which a range of survey questions was used to establish the data management maturity of the participant's organisations. The workshop was attended by participants from a broad range of industries and backgrounds. Some practical ideas in relation to the application of maturity surveys were given by Jess.

"**Fostering the development of Engineering Asset Management Programs at Higher Educational Institutions via Recognition**" was the subject of the final workshop. A highly interactive discussion was led by Prof. Joe Amadi-Echendu, Chair of ISEAM. Participants discussed the merits of recognition rather than accreditation of programs, the development and custodianship of the Asset Management Body of Knowledge and made comparisons between Asset Management and the Project Management Institute and their development path.

Three minicourses were conducted on the final day of the Congress. Alan Johnston from MIMOSA led the training session entitled "**The open industrial interoperability ecosystem, a supplier-neutral digital ecosystem, enabling critical infrastructure and industrial asset management**". Professor Romuald Rzadkowski from the Polish Academy of Sciences conducted the course "**Life estimation and exact time of failure of last stage steam turbine blades**". Dr. Carla Boehl from Curtin University held a training session on "**Mine Autonomous Haul System: Assessing the Impact in Asset Management**". Each of the minicourses was well attended.

The generous support of our Sponsors is acknowledged. They were our Platinum Sponsor, Mainpac, our Silver Sponsors, Brisbane City Council, Schneider Electric and the Fredon Group. Bronze Sponsors: QUT, Synengco, NMEMS Japan, K2Fly and Redeye, as well as our Exhibitors, Asset Finda, SAVTek, Springer and Request Direct.

Also acknowledged were our partners who assisted us in organising this event and providing some of our key speakers, i.e. MIMOSA, PPDM, the Vibration Institute of India, the IFAC A-MEST Working Group, the 2017 Global Business Challenge and NCCARF, Australia (Fig. 1).

This year ISEAM gave out two awards for the Best Paper Award, one for the management category and the other for the technical category and acknowledged the runner-ups as well. These were:

Management

Winner: "Analysing an Industrial Safety Process through Process Mining: A Case Study"

Authors: Anastasiia Pika, Arthur H. M. ter Hofstede, Robert K. Perrons, (QUT) and Georg Grossmann, Markus Stumptner, (University of South Australia), and Jim Cooley (Origin Energy).

Runner-up: "Semiparametric valuation of heterogenous assets".

Authors: Roar Adland and Sebastian Köhn, Norwegian School of Economics (NHH), Norway.

Fig. 1 Associate Professor Rob Perrons receiving the award from Adjunct Professor Mathew with Prof. Markus Stumptner on left

Fig. 2 Mr. Norihiko Ogura receiving the award from the Congress Chair

Technical

Winner: "Efficient Evaluation of Internal Concrete Damage of Steel Plate-bonded RC Slabs".

Authors: Norihiko Ogura, CORE Institute of Technology Corp, Hitoshi Yatsumoto, Hanshin Expressway Company Ltd. and Takahiro Nishida and Tomoki Shiotani, Kyoto University (Fig. 2).

Runner-up: Joint Optimization of Preventive Maintenance and Spare Parts Logistics for Multi-echelon Geographically Dispersed System.

Authors: Keren Wang and Dragan Djurdjanovic, University of Texas at Austin, USA.

The 2017 ISEAM Lifetime Achievement Award was presented to Emeritus Prof. Robert B Randall, University of New South Wales, Australia.

ISEAM's Lifetime Achievement Award recognises and promotes individuals who have made a significant contribution to research, application and practice of a discipline or in engineering asset management over a continued period of time (Fig. 3).

The Congress Chair was especially honoured by ISEAM at the Congress with a Lifetime Achievement Award. He is the 5th recipient of this award since the formation of the society in 2006.

In addition, ISEAM further honoured the Congress Chair with a Pioneer Award in recognition of his foundational contributions to ISEAM as its Founder and Chair

Fig. 3 Professor Randall receiving the Lifetime Achievement Award from the ISEAM Chair

Fig. 4 Professor Mathew receiving the Lifetime Achievement Award from the ISEAM Chair

of the Board in its first 11 years. The awards were presented by the ISEAM Chair, Prof. Joe Amadi-Echendu, University of Pretoria, South Africa.

The event was a resounding success judging from the feedback of the delegates and the program may be used in future WCEAMs' as a template for transnational, trans-sectoral and transdisciplinary knowledge sharing events in asset management going into the future (Fig. 4).

We are grateful for all the voluntary assistance provided by members of the International Scientific Committee, the Organisation Committee and the Conference Secretariat who are all acknowledged in the following sections.

Congress Chairs

Congress Chair: Adjunct Professor Joseph Mathew, Asset Institute, Queensland University of Technology, Australia

Co-chairs: Dr. Trudy Curtis, Professional Petroleum Data Managers (PPDM) Association, Canada
Alan Johnston, MIMOSA, USA
Professor C. W. Lim, City University of Hong Kong, Hong Kong
Associate Professor Marco Macchi, Politecnico di Milano, Italy

International Scientific Committee

Co-chairs: Dr. Michael E. Cholette, Queensland University of Technology, Australia
Dr. Pietro Borghesani, Queensland University of Technology, Australia
Adjunct Professor Joseph Mathew, Queensland University of Technology, Australia
Assistant Professor Siu-kai Lai, Hong Kong, China
Associate Professor Zhongxiao Peng, Australia
Dr. Allen Tam, Australia
Dr. Carla Boehl, Australia
Dr. Christos Emmanouilidis, UK
Dr. Diaswati Mardiasmo, Australia
Dr. Davide Crivelli, UK
Dr. Fahim Tonmoy, Australia
Dr. Hongkun Li, China
Dr. J. S. Rao, India
Dr. Joel Adams, UK
Dr. Paul Shantapriyan, Australia
Dr. Rob Schoenmaker, The Netherlands
Dr. Rifat Shahriar, Malaysia
Dr. Ruizi Wang, Australia
Dr. Samuel Paterson, Australia
Dr. Steven Pudney, Australia
Dr. Ype Wijnia, The Netherlands
Dr. Zuzana Dimitrovová, Portugal
Emeritus Professor Robert Randall, Australia
Mr. Don Sands, Synengco Pty Ltd, Australia
Ms. Helena Kortelainen, Finland
Professor Amy J. C. Trappey, Taiwan, China
Professor Andy Koronios, Australia
Professor Andy Tan, Malaysia
Professor Belle Upadhyaya, USA
Professor C. W. Lim, Hong Kong, China
Professor David Mba, UK
Professor Dimitris Kiritsis, Switzerland
Professor Fucai Li, China
Professor Jayantha P. Liyanage, Norway
Professor Joe Amadi-Echendu, South Africa
Professor Kerry Brown, Australia
Professor Lin Ma, Queensland University of Technology, Australia
Professor Marco Macchi, Italy
Professor Melinda Hodkiewicz, Australia
Professor Ming J. Zuo, Canada
Professor Shunming Li, China
Professor Tian Ran Lin, China

WCEAM VETOMAC 2017 Organising Committee

Chair: Don Sands, Synengco Pty Ltd, Australia
Co-chair: Professor Lin Ma, Queensland University of Technology, Australia
Dr. Jeremy Novak, Centaur Institute, Australia
Dr. Paul Shantapriyan, University of Tasmania, Australia
Dr. Steve Pudney, Queensland University of Technology, Australia—Industry sectors
Mr. Jess Kozman, PPDM, Singapore—Industry sectors
Mr. Lucas Skoufa, Queensland University of Technology, Australia
Mr. Mario Bojilov, MBS Conventions, Australia
Ms. Betty Goh, Asset Institute, Australia
Ms. Elise Sommer, Professional Petroleum Data Managers (PPDM) Association, Canada
Professor Andy Koronios, University of South Australia, Australia
Professor Andy Tan, Universiti Tunku Abdul Rahman, Malaysia
Professor Ashantha Goonetilleke, Queensland University of Technology, Australia
Professor Belle Upadhyaya, University of Tennessee, USA
Professor Kerry Brown, Edith Cowan University, Australia
Professor Robin Drogemuller, Queensland University of Technology, Australia
Professor Tian Ran Lin, Qingdao University of Technology, China

Congress Secretariat

Lauren Kerr and Rebecca Wood, QUT Conferences
Queensland University of Technology

Brisbane, Australia	Joseph Mathew
Hong Kong SAR, China	C. W. Lim
Brisbane, Australia	Lin Ma
Highgate Hill, Australia	Don Sands
Brisbane, Australia	Michael E. Cholette
Brisbane, Australia	Pietro Borghesani

Contents

Predicting the Remaining Life of Timber Bridges

T. Abbott, N. Gamage, S. Setunge and Weena Lokuge

Abstract This paper documents the current state of knowledge relating to the deterioration of timber bridges in Australia. The aim of this research, was to comprehend the present state of knowledge regarding maintenance of timber bridges and address any gap in knowledge. This involved: identifying key defects in timber, investigating the inspection methods utilised to detect these faults and finding the preventive measures used to mitigate bridge deterioration. Enclosed are figures which demonstrate how simple industry practices and procedures implemented by each states' governing authority can reduce these impacts and concludes with an empirical model for predicting the remaining lifespan of a bridge.

1 Introduction

Of the roughly 40,000 bridges in Australia, 27,000 of them are constructed of timber. Most of these are over 50 years old and in a weathered condition [4]. It is obvious (Fig. 1) that the timber bridge stock is depleting irrespective of the efforts in the rehabilitation processes [1]. From government documents and first hand correspondence with engineers in the industry it has been discovered that a lot of these bridges have been replaced by steel elements. The majority of remaining

T. Abbott · N. Gamage · S. Setunge
School of Civil, Environmental and Chemical Engineering, RMIT University, Melbourne, Australia
e-mail: tano_abb@hotmail.com

N. Gamage
e-mail: nirdosha.gamage@gmail.com

S. Setunge
e-mail: sujeeva.setunge@rmit.edu.au

W. Lokuge (✉)
School of Civil Engineering and Surveying, University of Southern Queensland, Darling Heights, Australia
e-mail: weena.lokuge@usq.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_1

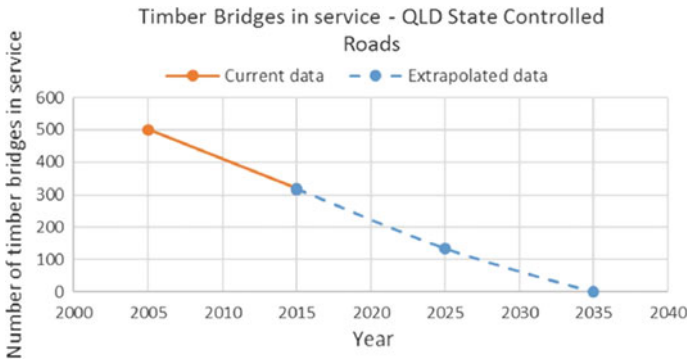


Fig. 1 Timber bridge stock in Queensland

timber bridges in Victoria, are of the girder deck system type and located on forest walk trails and tourist roads or are local roads which are controlled by municipalities [5]. There are a number of considerations which affect how a bridge will perform throughout its service life. The design load, environmental factors, type of timber and size of members used all contribute to which known defects each bridge element will be most susceptible to and the best method of detection [3]. With the knowledge contained in this report and previous history a matrix has been developed highlighting these key aspects. Inspectors can use this for their reference to check typical symptoms of decay and take appropriate actions.

The major defects that all Australian timber bridges are susceptible to are splitting and rot due to fungal attack. If the bridge is used as a water crossing, marine borers and flood are other possible factors. Termite infestation is not as predominate in southern states, although reports do suggest that in warmer states such as Queensland, this may be more of threat to timber structures. Inspectors have reported that it is more likely to see timber shrink and be pulled off their fasteners than it is to see the iron nails corrode, thus this form of deterioration is not very significant. The stringers of a bridge are subjected to substantial load distributed from the superstructure and may undergo excessive deflection over time in a phenomenon known as creep. Inspection techniques are divided into destructive and non-destructive. The most common and reliable destructive tool involves drilling into a timber member to grasp its interior condition. Currently in Australia, state governing bodies are responsible for the upkeep of their timber bridges intrust their skilled engineers to visually inspect the structures and decide which course of action to take.

2 Industry Practice

The following is a paraphrased excerpt from a conversation with two structural engineers who have decades of experience in inspecting timber bridges. Their knowledge outlines the current Victorian practice and highlights their conclusions

about deterioration they gathered from practical inspections in the field. The governing authority employed a number of cost-effective methods of preventing damage to timber elements including a range of epoxies and paints to protect the outside surface from moisture and debris build up which could lead to decay. The implementation of anti-split bolts and washers were installed on the piles to reduce the impact deep checks would have to the column. Petroleum jelly was also used as a water proof barrier. These measures are implemented on a case by case manner as recommended by the inspecting engineer, without any formal standard to follow.

There are a number of major factors to look for when attempting to estimate the remaining lifespan of a bridge. It is imperative to inspect the condition of every elements' surface, checking for moisture and debris accumulation which could lead to decay and risk in case of fire. It is important to check for splits and cracks in corbels and cross beams. There is no cause for alarm if sapwood is seen rotting as this always rots first and doesn't contribute to section capacity. It is humidity, soil presence and moisture govern how a timber will deteriorate. Timber bridges were abundant in Victoria during the 1940s due to availability of wood. The mindset back then was to aim for fifty years life from the structure through inexpensive, low level maintenance.

2.1 Treatment Methods

Treatment methods are often referred to as maintenance because of the existing proactive understanding in the industry. The ethic is that it is more beneficial to increase the lifespan of a bridge and its components through regular inspection and maintenance than to replace decrepit members on a need to basis. Frequent upkeep and inspection can report on the derogation rate and suggest recommendation for planned works in the future. The method of replacing members once they have decayed beyond repair is a costly exercise and by the time the deterioration is detected the operational effectiveness of the structure can be compromised, resulting in serviceability failure and a risk to safety [2].

The maintenance methods can be divided into three branches which are distinguished by the cost, timeframe and level of work performed. The simplest and cheapest form of treatment is Routine Maintenance/Preventative Treatment. This consists of mainly minor reactive works which are anticipated and allocated for in the budget and planned on a short term basis, usually about two weeks or less [2]. This commonly includes controlling factors which provide favourable conditions for decay such as moisture content. Periodic Maintenance/Early Remedial Treatments are carried out at regular intervals of longer than one year. These are designed to fix problems associated with early stage defects, such as rot. It is undertaken on a proactive rather than reactive basis [2]. Examples include baiting systems to deter termites, the installation of pile jackets and more rigorous application of sealants. Specific works/major maintenance occurs when decay is so advanced that the member simply does not have the structural strength to support

the loads acting on it. Expensive replacement type improvement maintenance or rehabilitation maintenance is used when a member is so decayed regular treatments will not repair its structural integrity. These include one-off repairs, refurbishment and upgrade works to retain the bridge as close as possible to its original condition.

3 Defect Treatment Protocol

Deterioration is a cause and effect process. Each type requires certain factors to manifest. These can be detected with different methods and categorised into routine maintenance mitigation, periodic maintenance and major works or rehabilitation maintenance. The following (Fig. 2) is an amended defect treatment matrix adopted from the Queensland Department of Transport and Main Road's Timber Bridge Maintenance Manual [6].

Matrix shown in Fig. 2 is a tabulated illustration condensing all the information found in the maintenance manuals. This figure details the type of defects, how they are identified and how they are mitigated through maintenance techniques. It provides a snapshot to inspectors at what to look for, how to detect and how to treat various forms of deterioration. The matrix does not give an indication to how long a bridge may last until it is unsafe, it only gives protocols to follow at the time of inspection. Therefore current practice shows that there is no concise methodical approach to quickly ascertain the required information about how to mitigate or predict the rate of timber bridge deterioration, which is reinforced by other researchers [4]. Equipped with this knowledge, a procedure collaborating all the necessary data has been developed to address the gap in current knowledge. It is anticipated that these tables can be used by inspectors to empirically rate and score the level of deterioration of a bridge.

Taking a closer look at stringers the usual causes for deterioration are: pipe rot, splitting and termite infiltration. Depending on the design and member dimensions, excessive deflection may also prematurely weaken the member. Figure 3 shows a worst case scenario for the prognosis of a stringer in a poorly fabricated and executed maintenance regime. Using the stringer member as an example, the flowchart shown above displays the major contributors to the element's structural demise as well as indicating the cause of each defect. The exterior boxes describe the factors present which cause the four most common types of decay to the stringer element. Interestingly, termite infestation occurs when rot breaks down the wood fibers allowing the organisms to enter the dark cavities, which occurs when timber shrinks releasing water attracting fungal spores to grow and start decaying the wood. Therefore, for termites to be present, the wood must first have been weakened by fungal rot to improve the conditions for termites to thrive. This shows a more appealing outcome for the condition of a bridge. With an effective mitigation

Defect	Caused by	Inspection Technique	Treatment Category	Treatment Action
Severe Splitting	Internal tensile stresses built up due to drying	Visual- check ends, occurs along the grain, be weary of water presence	Specific works/ replacement	Replace timber girder
Significant splitting			Periodic/ early remedial	Install anti-split bolts for mitigation
Fasteners corroded	Rust (rain, oxygen)	Visual, half-cell voltage	Specific works/ replacement	Replace or install bolts
Girder to cap beam bolts loose due to shrinkage	Uneven drying of timber member	Visual, physically testing strength	Routine/ preventative	Tighten existing bolts
Termite infestation	Existing decay, no light, humid moist atmosphere	Acoustic, coring,	Periodic/ early-remedial	Eradicate nest
Localised termite presence		Acoustic, visual, chipping	Routine/preventative	Inject poison into member
Crushing at end of span	Excessive shear force	Visual deformation present	Periodic/ early-remedial	Add supporting member
Light to medium Fire damage	Flammable environment	Visual	Periodic/ early- remedial	Apply chemical preservative
Fungal decay (up to 30% diameter loss at mid-span and 20% at ends)	Supply of food, moist atmosphere, suitable temperature and oxygen	Resistograph, coring, visual (if decay is external)	Routine/ preventative	Apply chemical preservative to slow decay rate
Fungal rot (up to 70% diameter loss at mid-span and 50% at ends)		Stress wave timing, chipping, coring, acoustic	Specific Works/ major maintenance or replacement	Replace timber stringer or install steel girder
Any defect causing more than 70% loss of section regardless of cause	Failure to implement an effective maintenance program	Measuring residual strength using more accurate and intricate methods (coring and stress wave timing)	Specific Works/major maintenance or replacement	Seek advice from structures division to decide whether to demolish or attempt to restore

Fig. 2 Defect treatment matrix

protocol in place, this bridge will outlast its above counterpart. One noteworthy remark is that this flowchart only shows the major causes affecting the deterioration rate. Other less common factors such as flooding or fire may contribute to the ultimate design life.

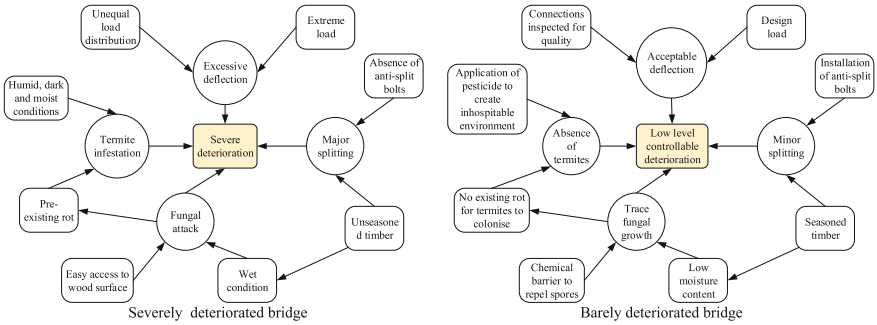


Fig. 3 Flowchart of bridge, highlighting the causes and factors contributing to respective states

4 Bridge Deterioration Prediction Model

After researching the current state of knowledge and delving into the information provided in the states' maintenance manuals, it became apparent that there was no documented way to predict the remaining service life of a bridge based on environmental factors which manifested the identifiable defects within it. Within this section is a prototype of an empirical model to produce an educated estimate of the amount of time, in years, until the overall condition of a timber bridge surveyed reduces to condition state 4.

This model depends on the type of decay present to predict the serviceability of the bridge. Therefore the first step is identifying the modes of deterioration. The five main constituents are fungus, splitting, termites, corrosion of fasteners and marine borers, which affect the three elements differently. Below are clauses to be considered when applying this model.

- Clause (a) Termites are only considered for bridges in warm climates. For the purpose of this model, "warm climates" include Queensland, Western Australia and the Northern Territory regions.
- Clause (b) Marine Borers are only considered for bridges where there is significant water presence that has prolonged contact with timber surfaces, such as rivers and creeks.
- Clause (c) Each form of decay affects each bridge element to various degrees of significance. Bridge Stringers will be susceptible to: Fungal attack, termite attack (if located in region defined in clause a) and splitting. Bridge Piles will be susceptible to: Fungal attack, termite attack (if located in region defined in clause a), splitting and marine borers (if situation satisfies clause b). Bridge Corbels will be susceptible to: Fungal attack, termite attack (if located in region defined in clause a), splitting and fastener corrosion.

- Clause (d) any defect which causes more than 50% loss of section (beyond the parameters of Condition state 4) is said to be unsafe as it has failed and requires immediate replacement of the member.

4.1 Limitations

This prototype focuses on three critical structural members that make up the sub-structure of the bridge. These are the stringers, corbels and piles. For this model to be implemented accurately, the inspector has to comment on the presence of each type of decay prone to a particular member. That is, they should not conclude a condition state based purely on one mode of deterioration, but should endeavour to observe other forms applicable to the element which may accelerate decay. The values assigned to each condition state are purely estimated with no laboratory testing to authenticate the figures delegated. It is also possible for one element to have a different condition state to another element of the same type. When this happens, it is up to the inspector's discretion to implement a value which best fits the description for the entire number of members of that type. For simplicity of the model and due to resource restrictions, it is assumed that degradation of a bridge is linear in nature.

The prediction model is based on the observations made during inspections. However it can accommodate quantitative observations or measurements as well using structural health monitoring devices. Further development of the model can be done based on differentiation between the wood species.

4.2 Model

The following tabulates all of the steps required to classify the condition of a bridge via its empirical score (S) and therefore compute its remaining lifespan. This model encompasses all information which has been gathered from this research and will need to be referred to acquire figures. Table 1 demonstrated the allocated distribution for each decay type for each element. The element in question and number of

Table 1 Deterioration weighted splits

Decay type	Weight as a percentage			
Fungus	60	45	50	40
Splitting	40	30	30	30
Termite	–	10	20	15
Marine Borer	–	15	–	–
Corrosion	–	–	–	15
Element	Stringer/pile	Pile	Stringer	Corbel

applicable methods of degeneration determine how much each type of decay contributes to the total amount of deterioration; see procedure clauses a, b and c.

As shown in Table 2, points are assigned for each condition state. Using the condition states listed earlier in this report the inspector can match the description of the bridge element to the most suitable condition state. This process will have to be repeated for each element (stringer, corbel and pile). See also procedure clause d. There may not be four decay types for every element based on environment.

The total weight for each element thus found is inputted into the cell under the respective element. The average is then calculated. Before the final score is given the entire bridge can be subjected to environmental factors which may adjust its overall score (Table 3).

Sum the values of both columns that are applicable based on inspector's general observations. Table 4 provides a numerical and descriptive indication of a bridge's integrity. This result can be compared to future inspections to summarise a trend and determine the rate of deterioration.

Following equation can be used to determine the time the bridge has until it reaches condition state 4. The equation is based on the scores for two extreme ends of the spectrum in Table 4. As shown in Table 2, a bridge in condition four still receives points and allocated a score of 3 in Eq. 1.

$$\upsilon = \frac{S-3}{10-S} * T \tag{1}$$

Table 2 Points assigned for each condition state

Condition state	1	2	3	4
Points assigned	9	6	4	2

Table 3 General observations modification adjustment

Discolouration	−0.1	Fresh, new coating	+0.1
Loss of fill in abutments	−0.1	Deck well maintained	+0.1
Deck cracking	−0.2	Sapwood still intact	+0.2
Undersize members	−0.3	No visible deflection under load	+0.2
Visible decay in other bridge members	−0.3	No debris on any bridge component	+0.3

Table 4 Timber bridge classes

Score (S)	0–3	3.1–5	5.1–7.5	7.6–10
Classification	Poor	Fair	Good	Excellent

υ = Remaining time in years until condition state 4 is reached; S = Score calculated from Table 6; and T = Age of bridge when inspection was conducted, in years. Inspectors would then refer to the defect treatment matrix for maintenance options and recommended mitigation techniques to preserve existing infrastructure to prolong the service life of the bridge.

4.3 Demonstration of the Model

20 year old two span simple beam wooden bridge constructed from seasoned F grade timber is selected to demonstrate the developed model. Bridge deck consists of timber planks installed transversely with bitumen sealed to support loads up to 20 t. Mid-span piles are submerged up to half the length of the pile in a river. Considerable discolouration and decay (~20%) was observed throughout all elements of the superstructure and substructure. Deck is still within serviceability limit state and deck and top of headstocks were littered with debris. Pile is in a more serious decayed state due to multiple types of decay present.

Tables 5 and 6 demonstrates the calculations for the final score and using Eq. 1 the remaining life of this bridge can be calculated as 13.57 years [(5.83–3)/(10–5.83)]. Treatment matrix such as the one shown in Fig. 2 can be used to determine the maintenance process for this bridge.

Table 5 Element designation and point matrix

Element	Decay	Description	Condition	Points	Weight
Stringer	Fungus	~20% pipe rot	2	6	3
	Splitting	Minor-medium splitting	2	5	1.5
	Termite	Minor presence	1	9	1.8
Total				20	6.3
Corbel	Fungus	~20% pipe rot	2	6	2.4
	Splitting	Minor splitting	2	6	1.8
	Termite	Minor presence	1	9	1.35
	Corrosion	Fasteners slightly loose	2	6	0.9
Total				27	6.45
Pile	Fungus	~30% pipe rot	3	5	2.25
	Splitting	Moderate splitting	3	5	1.5
	Termite	None	1	10	1
	Marine	Significant activity	3	4	0.6
Total				24	5.35

Table 6 Computation of final score (S)

Stringer weight	Corbel weight	Pile weight	Average weight
6.3	6.45	5.35	6.03
General observations modifications	No visible deflection +0.2 Visible decay −0.3 Discolorations −0.1	Final score (S) classification	5.83 Good

5 Conclusions

This research has been conducted through a highly theoretical and analytical approach. From the beginning the purpose of this research has been to report on the characteristics of timber as it degrades throughout its lifespan and what current authorities are doing to mitigate damage and maintain their infrastructure. The bridge deterioration prediction model has been completed which was the key deliverable in this paper. This model should satisfactorily address the gap in knowledge regarding this subject matter. Further research is necessary to access and validate the accuracy of the model. With more data and analysis of how deterioration develops, better time estimates can be recorded and the model can be recalibrated to improve its application throughout Australia's timber bridges. The best way to validate this research is to apply it to previously documented bridges or to conduct a study to observe and record in-service bridges' deterioration rates to check if they follow the predicted relationship.

Acknowledgements Authors would like to thank Mr. David Kempton and Mr. Ronald Hawken of VicRoads for their time and sharing their knowledge.

References

1. Cohen J (2016) Review of timber bridge maintenance practices on Queensland state controlled road networks. Honours thesis, University of Southern Queensland
2. Main Roads (2014) Detailed visual bridge inspection guidelines for timber bridges, Western Australia
3. Ranjith S (2010) Tools for diagnosis and prediction of deterioration of timber bridges in Australia. Masters thesis, RMIT University
4. Ranjith S, Setunge S, Gravina R, Venkatesan S (2013) Deterioration prediction of timber bridge elements using the Markov Chain. J Perform Constructed Facil 27:319–325
5. Road Structures Inspection Manual (2014) VicRoads, Melbourne, Australia
6. Timber Bridge Inspection Manual (2015) Department of Transport and Main Roads, Queensland Government

A Bibliographic Review of Trends in the Application of 'Criticality' Towards the Management of Engineered Assets

Joel Adams, Ajith Parlikad and Joe Amadi-Echendu

Abstract Increasing budgetary constraints have raised the hiatus for allocation of funding and prioritisation of investments to ensure that long established and new assets are in the condition to provide uninterrupted services towards progressive economic and social activities. Whereas a key challenge remains how to allocate resources to adequately maintain infrastructure and equipment, however, both traditional and conventional practices indicate that decisions to refurbish, replace, renovate, or upgrade infrastructure and/or equipment tend to be based on negativistic perceptions of criticality from the viewpoint of risk. For instance, failure modes, failure effects, and criticality analyses is well established and continues to be applied to resolve reliability and safety requirements for infrastructure and equipment. Based on a bibliographic review, this paper discusses trends in meaning, techniques and usage of the term 'criticality' in the management of engineered assets that constitute the built environment. In advocating the value doctrine for asset management, the paper proposes a positivistic application of criticality towards prioritisation of decisions to invest in the maintenance of infrastructure and equipment.

Keywords Criticality definition · Criticality analysis · Criticality application

J. Adams · A. Parlikad
Institute for Manufacturing, University of Cambridge, 17 Charles Babbage Road, CB3 0FS Cambridge, UK
e-mail: ja579@cam.ac.uk

A. Parlikad
e-mail: aknp2@cam.ac.uk

J. Amadi-Echendu (✉)
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

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_2

1 Introduction

In the context of this paper, we define criticality in terms of relative importance of an item to a decision maker. Something is critical if it is most important in a situation. It is in this context that the term has become colloquially adopted, especially since the 1940s as part of the design methodology referred to as failure modes, effects and criticality analysis (FMECA) (see MIL-P-1629, 1949). The use of FMECA was primarily to ensure the safety and reliability of products in a wide range of industries, especially in the aerospace, automotive, biomedical and nuclear sectors [18, 57]. This further spread to civil aviation and automotive sectors in the same period. Since then, the FMECA has become a common approach in reliability theory and practice (refer to IEC 60812; BS 5760–5; USM [67]). These standards describe techniques and approaches for criticality analysis as an essential function in the design and development of engineered components, equipment, and systems.

The rest of the paper is organised as follows: Section 2 presents the results of an extensive review of literature published between 1950 and 2016 using 'criticality' as the search criterion. Empirical data derived from the literature review is presented in Sect. 3, while Sect. 4 includes some conclusions.

2 Review of Literature

The source used for our study was academic journal articles published between 1950 and 2016. The search initially focused on articles indexed in Scopus and Google Scholar but also extended to citation search (both forward and backward) on the primary sources. This indirect search pointed to articles related to "criticality analysis" in terms of definition, technique and usage of criticality. This implies that articles merely describing the criticality analysis process have not been included.

2.1 Trend in the Definition of Criticality

The meaning of criticality has changed over the years. Sometimes even within a single organisation, different individuals may have different interpretations of equipment criticality [59]. There are many unanswered questions about what asset criticality means.

EN 13306 [8] defines criticality as "numeric index of the severity of failure or fault combined with the probability or frequency of its occurrence". According to USM Standard [67], criticality is a relative measure of the consequences of a failure mode and its frequency of occurrence. Some authors agree with this view of criticality in terms of risk of failure, (e.g., Cooper [23]; Wilson and Johnson [69]; Elperin and

Dubi [26]; McKinney and Iverson [41]; Kim et al. [35]; Bishop et al. [15]; Bertolini and Bevilacqua [14]; Benjamin et al. [13]).

Other methods for computing criticality focus on the cost consequences of failure (e.g., Spencer [60], Kendrick [34], Pappas and Pendleton [50], Gilchrist [30], Gajpal et al. [28], Alvi and Labib [5], Al-Najjar and Alsyouf [4], Walski et al. [68], Baruah and Vestal [12], Theoharidou et al. [65], Ahmadi et al. [2], Canto-Perello et al. [21]). BS 3811 (1984) and Norsok [48] define criticality analysis as a quantitative evaluation of events and faults and the ranking of these in order of the seriousness of their consequences. Pschierer-Barnfather et al. [54] see asset criticality as a comparative measure of consequences.

Extant literature suggests that asset criticality implies failure risk. If so, can asset criticality also be viewed from a value-based perspective? In attempting to answer this question, we developed a classification of criticality in terms of risk-based, cost-based or value-based.

2.2 *Trend in Techniques for Computing Criticality*

To understand the trend in the different techniques for calculating criticality, we used framework proposed in Liu et al. [37] for classifying the methods that have been identified in the literature. In this review, we divide the methods used in the literature into five main categories which are multi-criteria decision making (MCDM), artificial intelligence (AI), simulation (S), integrated approaches (IA), other approaches (OA). The five categories, each with some the related techniques and references, are reported in Table 1.

Table 1 Classification of evaluation methods for criticality analysis

Categories	Techniques	References
MCDM	AHP/ANP	Alvi and Labib [5], Gajpal et al. [28], Molenaers et al. [44], Stoll et al. [62], Goossens et al. [31], Nyström and Söderholm [49]
	Hazop	Bishop et al. [15]
	RBC	Theoharidou et al. [65]
	RPN	Shin et al. [58]
	Decision Tree	Dong et al. [25]
IA	Delphi-Color Coded-AHP	Canto-Perello et al. [21]
	FTA-SHA	Ye and Kelly [72]
	AHP-TOPSIS-VIKOR	Ahmadi et al. [2], Al-Najjar and Alsyouf [4]
	Fuzzy-MCDM	Arunraj and Maiti [7], Bertolini and Bevilacqua [14], Langkumaran and Kumanan [33]
	AHP-GP	Marriott et al. [38]
	Fuzzy-AHP-TOPSIS	
	PAM-FMEA	

(continued)

Table 1 (continued)

Categories	Techniques	References
AI	Fuzzy Logic	Braglia et al. [19], Mechefske and Wang [42], Feng and Chung [27], Pelaez and Bowles [52], Bowles and Peláez [18], Xu et al. [71]
S	EPS	Walski et al. [68]
	Monte Carlo	Liu and Frangopol [36], Neves and Frangopol [47], Borgovini et al. [16], Banaei-Kashani and Shahabi [10]
	MCNP	McKinney and Iverson [41]
OA	Input–Output Model	Benjamin et al. [13]
	FMEA	Cooper [23], Tsakatikas et al. [66], Abdul-Nour et al. [1]
	EUAC	Flores-Colen and de Brito (2010)
	CBC	Moore and Starr [45]

2.3 *Trend in the Usage of Criticality*

To understand the trend in the different usage of criticality analysis, we divide the methods found in the literature into six main categories as shown in Table 2.

Table 2 Classification of the uses of criticality analysis

Uses of criticality	References
Criticality analysis for prioritizing failure modes during design	Bishop et al. [15], Ye and Kelly [72], [23], Staat et al. [61], McGinnis [40], Moskowitz [46], Babb [9], Ahsen and von Ahsen [3]
Criticality analysis for selection/ planning of maintenance	Alvi and Labib [5], Benjamin et al. [13], Canto-Perello et al. [21], Abdul-Nour et al. [1], Ahmadi et al. [2], Al-Najjar and Alsyouf [4], Arunraj and Maiti [7], Bertolini and Bevilacqua [14], Mechefske and Wang [42], Dong et al. [25], Ilangkumaran and Kumanan [33], Nyström and Söderholm [49]
Criticality analysis for spare parts management	Gajpal et al. [28], Molenaers et al. [44], Stoll et al. [62], Tsakatikas et al. [66], Botter and Fortuin [17], Braglia et al. [20], Dekker et al. [24], Gelders and Van Looy [29], Huiskonen [32], Partovi and Anarajan [51], Porras and Dekker [53], Ramanathan [55], Syntetos et al. [63], Teunter et al. [64], Tsakatikas et al. [66], Zhou and Fan [73]

(continued)

Table 2 (continued)

Uses of criticality	References
Criticality analysis for prioritization of maintenance tasks	Baruah and Vestal [12], Marriott et al. [38], Moore and Starr [45], Nyström and Söderholm [49], Ming Tan [43], Cerrada et al. [22]
Criticality analysis for prioritizing decisions to acquire new equipment	Theoharidou et al. [65], Marriott et al. [38], Pschierer-barnfather et al. [54], Barnfather et al. [11], Alvi and Labib [5], Andersen et al. [6]
Criticality analysis for reliability improvement	Braglia et al. [19], Walski et al. [68], Shin et al. [58], McKinney and Iverson [41], Ramirez-Marquez and Coit [56], Pappas and Pendleton [50], Xu et al. [70], Maucec et al. [39], Ćatić et al. [74]

3 Findings and Observations

From a review of 94 articles published between 1950 and 2016, we make the following findings and observations. The bar graph in Fig. 1a shows that, from 1950 to 1985, the dominant definition of criticality was risk-based followed by the cost perspective. However, from 1998 onwards, a new perspective emerges of criticality in terms of the impact of decisions on values of an organization. As illustrated in the bar graph of Fig. 1b, multi-criteria decision making techniques are most frequently applied to criticality analysis. Figure 1c also indicates that criticality is primarily applied to prioritize failure modes, formulate strategy for maintenance interventions and reliability improvements, and to improve the management of equipment spares.

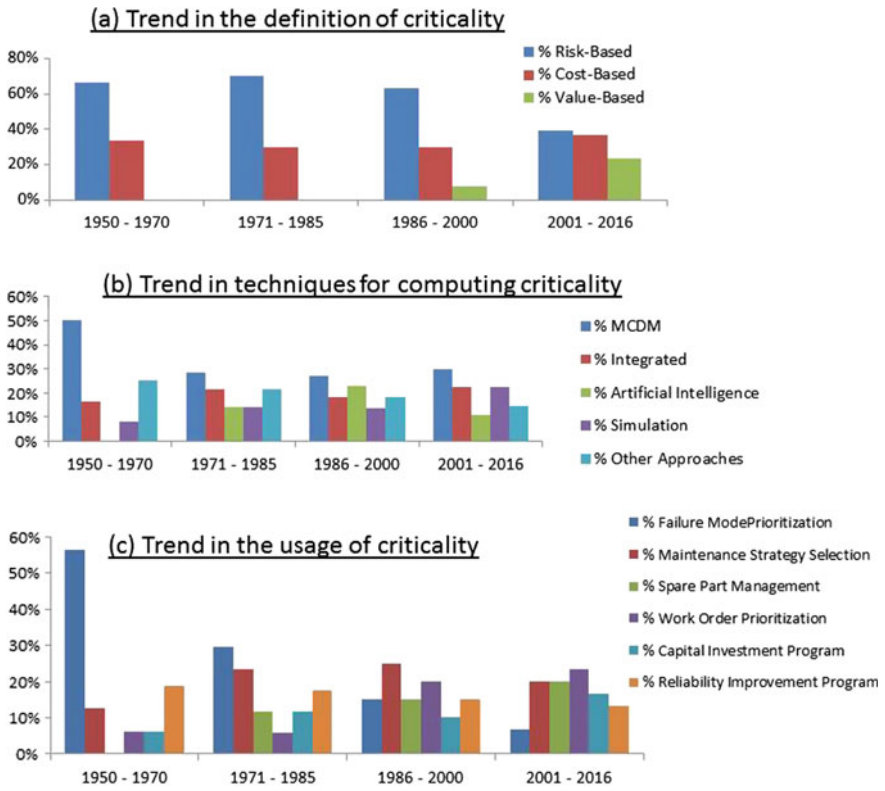


Fig. 1 Trends in the definition, computation and usage of criticality

4 Conclusions

This paper is based on a literature review from 1950 to 2016 on the meaning, usage and techniques for computing criticality. First, it was found that criticality analysis, which was first developed to meet obvious reliability and safety requirements, has been always been seen from a negativistic point of view. Second, although several approaches are proposed for computing criticality, however, the reviewed literature suggests that multi-criteria decision methods predominate. Third, the most common use of criticality rating was for prioritizing failure modes, developing maintenance strategy and reliability improvement programs. Remarkably, in practice, the use of criticality for the maintenance purpose is still at a nascent stage in many organisations. Other increasingly becoming popular applications of criticality are for spare parts management, work order prioritization and capital investment decisions. This suggests a positivist view of criticality embodied in the value doctrine for effective management of engineered assets that constitute our built environment.

Acknowledgements The research work was performed within the context of Sustain Owner ("Sustainable Design and Management of Industrial Assets through Total Value and Cost of Ownership"), a project sponsored by the EU Framework Programme Horizon 2020, MSCA-RISE-2014: Marie Skłodowska-Curie Research and Innovation Staff Exchange (RISE) (grant agreement number 645733—Sustain-Owner—H2020-MSCA-RISE-2014). We thank the Petroleum Technology Development Fund (PTDF) Nigeria, for sponsorship of the 1st author's PhD research.

References

1. Abdul-Nour G, Beaudoin H, Ouellet P, Rochette R, Lambert S (1998) A reliability based maintenance policy; a case study. Comput Ind Eng 35(3–4):591–594. https://doi.org/10.1016/S0360-8352(98)00166-1
2. Ahmadi A, Gupta S, Karim R, Kumar U (2010) Selection of maintenance strategy for aircraft systems using multi-criteria decision making methodologies. Int J Reliab Qual Saf Eng 17(3):223–243. https://doi.org/10.1142/S0218539310003779
3. Ahsen A Von, von Ahsen A (2008) Cost-oriented failure mode and effects analysis. Int J Qual Reliab Manage 25(5):466–476. https://doi.org/10.1108/02656710810873871
4. Al-Najjar B, Alsyouf I (2003) Selecting the most efficient maintenance approach using fuzzy multiple criteria decision making. Int J Prod Econ 84(1):85–100. https://doi.org/10.1016/S0925-5273(02)00380-8
5. Alvi AU, Labib AW (2001) Selecting next-generation manufacturing paradigms—an analytic hierarchy process based criticality analysis. Proc Inst Mech Eng: Part B: J Eng Manuf 215(12):1773–1786. https://doi.org/10.1243/0954405011519493
6. Andersen GR, Chouinard LE, Hover WH, Cox CW (2001) Risk indexing tool to assist in prioritizing improvements to Embankment Dam inventories. J Geotech Geoenvironmental Eng 127(4):325–334. https://doi.org/10.1061/(ASCE)1090-0241(2001)127:4(325)
7. Arunraj NS, Maiti J (2010) Risk-based maintenance policy selection using AHP and goal programming. Saf Sci 48(2):238–247. https://doi.org/10.1016/j.ssci.2009.09.005
8. BSI (2001) Maintenance terminology. In: Bs En 13306:2001. ISBN: 978 0 580 64184 8
9. Babb AH (1974) Failure mode effects and criticality analysis (FMECA) and planned maintenance applied to TXE 4 electronic switching system. In Proceedings of the international switching symposium, p 8
10. Banaei-Kashani F, Shahabi C (2003) Criticality-based analysis and design of unstructured peer-to-peer networks as "Complex systems"'. In: CCGrid 2003. 3rd IEEE/ACM international symposium on cluster computing and the grid, 2003. Proceedings, pp 351–358. https://doi.org/10.1109/ccgrid.2003.1199387
11. Barnfather P, Hughes D, Wells R (2014) Condition-based risk management of physical assets within the electrical power sector
12. Baruah S, Vestal S (2008) Schedulability analysis of sporadic tasks with multiple criticality specifications. Real-Time Systems, 2008. ECRTS'08. Available at http://ieeexplore.ieee.org/abstract/document/4573111/. Accessed 17 Feb 2017
13. Benjamin MFD, Tan RR, Razon LF (2015) A methodology for criticality analysis in integrated energy systems. Clean Technol Environ Policy 17(4):935–946. https://doi.org/10.1007/s10098-014-0846-0
14. Bertolini M, Bevilacqua M (2006) A combined goal programming—AHP approach to maintenance selection problem. Reliab Eng Syst Saf 91(7):839–848. https://doi.org/10.1016/j.ress.2005.08.006
15. Bishop P, Bloomfield R, Clement T, Guerra S (2003) Software criticality analysis of COTS/SOUP. In Reliability Engineering and System Safety, 291–301. https://doi.org/10.1016/s0951-8320(03)00093-0

16. Borgovini R, Pemberton S, Rossi M (1993) Failure mode, effects and criticality analysis (FMECA). reliability analysis center (B). Available at http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA278508. Accessed 7 Feb 2017
17. Botter R, Fortuin L (2000) Stocking strategy for service parts—a case study. Int J Oper Prod Manag 20(*5–6*), 20:656–674
18. Bowles JB, Peláez CE (1995) Fuzzy logic prioritization of failures in a system failure mode, effects and criticality analysis. Reliab Eng Syst Saf 50(2):203–213. https://doi.org/10.1016/0951-8320(95)00068-D
19. Braglia M, Frosolini M, Montanari R (2003) Fuzzy criticality assessment model for failure modes and effects analysis. Int J Qual Reliab Manag 20(4):503–524. https://doi.org/10.1108/02656710310468687
20. Braglia M, Grassi A, Montanari R (2004) Multi-attribute classification method for spare parts inventory management. J Qual Maintenance Eng 10
21. Canto-Perello J, Curiel-Esparza J, Calvo V (2013) Criticality and threat analysis on utility tunnels for planning security policies of utilities in urban underground space. Expert Syst Appl 40(11):4707–4714. https://doi.org/10.1016/j.eswa.2013.02.031
22. Cerrada M, Cardillo J, Aguilar J, Faneite R (2007) Agents-based design for fault management systems in industrial processes. Comput Ind 58(4):313–328. https://doi.org/10.1016/j.compind.2006.07.008
23. Cooper J (1971) LEAM failure mode effect and criticality analysis. Available at https://repository.hou.usra.edu/handle/20.500.11753/380. Accessed 6 Feb 2017
24. Dekker R, Kleijn M, de Rooij P (1998) A spare parts stocking policy based on equipment criticality. Int J Prod Econ, 69–77
25. Dong YL, Gu YJ, Dong XF (2008) Selection of optimum maintenance strategy for power plant equipment based on evidential reasoning and FMEA. In: 2008 IEEE international conference on industrial engineering and engineering management. IEEE, pp 862–866. https://doi.org/10.1109/ieem.2008.4737992
26. Elperin T, Dubi A (1985) On the Markov chain analysis of source iteration Monte Carlo procedures for criticality problems: I. Nuclear science and engineering. Available at http://www.ans.org/pubs/journals/nse/a_17128. Accessed 6 Feb 2017
27. Feng C, Chung C (2013) Assessing the risks of airport airside through the fuzzy logic-based failure modes, effect, and criticality analysis. Mathematical problems in engineering. Available at https://www.hindawi.com/journals/mpe/2013/239523/abs/. Accessed 17 Feb 2017
28. Gajpal PP, Ganesh LS, Rajendran C (1994) Criticality analysis of spare parts using the analytic hierarchy process. Int J Prod Econ 35(1–3):293–297. https://doi.org/10.1016/0925-5273(94)90095-7
29. Gelders LF, Van Looy PM (1978) An inventory policy for slow and fast movers in a petrochemical plant: a case study. J Oper Res Soc 29(9):867–874
30. Gilchrist W (1993) Modelling failure modes and effects analysis. Int J Qual Reliab Manag Emerald 10(5). https://doi.org/10.1108/02656719310040105
31. Goossens, AJM, Basten RJI, Dongen LAM Van (2014) Exploring the use of the Analytic Hierarchy Process for maintenance policy selection. Safety, reliability and risk analysis: beyond the horizon—proceedings of the European safety and reliability conference, ESREL 2013. Shers, pp 1027–1032. Available at http://www.scopus.com/inward/record.url?eid=2-s2.0-84900007725&partnerID=tZOtx3y1
32. Huiskonen J (2001) Maintenance spare parts logistics: special characteristics and strategic choices. Int J Prod Econ 71:125–133
33. Ilangkumaran M, Kumanan S (2009) Selection of maintenance policy for textile industry using hybrid multi-criteria decision making approach. J Manuf Technol Manage 20(7):1009–1022. https://doi.org/10.1108/17410380910984258
34. Kendrick EDJ Jr (1966) ELK river reactor operations analysis program. Spent fuel shipping cask criticality analysis: task 617. U.S. Atomic Energy Commission. https://doi.org/10.2172/4532922

35. Kim T, Singh B, Sung T, Park J, Lee Y (1996) Failure mode, effect and criticality analysis (FMECA) on mechanical subsystems of diesel generator at NPP. Available at https://inis.iaea.org/search/search.aspx?orig_q=RN:28018603. Accessed 6 Feb 2017
36. Liu M, Frangopol DM (2004) Optimal bridge maintenance planning based on probabilistic performance prediction. Eng Struct 26(7):991–1002. https://doi.org/10.1016/j.engstruct.2004.03.003
37. Liu H-C, Liu L, Liu N (2013) Risk evaluation approaches in failure mode and effects analysis: a literature review. Expert systems with applications, pp 828–838. https://doi.org/10.1016/j.eswa.2012.08.010
38. Marriott B, Arturo Garza-Reyes J, Soriano-Meier H, Antony J (2013) An integrated methodology to prioritise improvement initiatives in low volume-high integrity product manufacturing organisations. J Manuf Technol Manage 24(2):197–217. https://doi.org/10.1108/17410381311292304
39. Maucec M, Ravnik M, Glumac B (1998) Criticality analysis of the multiplying material inside the Chernobyl sarcophagus. Nuclear Technology. Available at http://www.ans.org/pubs/journals/nt/a_2867. Accessed 6 Feb 2017
40. McGinnis P (1969) Failure mode, effects and criticality analysis of ASE EMI modifications. http://www.lpi.usra.edu/lunar/ALSEP/pdf/31111000667079.pdf. Bendix Aerospace Systems Division
41. McKinney G, Iverson J (1995) MCNP perturbation technique for criticality analysis. ICNC Meeting, Albuquerque, NM. Available at http://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-95-2015. Accessed 6 Feb 2017
42. Mechefske CK, Wang Z (2003) Using fuzzy linguistics to select optimum maintenance and condition monitoring strategies. Mech Syst Signal Process 17(2):305–316. https://doi.org/10.1006/mssp.2001.1395
43. Ming Tan C (2003) Customer-focused build-in reliability: a case study. Int J Qual Reliability Management 20(3):378–397. https://doi.org/10.1108/02656710310468560
44. Molenaers A, Baets H, Pintelon L, Waeyenbergh G (2012) Criticality classification of spare parts: a case study. Int J Prod Econ 140(2):570–578. https://doi.org/10.1016/j.ijpe.2011.08.013
45. Moore WJ, Starr AG (2006) An intelligent maintenance system for continuous cost-based prioritisation of maintenance activities. Comput Ind 57(6):595–606. https://doi.org/10.1016/j.compind.2006.02.008
46. Moskowitz L (1971) Array E PCU failure modes, effects and criticality analysis. Available at https://repository.hou.usra.edu/handle/20.500.11753/333. Accessed 6 Feb 2017
47. Neves LC, Frangopol DM (2005) Condition, safety and cost profiles for deteriorating structures with emphasis on bridges. Reliab Eng Syst Saf 89(2):185–198. https://doi.org/10.1016/j.ress.2004.08.018
48. Norsok T (2001) NORSOK STANDARD criticality analysis for maintenance purposes
49. Nyström B, Söderholm P (2010) Selection of maintenance actions using the analytic hierarchy process (AHP): decision-making in railway infrastructure. Structure and Infrastructure Engineering, pp 467–479. https://doi.org/10.1080/15732470801990209
50. Pappas G, Pendleton R (1983) Failure modes, effects and criticality analysis (FMECA) of type AN/GRN-27 (V) instrument landing system with traveling-wave localizer antenna. Available at http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA128930. Accessed 6 Feb 2017
51. Partovi FY, Anarajan M (2002) Classifying inventory using an artificial neural network approach. Computers and industrial engineering, pp 389–404
52. Pelaez CE, Bowles JB (1994) Using fuzzy logic for system criticality analysis. In Proceedings of annual reliability and maintainability symposium (RAMS), pp 449–455. https://doi.org/10.1109/rams.1994.291150

53. Porras E, Dekker R (2008) An inventory control system for spare parts at a refinery: an empirical comparison of different re-order point methods. Eur J Oper Res, 101–132
54. Pschierer-barnfather P, Hughes D, Holmes S (2011) Determination of asset criticality: a practical method for use in risk-based investment planning. CIRED2011 Frankfurt (1013), pp 6–9
55. Ramanathan R (2006) ABC inventory classification with multiple-criteria using weighted linear optimization. Comput Oper Res, 695–700
56. Ramirez-Marquez JE, Coit DW (2007) Multi-state component criticality analysis for reliability improvement in multi-state systems. Reliab Eng Syst Saf 92(12):1608–1619. https://doi.org/10.1016/j.ress.2006.09.014
57. Sankar NR, Prabhu BS (2001) Modified approach for prioritization of failures in a system failure mode and effects analysis. Int J Qual Reliab Manage Emerald 18(3):324–336. https://doi.org/10.1108/02656710110383737
58. Shin J, Jun H, Catteneo C, Kiritsis D (2015) Degradation mode and criticality analysis based on product usage data. Int J of Adv Mfg Tech. Available at http://link.springer.com/article/10.1007/s00170-014-6782-7. Accessed 17 Feb 2017
59. Smith R (2015) Equipment Criticality Analysis. Analysis, pp 1–20
60. Spencer D (1962) Thermal and criticality analysis of the plasma core reactor. Available at http://www.osti.gov/scitech/biblio/4812653. Accessed 6 Feb 2017
61. Staat J, Dallaire R, Roukas R (1970) ALSEP flight system 5 (array D) system level failure mode effects and criticality analysis ATM 906 ALSEP flight system 5 (array D) system level failure mode effects and criticality analysis. Aerospace System Division
62. Stoll J, Kopf R, Schneider J, Lanza G (2015) Criticality analysis of spare parts management: a multi-criteria classification regarding a cross-plant central warehouse strategy. Production Engineering. Available at http://link.springer.com/article/10.1007/s11740-015-0602-2. Accessed 17 Feb 2017
63. Syntetos A, Keyes M, Babai M (2009) Demand categorisation in a European spare parts logistics network. Int J Oper Prod Manage 29(34):292–316
64. Teunter R, Babai M, Syntetos A (2010) ABC classification: service levels and inventory costs. Production and Operations Management, pp 343–352
65. Theoharidou M, Kotzanikolaou P, Gritzalis D (2009) Risk-based criticality analysis. International conference on. Available at http://link.springer.com/chapter/10.1007/978-3-642-04798-5_3. Accessed 17 Feb 2017
66. Tsakatikas D, Diplaris S, Sfantsikopoulos M (2008) Spare parts criticality for unplanned maintenance of industrial systems. Eur J Ind Eng 2(1):94. https://doi.org/10.1504/EJIE.2008.016331
67. USM Standard (1980) Mil-Std-1629a Procedures for Performing a Failure Mode, Effects and Criticality Analysis. Military Standard. MIL-1629a, pp 1–54. doi:MIL-STD-1629A
68. Walski T, Weiler J, Culver T (2006) Using criticality analysis to identify impact of valve location. Proceedings of 8th annual water ..., pp 1–9. doi:https://doi.org/10.1061/40941(247)31
69. Wilson R, Johnson J (1983) Application of logic trees for criticality safety analysis at the ICPP. Trans Am Nucl Soc (United States). Available at http://www.osti.gov/scitech/biblio/5790793. Accessed 6 Feb 2017
70. Xu K, Tang LC, Xie M, Ho SL, Zhu ML (2002) Fuzzy assessment of FMEA for engine systems. Reliab Eng Syst Saf 75(1):17–29. https://doi.org/10.1016/S0951-8320(01)00101-6
71. Xu K, Zhu M, Luo M (2000) SAE TECHNICAL Turbocharger's failure mode criticality analysis using fuzzy logic. SAE Technical Papers. https://doi.org/10.4271/2000-01-1350
72. Ye F, Kelly T (2004) Criticality analysis for cots software components. In: Proceedings of 22nd international system safety conference (ISSC'04). Available at https://www.cs.york.ac.uk/~tpk/issc04a.pdf. Accessed 17 Feb 2017

73. Zhou P, Fan L (2007) A note on multi-criteria ABC inventory classification using weighted linear optimization. Eur J Oper Res
74. Ćatić D, Jeremić B, Djordjević Z (2011) Criticality analysis of the elements of the light commercial vehicle steering tie-rod joint. *Strojniški vestnik-Journal*. Available at http://ojs.sv-jme.eu/index.php/sv-jme/article/view/sv-jme.2010.077. Accessed 17 Feb 2017

Semiparametric Valuation of Heterogeneous Assets

Roar Adland and Sebastian Köhn

Abstract We propose a semiparametric valuation model for heterogeneous assets that fits within Generalized Additive Model (GAM) framework. Using micro data for individual asset sales we are able to estimate the impact and relative importance of macro market conditions and the influence of technical specifications and asset age. We apply our model to the valuation of oceangoing chemical tankers. Our empirical results suggest that asset valuation is a non-linear function of main drivers such as ship size, age, and market conditions, whilst other engineering parameters that are specific to the chemicals market such as tank coating grade and cargo diversity also play a significant role. An asset valuation model that can account for generic market factors as well as highly heterogeneous asset-specific characteristics is important for owners and financiers, particularly in markets with limited liquidity.

1 Introduction

The ocean transportation of cargoes such as chemicals and petrochemical gases is undertaken by vessels that are technologically advanced, highly specialised and capital intensive, with a wide range of technical specifications. The heterogeneous and relatively small global fleet of such vessels and a concentrated ownership structure leads to low liquidity in the second-hand asset markets. For these reasons, asset valuation is a much more challenging task than for the larger segments of commodity shipping, but no less important for the market players and financial institutions involved. Research into the formation of second-hand ship prices has hitherto been based on time series of values for generic vessels in the tanker or drybulk sectors (see for instance [3, 5–7]). As the sole exception, Adland and

R. Adland (✉)
Norwegian School of Economics (NHH), Bergen, Norway
e-mail: roar.adland@nhh.no

S. Köhn
Deloitte, Hamburg, Germany
e-mail: mail@sebastian-koehn.de

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_3

Koekebakker [2] propose a nonparametric ship valuation model based on sales data for Handysize bulk carriers, though they consider only the size and age of the ship and the state of the freight market.

In this paper, we extend the research on ship valuation using micro data on vessel transactions and technical specifications by proposing a semi-parametric generalized additive model. Under the assumption of separable factors, we can quantify the pricing effect of a large number of technical variables in the valuation of highly sophisticated chemical tankers. Such a vessel valuation model is particularly valuable for brokers, financiers and owners when performing "desktop valuations" of specialised ships where brokers estimates are costly or perhaps not available.

2 Methodology

A Generalized Additive Model is the extension of a generalized linear model to a combination of linear predictors and the sum of smooth functions of explanatory variables. In general, a model may look like

$$g(\mu_i) = X_i^* \theta + f_1(x_{1i}) + f_2(x_{2i}, x_{3i}) \ldots \tag{1}$$

where $\mu_i \equiv E(Y_i)$, Y_i is the response variable distributed according to some exponential family distribution, X_i^* is a vector of explanatory variables that enter the model parametrically, θ is a corresponding parameter vector and f_j are smooth functions of the variables that are modelled non-parametrically. GAMs provide enough flexibility to take non-linear relationships into account without making any specific assumptions about the functional form of these relations. They also provide reliable results in samples of moderate size.

The bases for our estimations are thin plate regression splines (TPRS) in combination with a general cross validation procedure (GCV). Standard bases for regressions splines, such as cubic splines, require the user to choose knot locations, i.e. the basis dimension. Furthermore, they allow only for the representation of the smooth of one predictor variable. TPRS surmount these problems and are in a limited sense 'optimal' with respect to these problems[1]. Given the problem of estimating $g(x)$ such that $y_i = g(x_i) + \varepsilon_i$, thin plate spline smoothing estimates $g(x)$ by finding the function f minimising

$$\|y - f\|^2 + \lambda J_{md}(f) \tag{2}$$

where $J_{md}(f)$ is a penalty function measuring the "wiggliness" of f and λ is a penalisation parameter. Instead of choosing the basis dimension to control the

[1] See Wood [8, pp. 154].

models smoothness, the trade-off between model fit and smoothness is controlled by the smoothing parameter λ. If $\lambda = 0$ the spline is unpenalized while $\lambda \rightarrow \infty$ leads to a straight-line estimate (over-smoothing). Furthermore, λ relates to the effective degrees of freedom (EDF) of a smooth term, i.e. the larger the complexity of a smooth term the larger the EDF and the lower λ.

The problem of choosing some optimal value for λ is solved via generalized cross validation (GCV). For more details on GAM's and the practical implementation the interested reader is referred to for example Härdle et al. [4] or Wood [8]. One disadvantage in using GAMs is that hypothesis testing is only approximate and that satisfactory interval estimation requires a Bayesian approach. Using the Bayesian posterior covariance matrix and a corresponding posterior distribution allows us to calculate p-values and confidence intervals. Typically, the p-values calculated this way will be too low, because they are conditional on the uncertain smoothing parameter [8]. Therefore, we are restrictive when interpreting results and significance levels.

Our independent variables include both macro and ship-specific variables and can be justified as follows (subscript i is omitted, but refers to the value of the variable for sales transaction i, or at the time of transaction i for the macro variables):

NB	Newbuilding price (USD/Compensated Gross Tonnes, CGT). The cost of ordering a brand new vessel (i.e. replacement value).
EARN	spot market vessel earnings ($/day) as calculated on the benchmark Houston–Rotterdam route basis $/tonne rates for 3000 tonnes 'easychem' parcels.
SIZE	deadweight carrying capacity of the vessel (tonnes). A larger vessel should attract a higher price due to higher earnings capacity, all else equal.
SPEED	design speed of the vessel (knots). A greater speed indicates higher efficiency, though this may come at a cost of higher fuel consumption.
AGE	Age of the vessel at the time of the sale (years). As vessels depreciate, older vessels have lower values, all else equal.
NOTANK	the number of cargo tanks. A higher number is increases the potential number of different chemical parcels carried simultaneously.
I^{HULL}	Dummy variable indicating hull configuration (double hull, double bottom, double sides and single hull).
I^{COAT}	Dummy variable indicating tank coating type (epoxy-, polyurethane-, zinc-, and stainless steel-coating). Higher-grade coating increases the cargo flexibility.
I^{IMO}	Dummy variable for the vessel's IMO classification of the environmental and safety hazard of the cargoes (Type 1, 2, 3 with Type 1 being most severe).
$I^{COUNTRY}$	Country of build as a proxy for perceived overall build quality.

CARGODIV an interaction variable representing cargo diversity of the vessel as measured by the product of the number of coatings and number of tanks.

PUMPDIV an interaction variable representing the ability and flexibility of cargo handling as the product of the number of discharge pumps and pump capacity

As an example, the most comprehensive model specification can be written as:

$$g(E(PRICE_i|.)) = \gamma_0 + s(NB_i) + s(EARN_i) + s(SIZE_i) + s(AGE_i) + I_i^{HULL} + s(CARGODIV_i) + s(PUMPDIV_i) + I_i^{IMO} + s(SPEED_i) + I_i^{COUNTRY} \quad (3)$$

All regressions are carried out using $g(.) = \log(.)$ as link-function and assumes that second hand prices follow a Gamma distribution, $PRICE_i \sim G(\alpha, \beta)$ (α, β). Experiments with the Normal distribution and different link-functions did not improve results.

3 Data and Empirical Results

Tables 1 and 2 show the descriptive statistics for our variables. The dataset obtained from Clarkson Research Ltd. includes 842 observations of chemical tanker sales since October 1990. We remove vessels sold under unusual circumstances, including those sold at auction, judicial sales, vessels sold with attached time charter contracts and en-bloc transactions, leaving 736 observations for further analysis.

Table 3 shows the regression results for four different specifications (A to D) of the general pricing model in Eq. 1. The upper panel shows the results for the non-parametric components as estimated degree of freedom (EDF) which reflects the degree of non-linearity present in the regressors and the significance of this explanatory factor. The lower panel provides information on the parametric components given as point estimate (PE) and its significance which can be interpreted directly.

Table 1 Data overview

Variable	Min	Average	Max
Price (mUSD)	0.25	11.09	100.00
NB price (USD/CGT)	829	1012	1315
Earnings (USD/day)	6054	17,205	40,984
Age	0	12	25
Size (DWT)	1032	17,485	50,600
No. tanks	4	17	43
No. pumps	2	12	43
Pump capacity	150	2890	8670
Speed (knots)	10.5	13.5	17.0

Table 2 Data distribution for build country, coating, hull and IMO type

Build country		Coating		Hull type		IMO type	
Japan	44.9%	Epoxy	46.5%	D/ Bottom	46.3%	IMO1	29.2%
S. Korea	10.7%	S. Steel	6.4%	D/Hull	29.4%	IMO2	46.6%
Croatia	7.9%	S. Steel Epoxy	6.0%	D/ Sides	2.7%	IMO3	3.9%
Norway	5.6%	S. Steel Epoxy Zinc	6.4%	S/Skin	8.0%	N/A	20.2%
Denmark	4.3%	S. Steel Poly.	11.1%	N/A	13.7%		
Germany	3.8%	S. Steel Zinc	7.1%				
Sweden	3.0%	Zinc	4.1%				
Others	19.8%	Misc.	12.6%				

Table 3 Regression results for models A through D

	A		B		C		D	
	EDF	Sig.	EDF	Sig.	EDF	Sig.	EDF	Sig.
NB	6.102	***	6.107	***	5.566	***	5.438	***
EARN	5.454	***	5.934	***	3.197	*	3.019	*
SIZE	8.373	***	8.225	***	7.607	***	6.701	***
AGE	7.413	***	6.109	***	6.723	***	6.962	***
NOTANKS	–		8.356	**	–		–	
CARGODIV	–		–		8.230	***	7.884	***
PUMPDIV	–		–		1.000	***	1.000	***
SPEED	–		–		–		7.228	***
	PE	Sig.	PE	Sig.	PE	Sig.	PE	Sig.
D/Bottom	–		−0.260	***	−0.217	***	−0.164	***
D/Sides	–		−0.418	**	−0.181	*	−0.207	*
S/Skin	–		−0.427	***	−0.298	***	−0.244	***
S. Steel	–		0.131		–		–	
Poly	–		0.042		–		–	
Epoxy	–		0.261		–		–	
Zinc	–		0.062	*	–		–	
IMO2	–				0.254	***	0.257	***
IMO3	–				0.299	**	0.303	***
Intercept	2.049	***	2.041	***	2.066	***	2.004	***
No. obs.	736		736		736		736	
Adj. R^2	81.9%		84.7%		85.7%		86.3%	

The results are broadly consistent across specifications and can be summarized as follows. Firstly, the relationships between asset value and the replacement cost (NB price), vessel age, earnings and vessel size are non-linear and highly significant. Secondly, non-double-hulled tonnage attracting a substantial discount. Thirdly, perhaps somewhat surprising, tank coating does not significantly affect asset values. Fourthly, our proxies for versatility and efficiency (CARGODIV, PUMPDIV, SPEED) are highly significant. Finally, IMO classification matters, albeit perhaps not in the way expected, as IMO2 and IMO3 vessels carry a premium compared to the technically more advanced IMO1 vessels. The explanatory power of the model is relatively high, starting out at 81.9% for the basic 'macro' model and increasing to 86.3% as we add more technical vessel variables.

Table 4 presents the results for our most comprehensive model (Eq. 3). The results from the earlier specifications remain robust. Additionally, vessels built in certain countries (Denmark, Germany and Norway) attract a quality premium, while Ukrainian-built tonnage has perceived lower quality reflected in asset values.

Figure 1 presents the smooth of new building prices, earnings, size and age to second-hand prices. The relationships are strongly non-linear, with small confidence bands. Second-hand prices increase with size, decrease with age, and are also broadly increasing with the replacement cost and spot market earnings. The latter effect is less clear for high values, due to fewer observations and mean reversion in rates [1].

Table 4 Regression results for final model

	Model E		Country of build		
	EDF	Sig.		PE	Sig.
NB	5.902	***	Belgium	0.342	*
EARN	3.514	***	China	0.134	
SIZE	4.114	***	Croatia	0.045	
AGE	5.871	***	Denmark	0.448	***
NOTANKS	–		Finland	0.282	*
CARGODIV	7.412	***	France	−0.130	
PUMPDIV	1.001	**	Germany	0.402	***
SPEED	7.785	***	Italy	0.227	*
	pe	Sig.	Netherlands	0.232	*
D/Bottom	−0.113	*	Norway	0.307	***
D/Sides	−0.413	***	Poland	−0.114	
S/Skin	−0.200	**	South Korea	−0.054	
IMO2	0.166	***	Spain	0.172	*
IMO3	0.218	*	Sweden	0.149	*
Intercept	1.938	***	Turkey	0.232	*
			Ukraine	−0.381	***
No. obs	736		UK	0.056	
Adj. R^2	88.0%		Other	0.039	

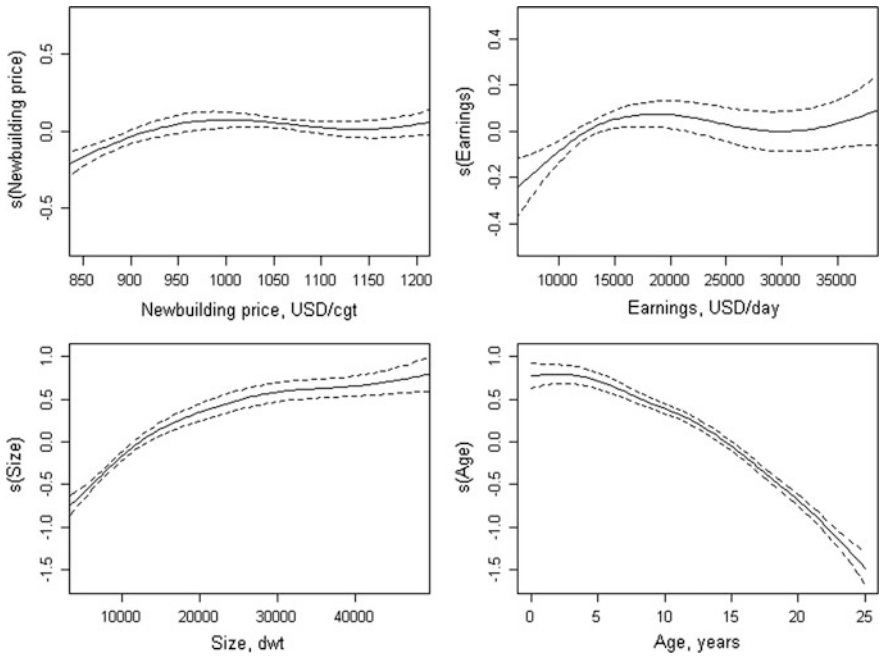


Fig. 1 Smooth of NB price, earnings, size, and age

Figure 2 illustrates the joint non-linear effect of vessel age and vessel size on second-hand prices, similar to Adland and Koekebakker [2].

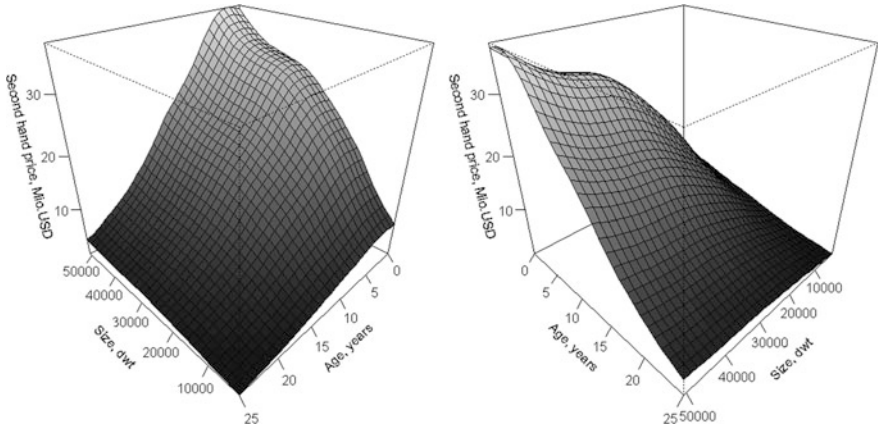


Fig. 2 3D plot of asset values against size and age

4 Concluding Remarks

We have developed a comprehensive multivariate semi-parametric framework for the estimation of chemical tanker second hand prices. Previous non-parametric models have shown that non-linear modelling is appropriate, but have suffered from the curse of dimensionality. Our model surmounts these issues and extends the existing literature by applying semi-parametric GAMs to a cross sectional dataset of actual sale and purchase transactions of chemical tankers. Even the heterogeneous nature of chemical tankers and the high variation in chemical tanker second hand prices can be satisfactorily modelled with this framework. Ship specific factors which have not been included in previous models are shown to have a significant impact on prices and the explanatory power of this model appears to outperform linear methods of estimation. Most of the factors turned out to show the expected effects on prices. To sum up, semi-parametric methods—especially GAMs—provide an appropriate framework to model asset values for highly heterogeneous assets.

References

1. Adland R, Cullinane K (2006) The non-linear dynamics of spot freight rates in tanker markets. J Transp Res E 42(3):211–224
2. Adland R, Koekebakker S (2007) Ship valuation using cross-sectional sales data: a multivariate non-parametric approach. Marit Econ Logistics 9:105–118
3. Glen DR (1997) The market for second-hand ships: further results on efficiency using cointegration analysis. Marit Policy Manage 24:245–260
4. Härdle W, Müller M, Sperlich S, Werwatz A (2004) Nonparametric and semiparametric models. Springer, Berlin
5. Kavussanos MG (1996) Price risk modelling of different size vessels in the tanker industry. Logistics Transp Rev 32:161–176
6. Kavussanos MG (1997) The dynamics of time-varying volatilities in different size second-hand ship prices of the dry-cargo sector. Appl Econ 29:433–443
7. Tsolakis S, Cridland C, Haralambides H (2003) Econometric modelling of second-hand ship prices. Marit Econ Logistics 5(4):347–377
8. Wood SN (2006) Generalized additive models—an introduction with R. Chapman & Hall/CRC, London

Fluid Induced Vibrations in Rotors Supported by Journal Bearings: A Case Study

Manish Agrawal and R. V. S. Krishnadutt

Abstract Rotors are supported by oil in fluid film bearings. Usually the oil film has a damping effect on the rotor. However, for some rotors, the oil film may become source of self-induced vibrations and cause instability in the system. Unlike the forced vibration response whose frequency is equal to the excitation force, self-exciting vibrations have a frequency equal to one of the rotor modes. As the source of vibrations in not linked to external excitation, vibration diagnostics of this kind of problems becomes complex. A case of a rotor failure due to self-exciting vibrations, its failure investigation and subsequent remedial measures are presented in this paper.

1 Introduction

Vibrations originating from fluid induced forces in the clearances between rotor and stator are associated with sub-synchronous frequency [1] and the phenomenon is called fluid whirl or fluid whip. The phenomenon is responsible for self-excited vibrations resulting in rotor instability. Fluid whirl frequency is at about 0.43–0.49 times the rotational speed as per Oliver [7] whereas fluid whip frequency gets locked to the first bending mode of the rotor. However, fluid whirl and fluid whip belong to the same phenomenon and are generated by the same source, i.e. fluid induced forces. Fluid whirl shows up if the instability threshold occurs at low speeds and transits into fluid whip at speed about twice the frequency of first bending mode. Fluid whip continues as the speed goes up and the frequency of sub synchronous vibrations gets locked to the frequency of first bending mode. Oil whip was first reported by Newkirk

M. Agrawal (✉) · R. V. S. Krishnadutt
Machine Dynamics and Failure Analysis, Bharat Heavy Electricals Limited
Corporate Research and Development, 500093 Hyderabad, India
e-mail: manishagrawal@bhel.in

R. V. S. Krishnadutt
e-mail: krishnadutt.rvs@gmail.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_4

and Taylor [6]. A detailed description of self-excited vibrations and evaluation of stability zones has been explained by Muszynska [3–5].

In a hydrodynamic fluid film bearing, the lubrication oil rotates along with the rotor at some circumferential velocity. This circumferential velocity is equal to the rotor velocity at the rotor surface and gradually decreases towards the bearing surface. The faster is the rotor speed, the higher is the fluid average velocity. Average circumferential velocity of the fluid in the annular space between rotor and bearing is approx. half of the rotor speed. Muszynska has shown that the destabilizing fluid force, pushing on to the rotor is proportional to the fluid average circumferential velocity. The magnitude of the sub synchronous vibrations due to this fluid force are self-exciting and grow till the magnitude of vibrations becomes equal to the oil clearance and the shaft touches the bearing surface.

Oil-whirl is usually a weak vibration phenomenon and not observed as the response is dominated by unbalance. Therefore, the oil whip frequency is commonly considered as the threshold of instability. However, it is essential to accurately identify the instability threshold at design stage to avoid any unexpected vibration behavior during operation, for high speed flexible rotors. This paper presents a case where oil instability induced vibrations caused large vibrations resulting in a rotor failure and the remedial measures taken up to mitigate the instability.

2 Problem Statement: Failed Rotor and Large Vibrations

The rotor was under balancing at the balancing tunnel in the factory. The rotor synchronous vibrations were within acceptable limits and it was decided to take up the over-speed test. However, the rotor experienced catastrophic failure at over-speed. Figure 1 shows the broken rotor assembly after the failure. Figure 2 shows the run up plot before the failure and run down plot recorded until complete failure. It can be seen from these plots that the synchronous vibrations do not seem to be the cause of failure and have gone up subsequent to the failure.

The failure investigation pointed out at the shrink fit release resulting in large unbalance and amplified response due to vicinity of over-speed to second critical

Fig. 1 Rotor assembly after failure at balancing tunnel for the failed rotor

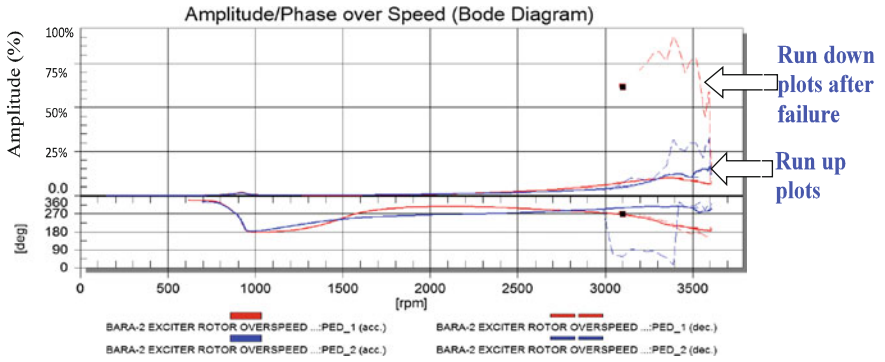


Fig. 2 Run up and run down plots showing amplitude of synchronous component for the failed rotor

speed of the rotor bearing system as possible causes of failure. The shrink fit was improved and critical speed was pushed up by reducing the bearing span. This time, the output of the vibration sensors was connected to a vibration analysis instrument and the output was recorded for diagnostic purposes.

The rotor could successfully clear the over-speed test with acceptable synchronous vibrations, however, extremely large overall vibrations were observed at higher speeds. Run up and run down plots for this rotor are shown in Fig. 3. The synchronous component is well controlled. The gap in OA and 1X vibrations is attributed to sub synchronous component, as shown by waterfall plots, Fig. 4. The sub-synchronous component shows up at around 2400 rpm and steadily grows with increasing speed. The frequency of the sub synchronous component is observed to be fixed and not varying with speed.

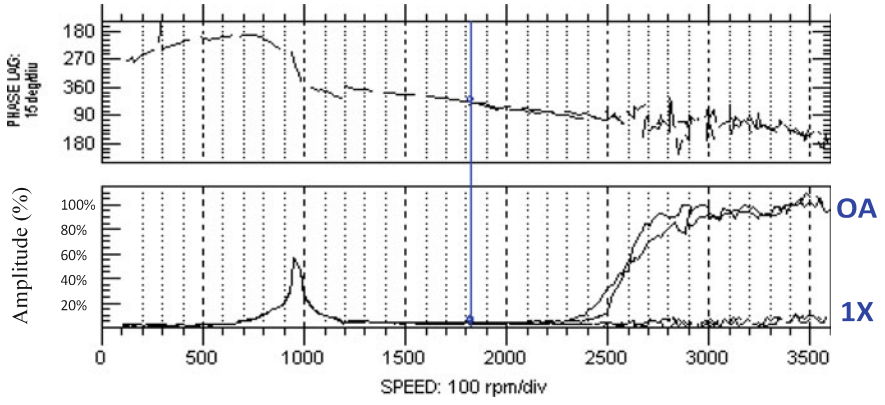


Fig. 3 Measured run up and run down plot for a successful rotor showing unusually large gap between OA and 1X, 2500 rpm onwards

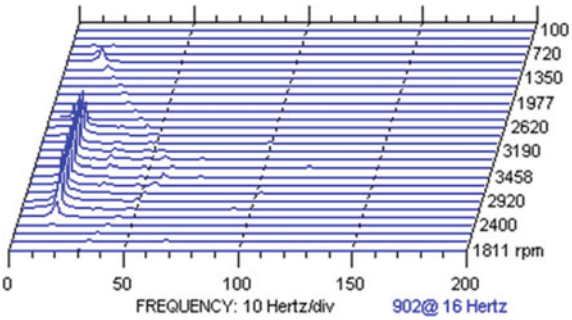


Fig. 4 Waterfall plot (non drive end) showing large sub synchronous component

3 Rotordynamic Analysis: Fluid Induced Instability

In view of the presence of a large sub synchronous component in the measurements, the rotordynamic behaviour of the rotor bearing system was reviewed. The primary aim of the analysis was to detect the instability threshold and identification of the response frequency, expected during the unstable behavior. MADYN2000 has been used for numerical simulations.

Figure 5 shows the first flexural mode of the rotor bearing system. The rotor is supported on elliptical bearings. The first bending mode at 890 rpm (14.8 Hz) is shown to have negative damping, indicating an unstable mode. Negative damping is confirmed by the Campbell and modal damping diagram, shown in Fig. 6.

Journal bearings are generally preferable for reducing the vibration response, due to the damping effect, however, insufficient design poses a risk of self-exciting vibrations above one of the rotor critical speed, usually the first critical speed. Run up plot for this case (Fig. 3) shows the harmonic resonance at around 950 rpm. When the speed increases beyond first resonance, self-exciting vibrations, known as oil whirl occurs. In the oil whirl condition, the rotor whirls like a rigid rotor in forward direction with small amplitude. The frequency of oil whirl is usually from 0.39 to 0.48 of rotational speed. The actual value is dependent upon type of bearing and the shaft location in the bearing. With further increase in the speed, the rotor

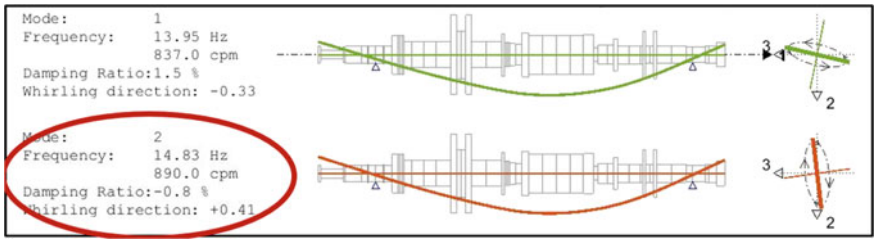


Fig. 5 Critical speeds of the rotor bearing system from eigenvalue analysis

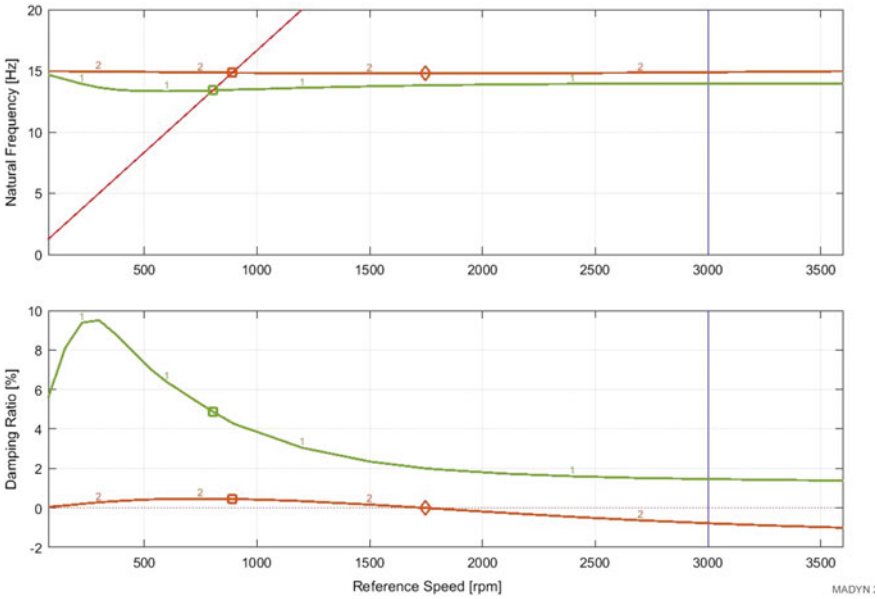


Fig. 6 Campbell and modal damping diagram showing negative damping for first bending mode

crosses the threshold of instability, which is usually at about twice the first critical speed. Transition of self-exciting from oil whirl to oil whip takes place. Frequency of self-exciting vibrations due to oil whip is locked to the frequency of first mode. For bearings with medium to heavy load, the oil whirl may not show up and oil whip may occur directly. This statement is shown to be true for the rotor under investigation. The bearings have specific load of approx. 1.0 N/mm^2 thus falling in the category of bearings with medium load. Measurements have shown the occurrence of oil whip directly without the appearance of oil whirl. Measurements show the threshold of instability at around 2200 rpm and the sub synchronous component rises violently with the increase in speed beyond this.

Nature and threshold of instability have been captured through non-linear transient rotordynamic analysis. Figure 7 shows amplitude versus time plots at 2550 rpm. The vibrations are self-exciting and very large in amplitude, reaching equal to the available oil clearance in the bearing.

4 Instability Due to Cross Coupling Stiffness Coefficients

Destabilizing force arising from the cross coupling stiffness terms must be suppressed by the dissipative force arising from the damping. However, acceleration of rotor results in rapid increase in cross coupling stiffness terms and the destabilizing force may surpass the dissipative force leading to instability of rotor [2].

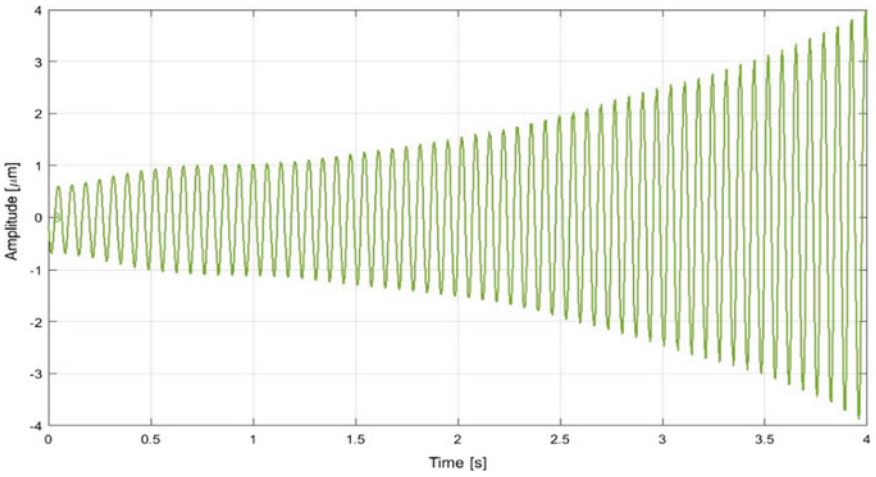


Fig. 7 Amplitude versus time plots at 2550 rpm, obtained from non-linear transient analysis

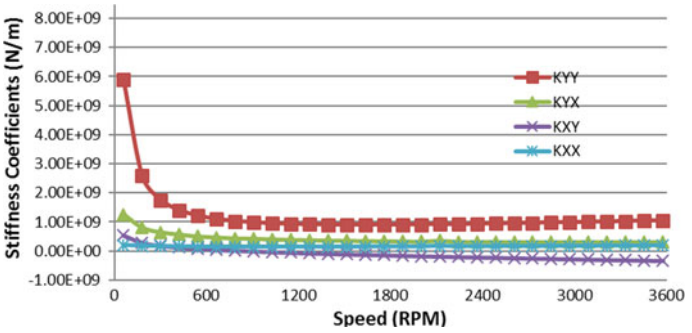


Fig. 8 Stiffness coefficients for elliptical bearing

Figure 8 shows the stiffness coefficients versus rotor speed graph for the elliptical bearing. Amplitude of the cross coupling stiffness term KXY increases rapidly with speed. This leads to destabilizing force dominating the dissipative force, causing the instability.

5 Remedial Measure: Use of Tilting Pad Bearings

The possible remedial measures to avoid instability include increasing the natural frequencies of the system by reducing the bearing span or increasing the shaft diameter, increasing the instability threshold, increase the bearing load, increase the

eccentricity ratio (eccentricity ratio = eccentricity/oil clearance) and using a tilting pad bearing. Tilting pad bearings do not have significant cross coupling terms and the destabilizing tangential force from oil film on the rotor.

Figures 9 and 10 show the eigenvalue analysis results and the Campbell and modal damping diagram for the rotor with tilting pad bearings. Damping ratio for the first bending mode, which is negative with original elliptical bearings, has now turned positive. However, the damping ratio at 0.4% indicates insufficient damping and hence, the sub-synchronous component is expected to be present though with a small amplitude.

Figure 11 (Estimated amplitude versus time plots at operating speed) shows that the self-exciting sub synchronous vibrations do not diminish, thus confirming the insufficient damping. Figure 12 shows the stiffness coefficients vs rotor speed graph. It can be seen that cross coupling stiffness terms KYX and KXY are insignificant as compared to the elliptical bearing.

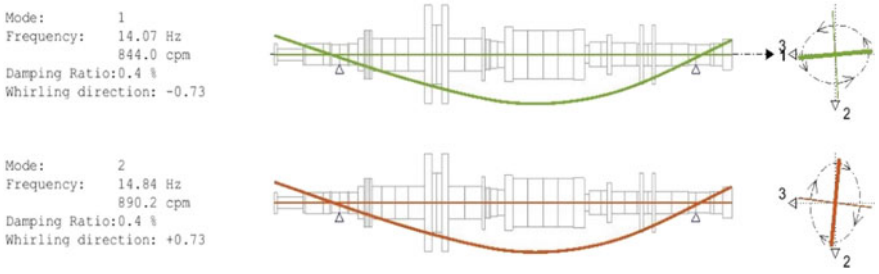


Fig. 9 Critical speeds of the system with tilting pad bearings

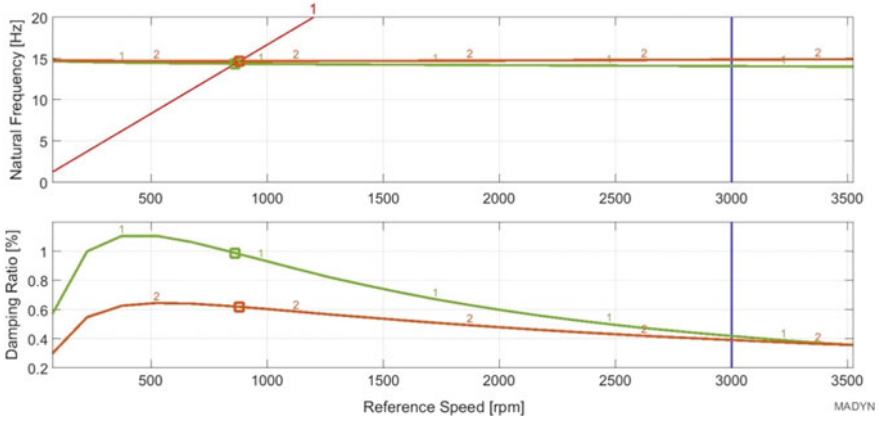


Fig. 10 Campbell diagram for the system with tilting pad bearings

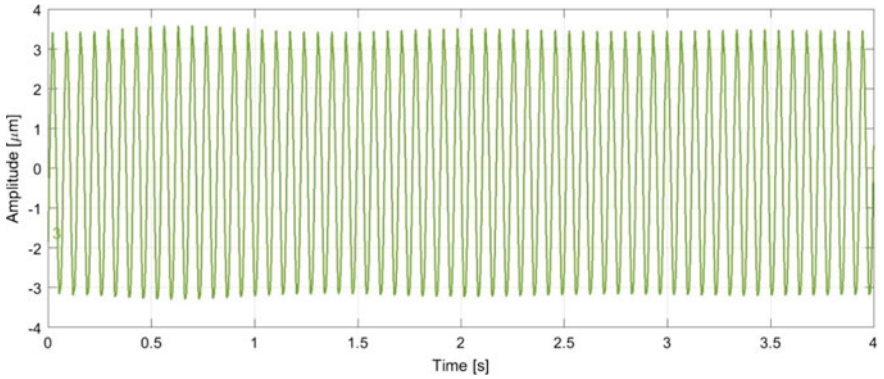


Fig. 11 Amplitude versus time plots at operating speed (3000 rpm)

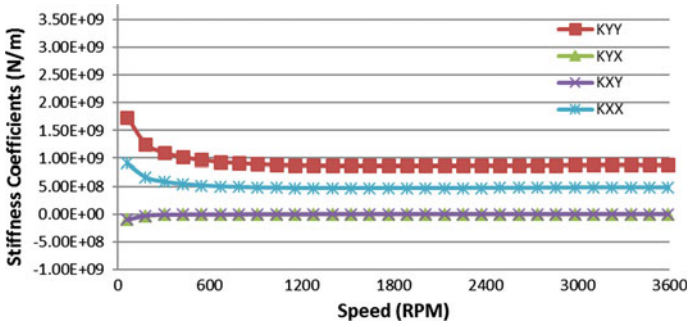


Fig. 12 Stiffness coefficients for tilting pad bearing

The bearings were replaced by tilting pad bearings and the rotor was operated in the balancing tunnel. Figures 13 and 14 show the measured run up and waterfall plots. The waterfall plot shows the presence of sub synchronous component in small amplitude, above 2000 rpm. This validates the results obtained from non-linear analysis.

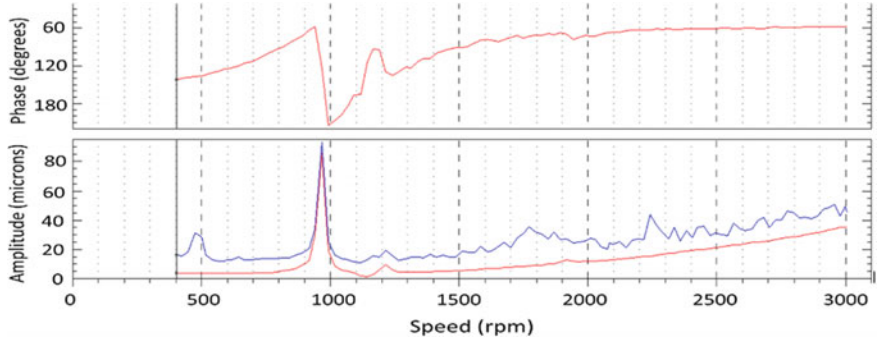


Fig. 13 Measured run up plot for the rotor with tilting pad bearings

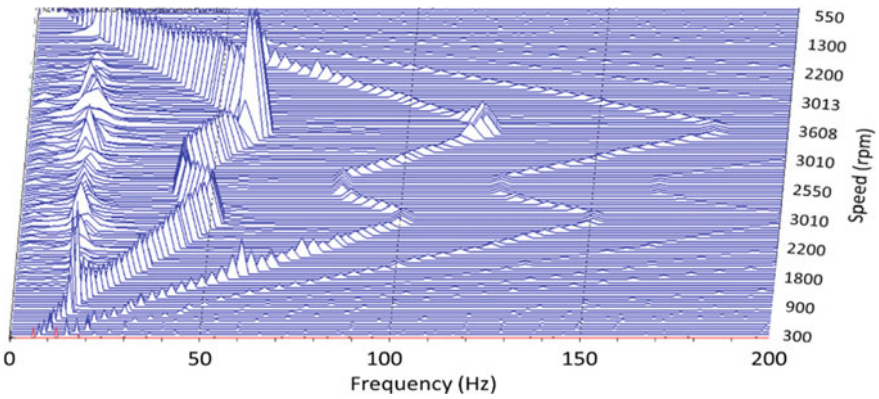


Fig. 14 Measured waterfall plot for the rotor with tilting pad bearings

6 Conclusions

Following conclusions are drawn from this case study:

1. Thresholds of instability and frequency of the unstable mode have been predicted through the rotordynamic analysis and are in good agreement with the measurements. The work successfully reiterates the need of rotordynamic analysis considering the stability of the rotor bearing system and prediction of any instability, present in the system at design stage to avoid the unpleasant surprises during operation.
2. The elliptical bearings with preload are known to be more stable than the cylindrical bearings. However, in this case, use of elliptical bearings has not helped in preventing the occurrence of instability.
3. Tilting pad bearings have been introduced replacing the two lobe elliptical bearings to mitigate the instability. However, sub synchronous component originating from oil whip is present even in case of the rotor supported on tilting pad bearings, though in small amplitude. Thus, it can be inferred that the self-exciting vibration mechanism cannot be eliminated by changing the bearings alone. This requires modification of the rotor design.
4. In the balancing tunnels world over, balancing teams consider the synchronous vibrations alone and usually do not monitor the overall vibration levels. However, it may be a good practice to monitor the overall vibrations and in case of significant non synchronous component, the diagnostic team can be alerted.

References

1. Allaire PE, Flack RD (1981) Design of journal bearings for rotating machinery. In: Proceedings of the tenth turbo machinery symposium, pp 25–45
2. Mendes RU, Cavalca KL (2014) Research article on the instability threshold of journal bearing supported rotors. Int J Rotating Mach vol 2014, Article ID 351261
3. Muszynska A (1986) Whirl and whip-rotor/bearing stability problems. J Sound Vib 110(3):443–462
4. Muszynska A (1988) Stability of whirl and whip in rotor/bearing systems. J Sound Vib 127(1):49–64
5. Muszynska A (1998) Oil whip of a rotor supported in poorly lubricated bearing, Orbit, September 1998, pp 4–8
6. Newkirk BL, Taylor HD (1925) Shaft whipping due to oil action in journal bearings. Gen Electric Rev 28:559–568
7. Oliver G (2001) An introduction to oil whirl and oil whip. Turbo Compon Eng Newsl: Bearing J 3(2):1–2

Flood Exposure and Social Vulnerability for Prioritizing Local Adaptation of Urban Storm Water Systems

Tanvir Ahmed, Abbas El-Zein, Fahim Nawroz Tonmoy, Federico Maggi and Kon Shing Kenneth Chung

Abstract Changes in rainfall patterns due to climate change are likely to increase weather-related hazards significantly, including floods and associated economic and social damage in urban areas. Adapting storm water systems to cope with these changes can be expensive, and often competes for resources with other priorities. Therefore, targeted adaptation of the infrastructure in a municipality, based on an assessment of both physical risks of flooding and socioeconomic vulnerability, can help decision-makers better allocate resources while developing adaptation plans. Conventional flood studies commissioned by municipal councils typically concentrate on geophysical aspects such as flooding frequency and areal extent. This paper goes beyond those and develops a Flood Social Vulnerability model (FSV) by combining hydrological and hydraulic analyses with social vulnerability analyses. The approach is illustrated by applying it to Marrickville, a local government area in Sydney's inner-west that is prone to flooding.

T. Ahmed (✉) · A. El-Zein · F. Maggi · K. S. K. Chung
School of Civil Engineering, University of Sydney, Sydney, Australia
e-mail: tanvir.ahmed1@sydney.edu.au

A. El-Zein
e-mail: abbas.elzein@sydney.edu.au

F. Maggi
e-mail: federico.maggi@sydney.edu.au

K. S. K. Chung
e-mail: ken.chung@sydney.edu.au

F. N. Tonmoy
National Climate Change Adaptation Research Facility (NCCARF), Griffith University, Gold Coast, Australia
e-mail: f.tonmoy@griffith.edu.au

F. N. Tonmoy
Griffith Centre for Coastal Management (GCCM), Griffith University, Gold Coast, Australia

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_5

1 Introduction

Global flood losses are increasing due to a number of factors, including changes in climate patterns as well as growing population and economic development in flood-prone areas [1, 2]. The reported average annual global losses between 1980 and 2012 exceeded $23 billion [3]. Climate change is likely to increase weather-related hazards significantly, including flood and associated economic and social damage [4] due to changes in rainfall patterns and increases in the frequency and intensity of extreme weather events [3]. A rich literature has helped us to understand the dynamics of floods and predict future flooding scenarios [5–8]. In addition, a number of studies have proposed and/or analysed flood management strategies aimed at reducing economic damage and assisting communities in adapting to flooding events [2, 9–11]. However, most flood studies are conducted at city or catchment scales that are important and useful for municipal governments, but not sufficient. Municipalities require locally-specific data and strategies that can identify populations at risk and develop specific measures to reduce vulnerability and increase resilience. Such strategies are typically inscribed within, and informed by, larger-scale strategies but are not entirely determined by them.

In some cases, local authorities have conducted their own flood studies. These studies typically concentrate on the geophysical aspects of flooding, by determining frequency and areal extent of floods, and by providing recommendations for improving drainage paths and upgrading storm water infrastructures (e.g. [12–14]). However, different communities and individuals may be at risk for different reasons (e.g., living in low-lying areas, poor mobility, poor access to financial resources in times of flood). Hence, for effective flood risk management and for better adaptation to floods, it is important to know not only how significant the aggregate flooding risk is, but who is at risk and what are the drivers of their vulnerability [15]. While the literature on flooding has recognized the importance of incorporating institutional and socio-economic factors in determining vulnerability to flooding, conducting such assessments at local scale remains limited.

This paper develops and applies a new methodology for assessing vulnerability to flooding in urban areas at local scale. The study aims to advance our understanding of the relative importance of geophysical, institutional and socio-economic factors within local communities that make them vulnerable to flooding. A hybrid approach, combining hydrological and hydraulic analyses with social vulnerability analyses, is adopted. The methods developed here are tested and illustrated by applying them to Marrickville, a local government area in Sydney's inner-west that is prone to flooding events.

2 Risk Assessment Framework

A number of different conceptual frameworks exist for analyzing flood risk and vulnerability. One such framework—used by a number of studies in the literature and adopted in this paper—defines flood risk as a pyramid in which the impact of floods is the outcome of the interaction between *hazard*, *exposure* and *social vulnerability* [16]. Hazard is typically characterized by its probability of occurrence and most commonly represented by the areal extent and depth of flooding in time and space under different scenarios [17, 18]. "Exposure" is the extent to which valued aspects of a community's life (e.g., health, prosperity, security) are likely to be affected by the flood [19]. On the other hand, social vulnerability (SV) refers to the intrinsic characteristics of the exposed elements which determine their potential to be harmed and their capacity to cope with, and adapt to, the hazard [20]. This helps explain why two communities equally exposed to the same hazard may experience its impacts in very different ways. In our adopted framework, the above three concepts are represented by the three dimensional risk pyramid shown in Fig. 1. Increasing the magnitude of any of these dimensions increases the volume of the pyramid, which in turn reflects a higher overall risk.

3 Development of Flood Social Vulnerability Model

In order to implement the above-mentioned risk framework in the context of local scale storm water management under climate change, a conceptual model of Flood Social Vulnerability (FSV) is developed here (Fig. 2). The three dimensions of the risk pyramid are operationalized by three sets of indices: (i) Flood Hazard Index (F_H); (ii) Flood Exposure Index (F_E); and (iii) Social Vulnerability Index (SoVI). Each of these indices are implemented within our framework using a set of indicators that capture major elements of an individual index. Each individual index is calculated as a weighted average of its indicators:

Fig. 1 Risk pyramid

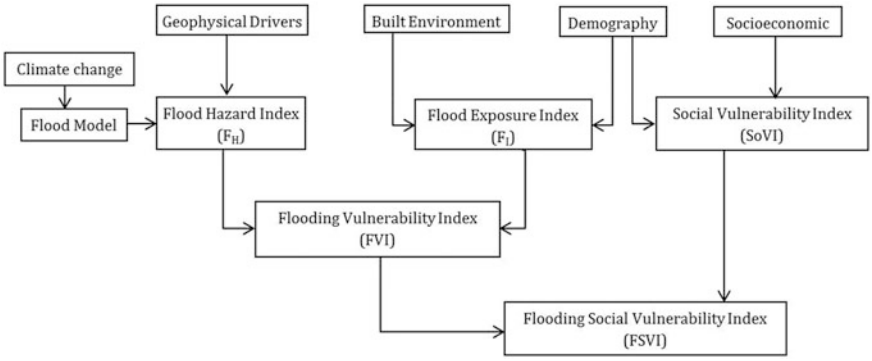


Fig. 2 Conceptual model for flood social vulnerability

$$F_m = \sum_{i=1}^{N_m} w_i I_i \tag{1}$$

Here F is the index of the dimension m (i.e., flood hazard, flood exposure, social vulnerability), w_i is the weight of indicator I_i. N_m is the number of indicators for dimension m.

The unit of analysis mostly depends on available data and can range from a single household to a large geographical area incorporating hundreds or thousands of households. F_H for a unit of analysis within a municipality is determined by certain characteristics of a flood event such as its aerial extent, flood depth, duration etc. Changes in rainfall patterns can influence these flood characteristics. On the other hand, F_E is captured by the demographic and built-environment characteristics of the unit of analysis. This can include population density, age and density of different types of built infrastructure such as residential, commercial and industrial properties, roads etc. Finally, SoVI is determined by the socioeconomic characteristics of the residents of the study unit, which may capture social capital, education, access to resources and technology, etc. A large number of socioeconomic indicators are available, that can be used to quantify social vulnerability. However often these indicators are highly correlated which can reduce the validity of the index since additive aggregation generally requires indicators to be algebraically independent of one another [21]. Therefore, a principal component analysis (PCA) is conducted to reduce an large initial set of selected social vulnerability indicators into a small set of principal components [22, 23]. SoVI is calculated by adding the principal components using equal weights. Finally, the three indices can be combined to calculate a Flood Social Vulnerability Index (FSVI) for each unit of analysis (Eq. 2).

$$FSVI_i = F_{Hi} \times F_{Ei} \times SoVI_i \quad (2)$$

The advantages of the multiplicative approach used in calculating the FSVI is that it magnifies cases in which more than one index is pointing to high vulnerability and yields a small value in cases in which any one of the indices indicates very low vulnerability.

4 Application of FSV in Marrickville Council

In order to test the applicability of FSVI, we applied the methodology to the suburb of Marrickville. Marrickville is a local government area (LGA) in Sydney's inner-west, which has experienced significant flooding as recently as 2012 [24]. It has medium-density residential and light-density industrial developments, as well as a number of major and minor roads. The storm water system of Marrickville valley sub-catchment drains its water through a curb/gutter system, to a pipe system, and finally into four major outfalls, including a tunnel, into the Cooks River. The Marrickville valley is divided into seven sub-catchments and the total drainage system consists of open channel and underground pipe systems [12]. According to the 2011 census data the total population of Marrickville LGA was 76,500, 65% of whom were engaged in full time employment with a median weekly household income of $1605. In addition, 34.6% of the population were enrolled in, or attended, tertiary education.

The unit of analysis for this study was adopted as Statistical Area Level 1 (SA1), which is designed by the Australian Bureau of Statistics (ABS) as the smallest unit for publicly available census data [25] typically including 200–800 residents. Table 1 shows the indicators used in the study. In order to generate indicators for F_H, an existing flood model was collected from the Inner West Council, which is the municipal authority responsible for the Marrickville area. Flood events were then simulated by means of a hydrological and a hydraulic modelling technique using TUFLOW and Drains software, respectively. The hydraulic model calculates flood levels and flow patterns, and simulates the complex effects of backwater, overtopping of embankments, waterway confluences, bridge constrictions and the effect of other hydraulic structures. Our analysis did not include the east-end catchment because data representing it were not available. The model was calibrated in a previous study against qualitative flooding hot spots, community questionnaire results and comparisons with earlier studies. No calibration based on the magnitude of past flood events was possible because of the lack of stream gauges in the catchment area. The model was used to generate a flood hazard index for an event equivalent to 100 year Annual Recurrence Interval (ARI) 2-h rainfall. In estimating F_H, the average flood depth was used (rather than the maximum depth) because of the large spatial variation of flood depth. The two indicators of F_H were normalized to a scale between 1 and 10. Figure 3a shows the spatial distribution of F_H.

Table 1 Indicators used for developing FSV for Marrickville council

Model	Indicators	Unit
Flood hazard	(i) Areal extent of flooding	As a % of total area
	(ii) Average flood depth	meter
Flood exposure	(i) Population density	count/km^2
	(ii) Building density	
	(iii) Old building density	
Social vulnerability[a]	(i) Number of children below 5 years of age	As a % of total population
	(ii) Number of people over 65 years of age	
	(iii) Number of people between 35 and 39 years of age	
	(iv) Number of people requiring special assistance	
	(v) Number of females	
	(vi) Number of non-citizens	
	(vii) Number of people moved residence in the last year	
	(viii) Number of people with low speaking proficiency in English	
	(ix) Number of people with weekly income less than $300	As a % of total population aged 15 years and over
	(x) Number of people with weekly negative or nil income	
	(xi) Number of people who never go to school	
	(xii) Number of single parent families with children under 15	As a % of total families
	(xiii) Number of couple families with more than 2 dependent children	
	(xiv) Number of unemployed families	
	(xv) Number of dwellings with no internet connection	As a % of total families
	(xvi) Number of dwellings with no motor vehicle	
	(xvii) Number of dwellings occupied by single individuals	
	(xviii) Number of dwellings occupied by 5 individuals or more	
	(xix) Number of dwellings owned with mortgage	
	(xx) Number of dwellings with median household income	

[a]This table shows 20 out of 31 indicators that satisfied the criteria for PCA

The National Exposure Information System (NEXIS) designed by Geoscience Australia provides aggregated information about residential, commercial and industrial structures. The aggregate information of population density, building density and old building density was extracted from the NEXIS database and used to generate an exposure index F_E. Figure 3b shows the spatial distribution of F_E. Socioeconomic data given by 31 indicators was collected from ABS (2011) and was next analysed with PCA to calculate SoVI for each SA1 (Fig. 3c). After a number of PCA iterations, 31 indicators were reduced to 20 that satisfied a number of criteria (e.g. multicollinearity, Kaiser-Meyer-Olkin measures, Bartlett's test of sphericity). A total of 5 components (each with 3–6 indicators) were produced, which explained 68% of all data variance. Components scores were then added using equal weights to determine SoVI. Next, F_H, F_E and SoVI were normalized between 1 and 10, where hazard, exposure and social vulnerability all increase with the increase of the index. Finally, F_H, F_E and SoVI were combined using Eq. 2 to calculate FSVI (Fig. 3d).

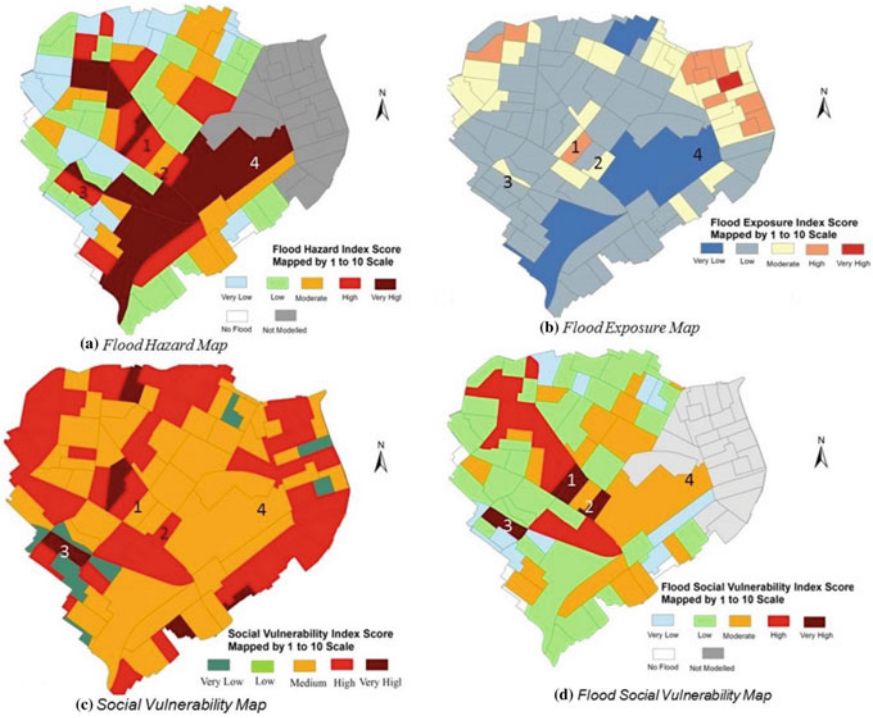


Fig. 3 Vulnerability profile of each study unit

5 Discussion

Units of analysis marked as 1, 2 and 3 in Fig. 3 were found to have the highest FSVI. High flood hazard, high exposure and medium SoVI resulted in a very high FSV for unit 1. This SA1 is located in the lower-elevation area of the Marrickville valley catchment, thus leading to high water depth and extent of flooded area (37% area flooded). This area also included a large number of residential buildings which increased its exposure to flooding. In the case of unit 2, high FSV was caused by the interaction of high flood hazard and moderate exposure index and SoVI. In unit 3, though the exposure index was low, the hazard index and SoVI were high. This SA1 was characterised by a relatively large percentage of aged and disabled population and low average income. On the other hand, despite its very high flood hazard, unit 4 had a medium FSV, mainly because of very low exposure (low population density and relatively newer houses).

All impacts of flooding are experienced locally and preparedness requires detailed information at a local scale that can help local government planners to develop specific, differentiated mitigation plans, rather than simply adopting flood mitigation measures based on larger-scale studies. The method proposed here can be particularly useful if the information it generates is used by local government to identify sections of the urban drainage system that need increase in capacity or to conduct more focused qualitative and quantitative assessments of drivers of vulnerability in areas identified as highly vulnerable.

In terms of future work, the effectiveness of the existing storm water infrastructure under future climate change will be tested where F_H will be developed by re-running the flood models for different climate change scenarios using The NSW and ACT Regional Climate Modelling (NARCliM) projection scenarios.

Acknowledgements This work was financed by a grant from the Civil Engineering Research Development Scheme (CERDS) at The University of Sydney, 2015–2016. The authors are grateful to the Inner-West Council, NSW, for its generous support, including discussion time, feedback, models and flood data, and to Watercom and BMT-WBM for providing access to modelling software.

References

1. Ashley RM, Balmforth DJ, Saul AJ, Blanskby JD (2005) Flooding in the future—predicting climate change, risks and responses in urban areas. Water Sci Technol 52(5):265–273
2. Muis S, Guneralp B, Jongman B, Aerts JCJH Ward PJ (2015) Flood risk and adaptation strategies under climate change and urban expansion: a probabilistic analysis using global data 538:445–457
3. Jongman B, Winsemius HC, Aerts JCJH, Perez EC, Aalst MK (2015) Declining vulnerability to river floods and the global benefits of adaptation. Proc Nat Acad Sci 112(18)
4. Muller M (2007) Adapting to climate change: water management for urban resilience. Int Inst Environ Dev (IIED) 19(1):99–113

5. Domingo NS, Refsgaard A, Mark O, Paludan B (2010) Flood analysis in mixed-urban areas reflecting interactions with the complete water cycle through coupled hydrologic-hydraulic modelling. Water Sci Technol 62(6):1386–1392
6. Hallegatte S, Green C, Nicholls RJ, Corfee-Morlot J (2013) Future flood losses in major coastal cities. Nat Clim Change 3(9):802–806
7. Patro S, Chatterjee C, Mohanty S, Singh R, Raghuwanshi NS (2009) Flood inundation modeling using MIKE FLOOD and remote sensing data. J Indian Soc Remote Sens 37(1):107–118
8. Zhou Q (2014) A review of sustainable urban drainage systems considering the climate change and urbanization impacts. Water 6(4):976–992
9. Garbutt K, Ellul C, Fujiyama T (2015) Mapping social vulnerability to flood hazard in Norfolk, England. Environ Hazards 14(2):156–186
10. Liu J, Hertel TW, Diffenbaugh NS, Delgado MS, Ashfaq M (2015) Future property damage from flooding: sensitivities to economy and climate change. Clim Change 132(4):741–749
11. Tavares AO, dos Santos PP, Freire P, Fortunato AB, Rilo A, Sá L (2015) Flooding hazard in the Tagus estuarine area: the challenge of scale in vulnerability assessments. Environ Sci Policy
12. Grey S (2011) Marrickville valley flood study. Project number: 110004, Client: Marrickville council, final draft report
13. Cornelius S, Mirfenderesk H, Chong E (2012) Experience and techniques in modelling urban stormwater networks and overland flow paths. Proceedings of Flood plain conference, New Castle, Australia
14. Reid A, Thomson R (2014) Floodplain risk management plan, Alexandra canal floodplain risk management study and plan, Job reference: w4948, prepared for City of Sydney
15. Koks EE, Jongman B, Husby TG, Botzen WJW (2014) Combining hazard, exposure and social vulnerability to provide lessons for flood risk management. Environ Sci Policy 47:47–52
16. Dwyer A, Zoppou C, Nielsen O, Day S, Roberts S (2004) Quantifying social vulnerability: a methodology for identifying those at risk to natural hazards. Geoscience Australia
17. Plate EJ (2002) Flood risk and flood management. J Hydrol 267(1):2–11
18. Schanze J (2006) Flood risk management—a basic framework. In: Flood risk management: hazards, vulnerability and mitigation measures. Springer, Netherlands, pp 1–20
19. Tonmoy FN, El-Zein A, Hinkel J (2014) Assessment of vulnerability to climate change using indicators: a meta-analysis of the literature. Wiley Interdisc Rev Clim Change 5(6):775–792
20. Cutter S, Emrich C, Morath D, Dunning C (2013) Integrating social vulnerability into federal flood risk management planning. J Flood Risk Manag 6:332–344
21. El-Zein A, Tonmoy FN (2015) Assessment of vulnerability to climate change using a multi-criteria outranking approach with application to heat stress in Sydney. Ecol Ind 48:207–217
22. Fekete A (2009) Validation of a social vulnerability index in context to river-floods in Germany. Nat Hazards Earth Syst Sci 9(2):393–403
23. Field A (2009) Discovering statistics using SPSS: (and sex and drugs and rock 'n'roll). Introducing statistical methods
24. Mckenny L, Saulwick J, Tovey J (2012) After the flood: Sydney counts the cost of record rains. The Sydney Morning Herald Environment, Fairfax Media. Web. 30 May 2016
25. ASGS (2010) Australian Statistical Geography Standard: Volume 1—main structure and greater capital city statistical areas, July 2011. Retrieved 1 June 2016 from http://www.abs.gov.au/ausstats/abs@.nsf/0/7CAFD05E79EB6F81CA257801000C64CD?opendocument

Enablers and Barriers of Smart Data-Based Asset Management Services in Industrial Business Networks

Toni Ahonen, Jyri Hanski, Matti Hyvärinen, Helena Kortelainen, Teuvo Uusitalo, Henri Vainio, Susanna Kunttu and Kari Koskinen

Abstract Recent academic research has paid particular attention to how digitalization disrupts current business models and business environments. Furthermore, servitization has gained significant attention. However, so far only a fraction of the wide range of opportunities related to digitalization has been realized. In this paper we aim to better understand the drivers, limitations and stakeholder expectations in different industrial business environments. In the proposed paper, we address digitalization in the area of engineering asset management from the following perspectives: (1) enablers and barriers of digitalized asset management service business, (2) availability and use of data for decision-making support, and (3) changes for business models. We also further contemplate which decision-making situations need to be supported by digital asset services. The paper is based on data received from a company workshop and a literature review.

1 Introduction

Digitalization can be defined as the use of digital technologies to change a business model and provide new revenue and value-producing opportunities; it is the process of moving to a digital business [3]. Digitalisation has already transformed many sectors of industry and society, for instance media, banking and communications. However, in the future, it is expected to transform the industries and society even further [8, 15]. Opportunities for asset management services offered by digitalization are discussed extensively in recent research publications (e.g. [9, 13, 15, 16]). In practice, for instance, digitalization enables better management of life cycle information and better availability of sensor, equipment and process information for

T. Ahonen (✉) · J. Hanski · H. Kortelainen · T. Uusitalo · S. Kunttu
VTT Technical Research Centre of Finland, Ltd., Tampere, Finland
e-mail: toni.ahonen@vtt.fi

M. Hyvärinen · H. Vainio (✉) · K. Koskinen
Tampere University of Technology, Tampere, Finland
e-mail: henri.vainio@tut.fi

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_6

different stakeholders. Therefore, new value potential for asset management services could be in real-time optimisation and predictive maintenance.

An increasing number of product and service providers aim to gain a competitive advantage through new smart data-based asset management services. It is acknowledged that the business potential is significant [7] and a change is called for, however, Porter and Heppelmann [10] state that it is a dangerous oversimplification to suggest that the Internet of Things "changes everything". They note that the rules of competition and competitive advantage still apply. Companies also need to have a clear strategy for how to use data and analytics to compete, and how to deploy the right technology architecture and capabilities such as analytics tools [1]. Thus, companies need to find their own perspective for the Industrial Internet and develop solutions that fit the concrete needs of their customers. Further information is required to understand the enablers and barriers related to different types of data-based asset management services.

In this paper, we address digitalization in the area of engineering asset management from the perspectives of enablers and barriers of digitalized asset management services. The paper is based on data received from a company workshop and a literature review. From the managerial perspective, this paper provides decision support for the development of asset management services by identifying different types of services and analysing the enablers and barriers for the different types. From the scientific perspective, this paper analyses the effects of digitalization on the asset management, a classification of smart asset management services and potential enablers and barriers for the development and delivery of asset management services.

2 Smart Data-Based Asset Management Services

Manufacturing companies are increasingly investing in developing industrial services and seeking new opportunities in a more customer value-oriented way. However, companies are in different positions and levels of preparedness with respect to the digitalisation and adoption of new technologies in asset management service business. Early adopters have started creating solutions for predictive maintenance and remote asset management, and new means for improving worker productivity, safety and working conditions. Maintenance services are a typical example of digitalized asset management services. One example of maintenance service is 'SKF Rotation For Life' provided by SKF, a global supplier of bearings, spindles and seals (SKF homepage). The service combines bearing technology, failure detectability and reliability services into an integrated package. Digital solutions can provide tools and methods to collect and analyze data to detect emerging failures in early phase and to give case-relevant repair instructions to field though e.g. mobile solutions. This allows for doing the right actions, on the right time, and correctly.

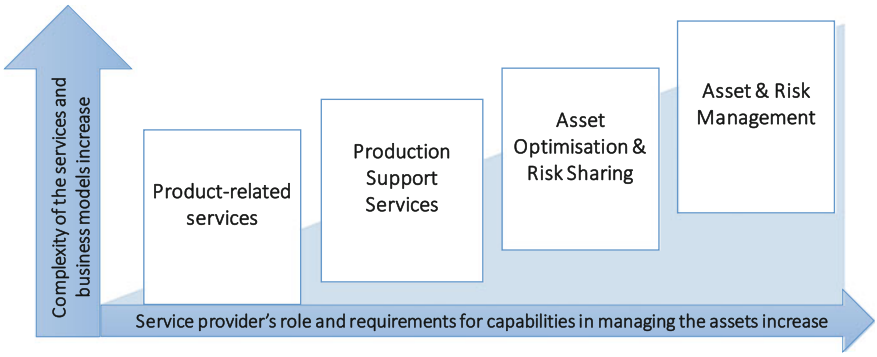


Fig. 1 Classification of smart asset management services

Another example of digitalized services is CIP (clean in place) wash optimization service, offered by a company called TTS-Ciptec (TTS-Ciptec homepages). The optimization of the CIP washes are conducted through remote monitoring of washing liquids and predictive analytics services related to these monitoring results. In this case, digitalized solutions reduce CIP times and so the CIP wash can be conducted much faster. CIP optimization includes also viewpoints from the resources savings, in terms of water, energy and chemical consumptions. Figure 1 presents a stepwise approach for holistic asset management services, with four service types; namely product-related services, production support services, asset optimization and risk sharing, and asset and risk management.

Product-related services exploit the sensor and ICT technologies to digitalize the current offering, including e.g. preventive maintenance services and spare parts management. Production support services cover the services with digital tools for preventive maintenance, failure diagnostics and information-based overall equipment efficiency optimization integrated with expert consultation. The CIP wash optimization service described above is an example of a service of this category. Asset optimisation and risk sharing includes utilisation of wide installed base data to support operation and maintenance, using methods such as benchmarking and predictive maintenance. The SKF Rotation for Life service is an example of a service of this category. Asset and risk management integrates holistic digital services and intelligent products, resulting in highly automated decision-making.

3 Methodology

The study follows hermeneutic worldview and qualitative research methodology. The goal of the research is to understand the barriers and enablers for digital transformation in the asset management services and construct new hypotheses based on the findings. The data for was collected by a literature review regarding

enablers and barriers for data-based asset management services and in a workshop using world café method.

Two workshop groups consisting of altogether 6 researchers from VTT and TUT and 8 company representatives from five companies discussed digitalization opportunities, barriers and enablers. Selective sampling was utilised in the study. All the participants are involved in a national research project aiming at supporting the development of new digital asset management services. All the participants represent small or medium sized companies that are currently actively involved in developing or offering digital asset management services. Researchers, who also wrote down all the ideas, facilitated the discussions. The generated ideas were discussed and enriched by the other group in an iterative process. At the end of the world café, the facilitators presented the opportunities, enablers and barriers and the workshop participants commented and further enriched them.

In the first phase of analysis, the data from the literature review and the workshop were analysed separately. In the second phase, the data from the different sources were compared. Identified enablers and barriers were divided using the framework that includes three crucial stages for service development; business models, decision-making support and data analytics. In the final stage of analysis, the identified key barriers and enablers were analysed through the lenses of the asset management services types presented in Fig. 1.

4 Results

We apply the following three areas as the parts of the framework for our analysis in the next section: (1) Business models, (2) Decision-making support, and (3) Data analytics. These areas together compose the path from data to development of profitable business. Evans and Annunziata [2] have stated that Industrial internet is a combination of the following key elements: intelligent machines, advanced analytics and people at work. Here, we address intelligent machines and advanced analytics together and address the people at work aspect by particularly considering the decision-making support for different processes, actions and decision-making situations at different levels of work. In addition, we consider the business model level as an integrating entity in which significant changes are expected.

4.1 Enablers and Barriers of Digital Asset Management Services

The recent literature discusses the barriers for digital transformation [5, 12, 13]. In addition, barriers for digital asset management services were identified in a

workshop setting described in Chapter "Semiparametric Valuation of Heterogeneous Assets".

Table 1 introduces some of the barriers for digital asset management.

Table 1 Barriers for digital asset management services

	Barriers identified from the literature review	Barriers identified in the company workshop
Business model	– No strong business case – Investment in current infrastructure – Industry specific financial KPI's – Industry beliefs – Legacy competences, processes and systems and equipment – Return on continuous improvement – Data security risks and ownership – Incumbents active resistance – Lack of trust in virtual environments – High barrier industry – Rivalry between existing competitors – Cost driven customers – High entry barrier industry – Long established stakeholder organizations – Oligopolistic market – Heightened complexity of the innovation process	– Selling value instead of a product challenging – Customers' focus too much on the acquisition price instead of asset lifecycle value – New collaboration models and value creation of services not easily understandable for customers, in particular procurement organisations – Earning logic: if service is given free of charge to customers, customers might not see the value – Identifying the specific digitalization opportunities – Technology-driven development efforts of companies – Companies afraid of cannibalizing their current business – Inadequate communication of the examples, benefits, opportunities and the risks related to digitalization – Customers do not recognise the potential for digital services when they see the supplying company only as a machine supplier – Risk of losing own analytics competences, when using external machine learning systems – Marketing digital services is underdeveloped, need for clear messages and stories
Decision-making	– Insufficient technical skills – Lack of interoperability of existing systems – Uncertain risks brought by investment in immature or untested technologies – Too many competing priorities – Lack of an overall strategy – Security concerns and data privacy – Lack of organizational agility	– Stepwise change, no time for incremental learning by doing – New, diverse know-how and skills are needed for the new digital services, but the current employees may not necessarily have it – Management lacks understanding of the nature of change in skills and competencies – Lack of change leaders in companies – Fear of losing job

(continued)

Table 1 (continued)

	Barriers identified from the literature review	Barriers identified in the company workshop
	– Lack of entrepreneurial spirit, willingness to take risks – Lack of collaborative, sharing culture – Lack of employee incentives – Cultural barriers – Legacy leadership and management (leadership from the past) – Lack of absorptive capacity – Next quarter pressure – Well established dominant logic – Complex organisations – Incrementalism – Hierarchical organisations – Impatience – Inexperienced management – Missing financial track record	– Internal opposition for change – Making right technology and platform choices – Maintenance and version management problems due to tailoring solutions for customer needs
Data analytics	– Lack of relevant history data – Data security – Data ownership confusion and conflict of interest – Management beliefs – Lack of competences – Privacy concerns – Lack of systemic understanding – Cost and availability of data transmission – Non sensored installed fleet – Lack of data governance rules	– Lack of relevant data and information for specific situations – Access to customer's system is not always easy – Lack of integration of data collected from different sources – Insufficient knowhow to analyse the available data efficiently – Environmental features and insufficient sensor technology cause challenges for data acquisition

Enablers for digital transformation based on World Economic Forum [15] and Sommarberg [12] and the company workshop are depicted in Table 2.

4.2 Analysis

The division of the identified enablers and barriers into the three categories was challenging as some of the enablers and barriers could be placed into two or even three categories. The main barriers identified by the workshop participants are related to value of services, business perspective and customer needs, readiness for

Table 2 Enablers for digital asset management services

	Enablers identified from the literature review	Enablers identified in the company workshop
Business model	– Digital entrants to traditional industries – Substitutive business models – Value driven customers and suppliers – Influential early adopters – Suppliers passing middle-men – Platform supply – Business model diffusion – Consumer business behaviour into B2B – Software intensity – Access to crowd knowledge – New level of customer insight – Powerful development tools – Quantum leap in productivity – Real time opportunity – Asset light – Slow growth of existing business – Start-up emergence – Crowdfunding – Common approach to addressing security concerns – The convergence of standards to support better interoperability – Collaboration to create technology testbeds, such as for testing technology interoperability – Bandwagon and Hype – Degree of external engagement (crowds and users)	– Partnerships in development of services, products and competences – The pilots and development projects should be at sites where the fluent communication is guaranteed and the needs can therefore be understood well – A combination of short and long term objectives for the development guarantees both the adequate ambition level of the whole efforts but also practical results to justify the business cases – Marketing new digital services to existing customers – The pioneers act as encouraging examples since when a larger company takes the first steps towards digital services first, it makes easier for the smaller players to follow – Emphasising the business perspective and customer need in the development of digital service
Decision-making	– Management consulting focus – Management rotation – Emerging digital strategy frameworks – Experiment culture	– Transparency and commitment – Internal information sharing and marketing, creating a clear message and narrative
Data analytics	– General digital technologies – Reducing human errors – Scalability of knowledge – Analysis speed – Advances in cognitive computing – Users opening their data – Platform tools	– Agile testing to complement long-term development – Education for the whole organisation – Cloud based solutions – Companies should offer frameworks/platforms where digital services can be easily used

transformation, supporting technologies, and competencies in service delivery, risk of losing competences, access to information required for services, marketing and communication and R&D. The enablers identified by the participants are related to partnerships, R&D, marketing and communication, supporting technologies and competencies, references, business perspective and customer need.

Main barriers related to business models identified in both literature and by the workshop participants are cost-driven customers, selling value instead of product, and earning logic (lack of strong business case). Common barriers in decision-making include lack of new diverse know-how, skills and leadership. Regarding the data analytics, common barriers include lack of relevant data, lack of capabilities to produce information for decision-making (e.g. lack of competencies in analytics), and data access and ownership issues. Business model related enablers identified in the literature and by the workshop participants are partnerships (collaboration and external engagement), testbeds and pilots, customer need emphasis (new level of customer insight) and influential early adopters (pioneers). Common decision-making related enablers are not identified. Common enablers related to analytics include utilisation of new digital technologies, such as cloud and platform solutions.

5 Discussion

Figure 1 presented a classification of smart asset management services. It suggests that, when moving towards asset and productivity management, service providers are expected to acquire more capabilities to manage complexity and take a significant role in creating valuable new knowledge to manage customer's asset in a holistic manner. The emphasis and role of different enabling and hindering factors changes according to the service types. There are enablers and barriers for digitalised asset management that have been emphasised in both literature and in our research. In Fig. 2, we present these most essential enablers and barriers identified for digitalised asset management. We address them with respect to four asset management service types presented in Sect. 2 and present how the enablers and barriers are differently emphasised in different services.

It is identified that the transformation from product-related services towards more holistic service types calls for new sort of capabilities. Thus, more collaborative relationships and continuous communication are required. Additionally, all the relevant stakeholders should agree on data governance rules and practices. Furthermore, transparency and traceability of data usage may provide an answer to security and ownership related barriers. Increased capability to produce data for specific situations and utilise the available data efficiently are also important. Technological capabilities can be seen relevant for all types of asset management services. However, issues related to customer and value understanding, capabilities to integrate to customers' decision-making processes and collaborative business models are more emphasised in the more holistic service types.

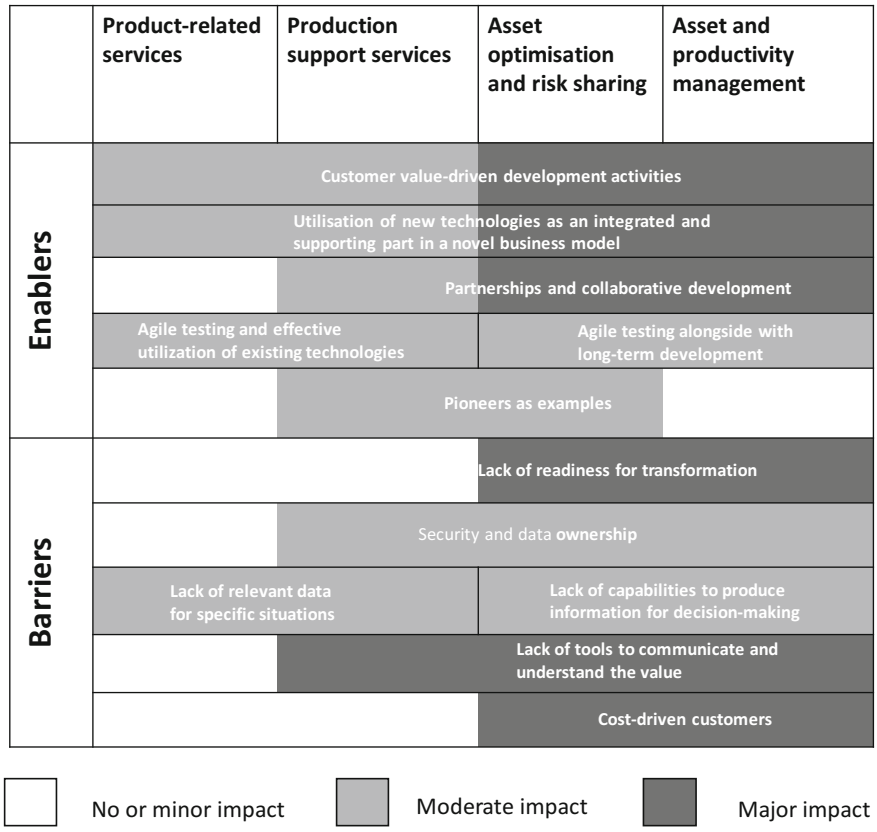


Fig. 2 The most significant enablers and barriers for asset management service types

6 Conclusions

Current paper proposes a framework for the classification of asset management services. Furthermore, the paper discusses the enablers and barriers of smart asset management services with a service classification as a framework, which seems rather little addressed in the scientific debate. Our results suggest that to answer to the challenges and opportunities of digitalization, companies need new capabilities at business model, technology, management and process and method levels. Significant part of the emerging industrial services rely on multi-source data. Large amounts of data are already collected, however, in many occasions only a fraction of the data is effectively exploited. Therefore, the utilisation of the available data and data analytics is an area where more know-how and competences would be needed. However, even more importantly, understanding of the customer and phenomena related to production and business environment will be a prerequisite for developing services that meet the customer needs. As the managerial

implication, the service classification framework and the enablers and barriers identified provides supporting knowledge for practitioners to plan their own route to digitalized services. We also conclude that the enablers and barriers specifically related to how data are converted into customer value need to be studied further.

Acknowledgements The authors gratefully acknowledge the Finnish Funding Agency for Technology and Innovation for funding the research project and the companies involved in the research. We also wish to thank the colleagues from VTT and Tampere University of Technology for their valuable comments.

References

1. Barton D, Court D (2012) Making advanced analytics work for you—a practical guide to capitalizing on big data. Spotlight on big data. Harvard Bus Rev, October 2012
2. Evans and Annunziata (2012) Industrial internet: pushing the boundaries of minds and machines. Gen Electr. Available at: https://www.ge.com/docs/chapters/Industrial_Internet.pdf.(02/2017)
3. Gartner (2016) IT Glossary. Available at: http://www.gartner.com/it-glossary/digitalization/
4. ISO 55000 (2014) Asset management. Overview, principles and terminology
5. Kane GC, Palmer D, Phillips AN, Kiron D, Buckley N (2015) Strategy, not technology, drives digital transformation. Becoming a digitally mature enterprise. MIT Sloan Management Review and Deloitte University Press, July 2015
6. Liyanage JP (2012) Smart engineering assets through strategic integration: seeing beyond the convention. 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, Netherlands
7. Manyika J, Chui M, Bughin J, Dobbs R, Bisson P, Marrs A (2013) Disruptive technologies: advances that will transform life, business, and the global economy. McKinsey & Company, McKinsey Global Institute, May 2013
8. OECD (2015) OECD digital economy outlook. OECD Publishing, Paris
9. Popescu GH (2015) The economic value of the Industrial Internet of things. J Self-Gov Manage Econ 3(2):86–91. ISSN 2329-4175
10. Porter ME, Heppelmann JE (2014) How smart, connected products are transforming competition. Harvard Bus Rev
11. SKF homepage. http://www.skf.com/group/services/service-contracts/rotation-for-life/index.html
12. Sommarberg M (2016) Digitalization as a paradigm changer in machine-building industry (vol 1436). Tampere University of Technology
13. Tihinen M, Kääriäinen J (eds) (2016) The Industrial Internet in Finland: on route to success? vol 278. VTT Technology, 84 p. Available in: http://www.vtt.fi/inf/pdf/technology/2016/T278.pdf. Accessed 7.2.2017
14. TTS Ciptec homepage. http://www.tts-ciptec.com/
15. World Economic Forum (2015) Industry agenda. Industrial internet of things. Unleashing the potential of connected products and services, January 2015, 40 p
16. Yoo Y, Richard J, Boland J, Lyytinen K, Majchrzak A (2012) Organizing for innovation in the digitized world. Organ Sci 23:1398–1408

Assessment of the Impact of Maintenance Integration Within a Plant Using MFD: A Case Study

Hatem Algabroun, Basim Al-Najjar and Anders Ingwald

Abstract In the recent decades the recognition of maintenance as an effective part of the company competitiveness has grown. In order to enhance maintenance performance and its positive impact, integration with the rest of the plant activities should be planned and performed. In this study, Maintenance Function Deployment model was applied on a real data of a case company. The model was used to analyse the company's production with respect to the business and economic variables. This was done first through finding loss causes of the strategic goals of the case company, then breaking down these causes and their costs into their root causes. Based on economic estimations, the results show that the major root causes behind losses are insufficient training of personnel and lack of maintenance integration. It is concluded that properly considering the integration of maintenance within the company's activities, reduces loss and improves the company's performance. Also, that applying MFD eases the identification of the root-causes behind losses as well as quantifying and prioritizing the economic losses.

1 Introduction

Nowadays, the challenge for any industrial company to survive in the market is to generate sufficient profit. Different activities such as, maintenance, quality, operation, etc., in a company are essential to achieve the company's goals. Changing in any of these working areas could influence other related areas due to their internal interactions. Thus, these activities should be integrated and synchronized to fulfil

H. Algabroun (✉) · B. Al-Najjar
Department of Mechanical Engineering, Linnaeus University, Växjö, Sweden
e-mail: hatem.algabroun@lnu.se

B. Al-Najjar
e-mail: basim.al-najjar@lnu.se

A. Ingwald
Department of Physics and Electrical Engineering, Linnaeus University, Växjö, Sweden
e-mail: anders.ingwald@lnu.se

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_7

company goals [1]. In general, maintenance activities are responsible to sustain the production system. But it also has a high influence on the general performance of the company, due to its impact on different areas, such as quality, production, safety, etc. Thus, a proper maintenance not only sustains the productivity, but it also improves the overall performance of the company [2]. Therefore, maintenance should be utilized effectively to its full extent. In order to enhance maintenance performance and its positive impact, integration with the relevant activities should be planned and performed. However, it is not easy to make an assessment of the impact of maintenance integration due to lack of relevant information within companies. This paper aims to apply Maintenance Function Deployment (MFD) model on a case company and show a quantitative assessment of the impact of maintenance integration. In the next section we discuss maintenance integration in companies and its impact. Section 3 describes the MFD model. In Sect. 4 the MFD is applied on a case company, and then the results are discussed in Sect. 5. Finally, Sect. 6 draws the conclusion of the study.

2 On the Integration and Its Impact on Maintenance Performance

Several researchers discussed the interdepartmental integration within organizations [1, 3, 4]. In this study we refer to the integration as the process of interaction and collaboration among working areas that unite the organization to achieve its goals [4]. The interaction is tangible and can be monitored easily; it represents the communication and the flow of information. It comes in several forms, e.g., meetings, emails, documents, etc. On the other hand, the collaboration is intangible. It shows the ability of departments to work together and it comes in the form of having mutual understanding, common vision, joint goals and sharing resources. However, the integration should be built in an effective way, e.g. too much interaction could prevent an effective relationship by burdening personnel with documentation efforts and information overload. In other word the integration should be built properly with consideration to the situation, as some tasks require high interaction and collaboration between departments and other tasks should be handled internally in the department [4].

Generally, failures in the operative level impact negatively on achieving these goals [5]. The role of maintenance activities is to reduce the probability of failure and unplanned stoppages in order to minimize the impact of failures on the companies' profitability and strategic goals. In general, failures in production process could occur due to technical issues or other elements involved in the production process, for example, human resources or systems' quality [2, 6, 7]. Maintenance integration with relevant working areas helps maintenance to obtain the relevant information. This will assist maintenance to detect, diagnose and prognose potential failures and define their root causes in order to make cost effective decisions.

Therefore, maintenance integration is a key factor in enhancing the maintenance outcomes and to achieve companies' goals.

3 Maintenance Function Deployment (MFD)

The MFD model is developed by Al-Najjar [8], using an equivalent structure that is used by the Quality Function Deployment model (QFD). QFD is a tool that is used to transfer customer demands, first into technical attributes, then from technical attributes into parts characteristics, then into (production) processes and lastly into instructions for production [9]. However, QFD does not identify system weakness or potential problems. MFD could be considered as a complementary tool to enhance QFD [8]. MFD aims to identify and analyze problems, assess losses, prioritize problems and causes behind losses in different working areas. MFD breaks down the economic losses in a backwards direction to reach the root-causes behind. These losses are usually spread in the operative level of different disciplines in the production process. MFD consists of four phases. Each phase contains a matrix of specific objectives i.e., What's (W) and attributes i.e., How's (H) to maintains these What's.

In the matrix, see Table 1, the value V_{mn} describes the share value of H_n (How's) (n = 1, 2, 3, …, j) in causing the loses of W_m (What's) (m = 1, 2, 3, …, i). The V_{mn} can be estimated by experts if a real value is not available. Then applying a summation of V_{mn} for each What's and How's to get a collective value for the What's importance (WT_m) and How's importance (HT_n), see Eqs. (1) and (2).

$$WT_m = \sum_{n=1}^{j} V_{mn} \tag{1}$$

Table 1 MFD matrix

		How's (H) i.e. How to achieve the objectives				Importance of What's (WT_m %)	Priority list (PL)
		H_1	… H_n	…	H_j		
What's (W) i.e. What are the objectives to be achieved	W_1	V_{11}	V_{1n}	…	V_{1j}		
	W_m	V_{m1}	V_{mn}	…	V_{mj}		
	W_i	V_{i1}	V_{in}	…	V_{ij}		
Importance of How's (HT_n %)							

$$HT_n = \sum_{m=1}^{i} Vmn \tag{2}$$

Priority list (PL) then can be formed by ranking the What's importance as follow: $P_{1st} > P_{2nd} > P_{3rd} > \cdots > P_i$. Where P_{1st} is the most important What's (the highest *WT*) and P_i is the least important (the lowest *WT*). Each How's from each phase then is used to form the new What's of the next phase. The priority of each What's and the importance of each How's will be the outcome of each phase. The four phases summarized below as follow:

I. Identifying the requirements to be maintained in order to achieve the strategic goals of the company i.e. How's. Then the share of the total losses in each strategic goal i.e. What's, that are distributed among the How's are assessed.
II. Identifying the needed tools in order to maintain the requirements in phase I. Each How's from step I will form the new What's and its total loss value is distributed among the new How's.
III. Identifying the needed activities in order to maintain and utilize the tools in phase II to the best of their extents. Each How's from step II will form the new What's and its total losses value is distributed among the new How's.
IV. Identifying the needed relevant factors to be implemented and considered in order to maintain the activities and support the integration of maintenance with plant business. Each How's from step III will form the new What's and its total loss value is distributed among the new How's.

4 Case Study

For simplicity, only the working areas of maintenance, operation and quality are considered. A real data from Auto CNC-Bearbetning i Emmaboda AB was collected based on semi-structured interview and then tested on the MFD model. The company is selected based on: (1) being an industrial company that produces countable items, and (2) being able to provide the required data. The selected company is located in Sweden and specialized in producing small mechanical parts for industrial products. The company has about 25 production stations. For simplicity, one production station is selected and one product, "sleeve", which is a component for water pumps. The product is one of many other products. It was selected based on the criteria of being a relatively expensive and produced in a relatively large quantity. The production process starts with a one-time set up of tools and different variables, e.g. speed and air pressure. Then the production process is as follow: (1) Raw material (3-m bar length) is fed into the bar feeder of the production station. (2) Shaping the first edge then cutting and shaping the second edge. (3) Cleaning. (4) Output and quality control. The roles of

maintenance, production and quality in the production process are presented in Table 2. The flow of information as well as the method of information exchange is shown in Table 3.

In order to apply the MFD, first, basing on the discussion with the company, the strategic goals were identified and they were considered to be: high product quality, on time delivery, competitive price, low violation environment, society acceptance and maintained assets condition. Then, the total losses in each strategic goal were calculated by subtracting the ideal production situation for the production line under

Table 2 The role of operation, quality and maintenance in each process

Production process	Operation	Quality	Maintenance
One-time setting up of the production station	Adjust the parameters e.g. speed	–	Check machines condition e.g. Grease
Raw material fed into the bar feeder of the production station	Feeding the bar feeder with the raw material	Check the dimension of the raw material	Continuous rotations are performed on all the production lines in order to check if there are any sign of problem during the production process
Shaping the first edge then cutting and shaping the second	–	–	
Cleaning	–	–	
Output and quality control	Organizing the output items	Quality check e.g. any burnt spot	

Table 3 The role of operation, quality and maintenance in each process

Information flow and direction	Information	The way of information transaction
Quality to operation	QC plan demanded from customer e.g. dimension	Production order sheet from computerized system
Operation to quality	Taken production time, faced problems, comments and suggestion for changes	Sheets or verbal communication
Operation to maintenance	Failures information e.g. failure mode, time, etc. and up normal signs e.g. noise	Verbal communication
Maintenance to operation	Machine limit information, e.g. speed, and setting up advices	Mostly verbal or written in some critical situations
Maintenance to quality	Limitations and capabilities of machines to produce the required quality Performed preventive maintenance	Verbal communication
Quality to maintenance	–	–

Table 4 The role of operation, quality and maintenance in each process

No.	Loss categories according to the strategic goals	Losses in units	Losses in %	Comments
1	Bad quality (product quality)	7.24	28.7	Losses due to internal causes, e.g. reworking, and external causes, e.g. compensations for customers
2	Less profit margin, which influences product price (competitive price)	18	71.3	Unnecessary production costs due to failures and disturbances
	Total	25.24	100	

study from the actual situation. For example in Table 4, "less profit margin" was calculated as follow: the ideal production rate (74 items/shift) was subtracted from the average of the actual production rate (56 items/shift) and therefore the loss was about 18 units, i.e., 71.3% of the total loss of the categories 1 and 2. Table 4 shows the losses distribution. Other losses in the strategic goals (e.g., losses due to delivery delays, environment violations, etc.) were insignificant according to the company's experience, so they have been neglected. Based on interviews and brainstorming with the company, What's and How's in each phase were formed from maintenance perspective in order to highlight its role. Then the values of each How's were estimated as percentage by the company.

In Table 5, the requirements to maintain the company's goals were extracted and their values were estimated. The loss of each strategic goal (see Table 4) was distributed on the How's based on the company experience due to lack of such data in the company databases. The highest loss share is shown to be from "Production machine condition" followed by "Production rate" with 50 and 27.9% respectively. The priority list shows the importance of the What's which were prioritized according to their economic impact.

The tools that are necessary to maintain the requirements and their value are shown in Table 6. Due to the importance of the How's "Inefficient maintenance policy" on the production process it acquires 38.9%. Then followed by the "Inefficient instruction for doing proper maintenance", which is responsible of 22.8% of the loss share. The activities that affect the tools are shown in Table 7, the highest loss is generated by the How's "Inefficient training to enhance staff competence" which shares the losses by 53.4%. The rest of the losses are distributed almost fairly among the rest of the How's.

In Table 8, the two activities in the What's "Inefficient training to staff competence" and "Inefficient technique for monitoring production rate and working environment" are considered to be the first and second priority due to their major economic impact. The major two How's (i.e. factors) that affect the first and second prioritized activities were the "Lack plan for integrating maintenance with plant business" and "Lack of knowledge in the production machines and processes" both are sharing the highest losses 47.3 and 20.1% respectively.

Table 5 The loss shares of the Hows to maintain the company's strategic goals

Losses category in the: outputs to be achieved, maintained and improved (What's)	Requirements to maintain the outputs (Hows)	Production machine condition	Machine tool condition	Working environment	Production rate	Competence of the operating and maintenance staff	Condition of the production logistics system	Total importance of Whats (WT%)	Priority list of the actions required for improvement (PL)
Bad quality		14.4	2.9	1.4	8.6	1.4	0	28.7	2nd
Lost profit margin (due to failures)		35.6	5	2.1	19.3	5	4.3	71.3	1st
Importance of Hows (Total %)		50	7.9	3.5	27.9	6.4	4.3	100	

Table 6 The shares of the tools in affecting the requirements

What's	How's	Inefficient maintenance policy	Inefficient measuring and analysis system	Inefficient cost effective and continuous improvement policy	Inefficient instruction for proper maintenance	WT (%)	PL
Production machine condition		20	10	10	10	50	1st
Machine tool condition		3.9	2.4	0	1.6	7.9	
Working environment		1.4	0.3	1.5	0.3	3.5	
Production rate		11.1	2.8	5.6	8.4	27.9	2nd
Competence of the operating and maintenance staff		1.9	1.9	0.7	1.9	6.4	
Condition of the production logistics system		0.6	1.7	1.3	0.6	4.3	
HT (%)		38.9	19.1	19.1	22.8	100	

Table 7 The shares of the activities in affecting the tools

What's	How's	Inefficient training to enhance staff competence	Inefficient technique for monitoring production rate and working environment	Inefficient measures for monitoring process performance and cost-effectiveness	Inappropriate data and knowledge base	WT (%)	PL
Inefficient maintenance policy		19.5	7.8	3.9	7.8	38.9	1st
Inefficient measuring and analysis system		11.5	3.8	1.9	1.9	19.1	
Inefficient cost effective and continuous improvement policy		13.3	1	2.9	1.9	19.1	
Inefficient instruction for doing maintenance properly		9.1	6.8	4.6	2.3	22.8	2nd
HT (%)		53.4	19.4	13.3	13.9	100	

Table 8 The shares of the factors affecting the activities

What's	How's	Lack of plan for integrating maintenance with plant business	Lack of capital for integrating maintenance with plant business	Lack of knowledge in the production machines and processes	Lack of managerial tools	Lack of criteria for selecting tools and policies	WT %	PL
Inefficient training to staff competence		26.7	5.3	10.7	5.3	5.3	53.4	1st
Inefficient technique for monitoring production rate and working environment		9.7	3.9	3.9	0.6	1.4	19.4	2nd
Inefficient measures for monitoring process performance and cost-effectiveness		6.7	1.3	1.3	2.6	1.3	13.3	
Inappropriate data and knowledge base		4.2	0.7	4.2	2.8	2.1	13.9	
HT%		47.3	11.2	20.1	11.3	10.1	100	

5 Discussion

In Table 2 the operation team is involved in the process of "setting up", "Raw material input" and "Output and quality control". The quality team is involved in the process "Raw material input" and "Output and quality control". Regarding maintenance, as problems may occur in any process, the maintenance is involved in all processes. This highlights the maintenance importance in supporting the production process. In agreement with the Pareto principal 80/20, in general, the highest loss of sharing is generated by few causes (one or at maximum two How's), while the rest of the losses are distributed among the rest of the How's. One clear example is in Table 5, where the values 50 and 28% of the losses are generated by "production machine condition" and "production rate" respectively, while the rest of the losses are distributed among the other How's. Concerning the activities integration, the analysis shows that there is a lack of integration between quality and maintenance. There is no direct feedback from quality to maintenance, see Table 3. It is interesting that this stands in agreement with the outcomes of Table 8, where "Inefficient plan for integrating maintenance with plant business" acquired 47.3% of the losses sharing which is relatively high. The integration between quality and maintenance could be very useful for the company performance. For example, this information could be used to map the machine's condition to the output quality. This will help in identifying problems that reduces the items' quality, and then, this can be used to reduce the loss due to "bad quality". Note that, the "bad quality" loss has a considerable impact on the profitability, as shown in Table 4, i.e., "bad quality" loss acquired 28.7% of the total loss of the production units. Also, Table 8 shows that the "Inefficient plan for integrating maintenance with plant business" shares almost 50% of the loss of both the first and second priority i.e., 26.7% out of 53.4% in the "Inefficient training to staff competence" (first priority) and 9.7% out of 19.4% in "Inefficient technique for monitoring production rate and working environment" (second priority).

6 Conclusion

This study shows a quantitative assessment of the impact of maintenance integration on a case company. Also, it shows a practical application of the MFD model, and its possible benefits. This model assists the user to pinpoint, quantify and prioritize the economic loss and their causes in a production process. The study shows the importance of maintenance on the company profitability and its effect on the other working areas. It shows the importance of the planning and considering maintenance integration with the rest of the plants activities. It is concluded that better maintenance functions can maximize the productivity and enhance the quality. To increase the profitability and the company's competitiveness the authors recommend improving the integration of maintenance with the quality activities.

Also, to increase the level of the training of personals to detect problems related to machine at an early stage. A future work could include developing a supporting tool to increase the objectivity of the gathered data for the MFD model as well as questionnaires or interview structure.

References

1. Al-Najjar B (2003) Modelling the design of the integration of quality, maintenance and production. In: Proceeding of the 16th international congress COMADEM 2003. Växjö, Sweden, pp 23–31
2. Waeyenbergh G, Pintelon L (2002) A framework for maintenance concept development. Int J Prod Econ 77(3):299–313
3. Hartmann J, Germain R (2015) Understanding the relationships of integration capabilities, ecological product design, and manufacturing performance. J Clean Prod 92:196–205
4. Kahn KB, Mentzer JT (1996) Logistics and interdepartmental integration. Int J Phys Distrib Logist Manage 26(8):6–14
5. Al-Najjar B, Algabroun H (2016) A model for increasing effectiveness and profitability of maintenance performance : a case study. In: 2016WCEAM international conference
6. 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. Int J Prod Econ 107(1):260–273
7. Al-Najjar B (1996) Total quality maintenance: an approach for continuous reduction in costs of quality products. J Qual Maint Eng 2(3):4–20
8. Al-Najjar B (2011) Maintenance impact on company competitiveness & profit. In: Asset management: the state of the art in Europe from a life cycle perspective. Springer Science Business Media B.V., Dordrecht, Heidelberg, London, New York, pp 115–141
9. Govers CPM (1996) What and how about quality function deployment (QFD). Int J Prod Econ 46(95):575–585

Comparative Study: Linear and Nonlinear Transient Stability Analysis for Meso Scale Rotor with Gas Foil Journal Bearings

Skylab P. Bhore and Ashish K. Darpe

Abstract A comparative study on linear and nonlinear transient stability analysis of a meso scale rotor supported on gas foil journal bearings is carried out. Steady and unsteady Reynolds equations are solved by using the control volume formulation with a power law scheme and the Gauss–Seidel method. The modal impedance and the time domain orbit simulation are used to obtain the stability parameter: the threshold rotor mass for the linear and nonlinear transient methods respectively. The modal impedance is calculated by using the linear stiffness and damping coefficients. However, the orbit simulation solves explicitly the journal equations of motion, the compressible unsteady Reynolds equation and the equation of foil deformation. The threshold rotor mass which gives the stability threshold value for a given operating speed and load is obtained. It is found that there is a significant difference between the values of threshold rotor mass predicted by the linear and nonlinear transient stability methods.

1 Introduction

The meso scale gas turbine is a newly designed portable power source. Such gas turbines work on hydrocarbon fuels and give high power density (power per unit volume of the system) and high energy density (energy per unit mass of the system). A meso scale gas turbine system consists of a radial inflow turbine, a centrifugal compressor and an annular combustion chamber. The operating speed range is between 500,000 and 1,000,000 rpm for meso scale gas turbines. It is envisaged that the meso scale gas turbines can be used as continuous power supply for several

S. P. Bhore (✉)
Department of Mechanical Engineering, Motilal Nehru National Institute of Technology Allahabad, Allahabad, India
e-mail: skylabpbhore@mnnit.ac.in

A. K. Darpe
Department of Mechanical Engineering, Indian Institute of Technology Delhi, Delhi, India
e-mail: akdarpe@mech.iitd.ac.in

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_8

applications viz. micro power generation for robots, micro satellites, micro systems, hybrid energy system with fuel cells, unmanned air vehicles (UAVs), etc. [1]. Vleugels et al. [2] presented a linear stability analysis of a bump foil bearing. Salehi et al. [3] have reported the operation of meso scale gas turbine simulator supported on foil bearings at speed of 700,000 rpm. The test was carried out with specially designed foil journal and thrust bearings. Lee et al. [4] have performed test on meso scale rotor supported with foil bearing with target speed of 300,000 rpm and reported that the designed foil journal and thrust bearings had enough load carrying capacity and stiffness and damping characteristics to maintain the stability of the system. Kim et al. [5] have designed and manufactured meso scale foil bearings using nonconventional methods such as X-ray lithography, ultraviolet lithography, electroplating and molding to manufacture the meso scale foil bearings using metals.

In case of macro scale gas foil bearing, San Andres and Kim [6] studied the nonlinear behaviour of gas foil bearing. Kim [7] also studied the nonlinear aspects of foil bearing and found significant differences in the estimated onset speeds of instability. Kim et al. [5] have presented the rotordynamic analysis of a meso scale rotor with foil bearings. They have not studied the stability aspects of meso scale foil bearings. Bhore and Darpe [8] have presented the nonlinear dynamics of a macro scale rotor with gas foil bearings. The literature on foil bearings mentions about its strong nonlinear behaviour. Therefore, stability analysis based on linear stiffness and damping coefficients may be less accurate. A comparative study on linear and nonlinear transient stability analysis is essential. As a meso scale rotor runs beyond the conventional speed range, its dynamic behavior and stability at such a high speed is an important concern in the design aspects of meso scale rotor bearing system.

In this paper a comparative study on linear and nonlinear transient stability analysis of a meso scale rotor with gas foil journal bearings is carried out. A stability parameter, the threshold rotor mass, is obtained by both linear and nonlinear transient method. The comparison of threshold rotor mass and detail analysis has been presented.

2 Compressible Reynolds Equation

In order to represent the phenomenon of hydrodynamic lubrication with the compliant effect of foil bearing, the model proposed by Heshmat et al. [9] and Kim [7] is used. The energy dissipation in the foil bearing occurs due to the relative displacement between bump foil and housing and between bump foil and top foil. It is modeled by an equivalent viscous damping [7]. The sagging effect in the top foil is proportional to the bump pitch and inversely proportional to the top foil thickness [2]. In case of meso scale foil bearings, the value of bump pitch becomes small. In addition, due to manufacturing constraints, the reduction of top foil thickness is limited therefore the resulting sagging effect will be smaller and can be neglected.

For the coordinate system shown in Fig. 1, the compressible unsteady Reynolds equation with isothermal condition is

$$\frac{\partial}{\partial\theta}\left(\bar{p}\bar{h}^3\frac{\partial\bar{p}}{\partial\theta}\right)+\frac{\partial}{\partial\bar{z}}\left(\bar{p}\bar{h}^3\frac{\partial\bar{p}}{\partial\bar{z}}\right)=\Lambda\frac{\partial}{\partial\theta}(\bar{p}\bar{h})+2\Lambda v\frac{\partial}{\partial\tau}(\bar{p}\bar{h}) \tag{1}$$

In Eq. 1, the dimensionless groups for pressure $\bar{p}$, film thickness $\bar{h}$, axial bearing length $\bar{z}$, time τ, speed ratio v and bearing number Λ are as follows,

$$\bar{p}=\frac{p}{p_a},\bar{h}=\frac{h}{C},\bar{z}=\frac{z}{R},\tau=\omega t,v=\frac{\omega_s}{\omega},\Lambda=\frac{6\mu\omega}{p_a}\left(\frac{R}{C}\right)^2 \tag{2}$$

Following the coordinate system in Fig. 1, the film thickness for the gas foil bearing is

$$\bar{h}=1-\varepsilon_X\cos(\theta)-\varepsilon_Y\sin(\theta)+\delta(\theta,\bar{z}) \tag{3}$$

where, $\varepsilon_X,\varepsilon_Y$ are eccentricity ratios in X and Y direction respectively, $\delta(\theta,\bar{z})$ is a foil deformation. The foil deformation depends on the applied pressure $\bar{p}$ and compliance α of foil. The deformation of bump foil at steady state is

$$\delta(\theta,\bar{z})=\alpha(\bar{p}(\theta,\bar{z})-1) \tag{4}$$

where, α is a compliance of the bump foil [9] and expressed in dimensionless group as follows,

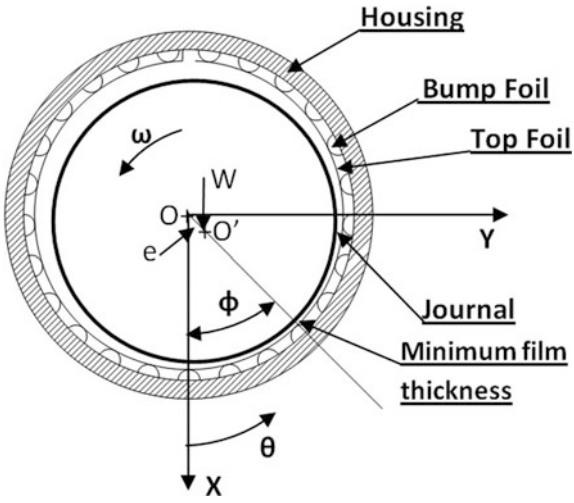


Fig. 1 Co-ordinate system for bump foil bearing

$$\alpha = \frac{2p_a S_o}{CE}\left(\frac{l_o}{t_b}\right)^3 (1-\upsilon^2) \tag{5}$$

The deformation of the bump foil at unsteady state [7] is

$$\delta^{\tau} = \frac{C_B}{C_B + K_B \Delta\tau}\delta^{\tau-\Delta\tau} + \frac{\Delta\tau}{C_B + K_B \Delta\tau}\bar{p}^{\tau}(\theta, \bar{z}) \tag{6}$$

where, K_B and C_B are the stiffness and damping coefficients respectively.

$$K_B = \frac{1}{\alpha}, \quad C_B = \left(\frac{\eta}{v}\right)K_B \tag{7}$$

where, α is a compliance of the bump foil, η is a loss factor of the bump foil material and v is a speed ratio.

For solving Eq. 1, the following boundary conditions are used.

$$\left.\begin{array}{l} \bar{p}(\bar{z} = \pm L/2R, \theta) = 1 \\ \bar{p}(\bar{z}, \theta = \pi) = 1 \\ \bar{p}(\bar{z}, \theta = 0) = \bar{p}(\bar{z}, \theta = 2\pi) \end{array}\right\} \tag{8}$$

Equation 1 is solved by using the control volume method and pressure distribution is obtained. The pressure distribution field $\bar{p}(\theta, \bar{z})$ is integrated over the bearing surface area using Simpson's 1/3rd rule [10].

The bearing reaction forces along X and Y direction are

$$f_X = -p_a R^2 \int_0^{2\pi}\int_0^{L/R} \bar{p}(\theta, \bar{z})\cos(\theta) d\theta d\bar{z} \tag{9}$$

$$f_Y = -p_a R^2 \int_0^{2\pi}\int_0^{L/R} \bar{p}(\theta, \bar{z})\sin(\theta) d\theta d\bar{z} \tag{10}$$

Equation 1 is also used for the perturbation method which gives linear stiffness and damping coefficients. In this, an infinitesimal perturbation amplitude is used and it is based on Taylor series (higher order term neglected). The method is used by Kim and San Andres [11] for obtaining the linear stiffness and damping coefficients of a gas foil bearing. A similar approach has been used by the authors in their earlier research work [10]. In the present paper, the same method is used for the calculation of linear stiffness and damping coefficients.

3 Linear Stability for Meso Scale Rotor with Gas Foil Journal Bearings

The linear stability analysis based on linear stiffness and damping coefficients is performed. Following the procedure in [10], the excitation frequency dependent stiffness and damping coefficients for different rotor speed and eccentricity ratio ε (or applied load *fe*) are obtained. For a Jeffcott rotor having two degrees of freedom and with no external force, the equations of motion in matrix form can be written as,

$$\begin{bmatrix} m_r & 0 \\ 0 & m_r \end{bmatrix} \begin{Bmatrix} \Delta\ddot{X} \\ \Delta\ddot{Y} \end{Bmatrix} + \begin{bmatrix} B_{XX} & B_{XY} \\ B_{YX} & B_{YY} \end{bmatrix} \begin{Bmatrix} \Delta X \\ \Delta Y \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{11}$$

where, $B_{IJ} = k_{IJ} + jc_{IJ}\omega_s$ is the impedance of the foil bearing in which stiffness and damping coefficients vary with the excitation frequency.

With the assumption of nonzero Eigen modes $\{\Delta X_m \, \Delta Y_m\}^T_{G=bacward/forward}$, $B_G = -m_r\lambda$ and $\lambda = (\alpha \pm j\omega_s)^2$, an Eigen value problem can be written as,

$$\begin{bmatrix} B_{XX} - B_G & B_{XY} \\ B_{YX} & B_{YY} - B_G \end{bmatrix} \begin{Bmatrix} \Delta X_m \\ \Delta Y_m \end{Bmatrix}_{G=backward/forward} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{12}$$

After solving Eq. 12, the modal impedance for different modes can be written as,

$$B_{G=backward/forward} = \frac{B_{XX} + B_{YY}}{2} \pm \sqrt{\left(\frac{B_{XX} - B_{YY}}{2}\right)^2 + B_{XY}B_{YX}} \tag{13}$$

The modal impedance is a complex number; the real and imaginary part represent modal stiffness and modal damping respectively. When the modal damping becomes zero; the corresponding excitation frequency ratio and modal stiffness are called as v_{eigen} and critical modal stiffness k_{s_modal} respectively. Thus critical rotor mass m_{crit} for given rotor speed ω and eccentricity ratio ε can be obtained from the following relation,

$$m_{crit} = \frac{k_{s\_modal}}{\left(v_{eigen}\omega\right)^2} \tag{14}$$

4 Nonlinear Transient Stability for Meso Scale Rotor with Gas Foil Journal Bearings

A nonlinear transient stability analysis is performed using orbit simulation. Orbit simulation is a time domain analysis technique. In this, the locus of the journal centre is traced by solving the journal equations of motion, the compressible unsteady Reynolds equation and the foil deformation equation simultaneously. To perform nonlinear transient stability analysis (discussed in the next section), a Jeffcott rotor with 2 DOF is used. For this, the journal equation of motion in X and Y direction [12] are

$$\ddot{\varepsilon}_X = \frac{p_a R^2}{m_r C \omega^2}\left(-\int_0^{2\pi}\int_0^{\frac{L}{R}} p(\theta, z)\cos(\theta) d\theta dz + \frac{fe_Y}{p_a R^2}\right) \tag{15}$$

$$\ddot{\varepsilon}_Y = \frac{p_a R^2}{m_r C \omega^2}\left(-\int_0^{2\pi}\int_0^{\frac{L}{R}} p(\theta, z)\sin(\theta) d\theta dz + \frac{fe_X}{p_a R^2}\right) \tag{16}$$

In Eqs. 15–16, the first term in the bracket is representing the bearing reaction forces and the second term is the external load including rotor weight $fe = \sqrt{fe_X^2 + fe_Y^2}$. Using Eqs. 15–16, the velocity and displacement of journal center at each time step are obtained as follows,

$$\dot{\varepsilon}_X(\tau + \Delta\tau) = \dot{\varepsilon}_X(\tau) + \ddot{\varepsilon}_X(\tau + \Delta\tau)\Delta\tau \tag{17}$$

$$\dot{\varepsilon}_Y(\tau + \Delta\tau) = \dot{\varepsilon}_Y(\tau) + \ddot{\varepsilon}_Y(\tau + \Delta\tau)\Delta\tau \tag{18}$$

$$\varepsilon_X(\tau + \Delta\tau) = \varepsilon_X(\tau) + \dot{\varepsilon}_X(\tau + \Delta\tau) + \frac{1}{2}\ddot{\varepsilon}_X(\tau + \Delta\tau)\Delta\tau^2 \tag{19}$$

$$\varepsilon_Y(\tau + \Delta\tau) = \varepsilon_Y(\tau) + \dot{\varepsilon}_Y(\tau + \Delta\tau) + \frac{1}{2}\ddot{\varepsilon}_Y(\tau + \Delta\tau)\Delta\tau^2 \tag{20}$$

4.1 Procedure of Finding Threshold Rotor Mass

The threshold rotor mass is obtained through the following steps,

1. Select a trial rotor mass m_r
2. Find the bearing reaction forces (f_X, f_Y given by Eqs. 9–10) for the required speed N and eccentricity ratio ε or load fe.

3. Apply the external load fe_X and fe_Y (equivalent to bearing reaction force f_X and f_Y respectively) in opposite direction to get the static equilibrium position of the rotor (Eqs. 15–16).
4. Disturb the rotor from its equilibrium position by a small perturbation $\Delta\varepsilon_X = 0.001\varepsilon_X$ in X and $\Delta\varepsilon_Y = 0.001\varepsilon_Y$ in Y directions. At $\tau = 0, \varepsilon_X = \varepsilon_X + \Delta\varepsilon_X$ and $\varepsilon_Y = \varepsilon_Y + \Delta\varepsilon_Y$ are the initial perturbations.
5. Trace the path of the rotor center (ensuing motion of the rotor after initial perturbation is given).
6. Check the trace of rotor center, if the rotor center returns to its static equilibrium position then the rotor is stable; otherwise it qualifies as unstable rotor.
7. Repeat the procedure starting from step-1 with an increased value of rotor mass to find the stability behavior.
8. The iteration stops when the new value of (increased) rotor mass makes the rotor unstable.

The mass of the rotor at which the rotor becomes unstable is called threshold or critical rotor mass m_{crit}.

5 Results and Discussion

5.1 Numerical Simulation and Validation

In order to solve Reynolds Eq. 1 and find the pressure field $\bar{p}(\theta, \bar{z})$, a control volume with a power law hybrid scheme and Gauss–Seidel iterative method are used [8, 10]. The code is written in Matlab-7.12 and the following convergence criterion is implemented.

$$\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left|P_{i,j}^{iter+1} - P_{i,j}^{iter}\right|}{\sum_{i=1}^{m}\sum_{j=1}^{n}\left|P_{i,j}^{iter+1}\right|} \leq 10^{-5} \tag{21}$$

The comparison of load carrying capacity with results given in Peng and Carpino [13] is carried out for different eccentricity ratios. It is found that there is a close agreement between the results (Fig. 2).

5.2 Comparison of Threshold Rotor Mass

Results are obtained for a compliance of the bump foil $\alpha = 0.1$, a loss factor $\eta = 0.25$ and an eccentricity ratio $\varepsilon = 0.6$. Following the procedure given in Sect. 3, the variation of modal impedance with excitation frequency ratio v for rotor speed N = 100,000–700,000 rpm are obtained. Figure 3 shows the variation of

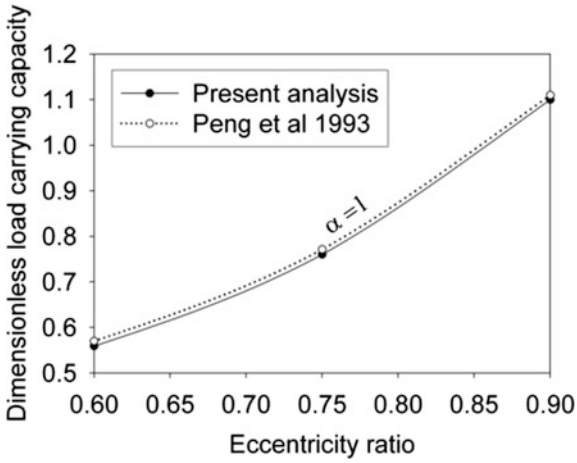


Fig. 2 Validation of load carrying capacity with reference Peng et al. [13]

modal impedance with the excitation frequency ratio a for rotor speed N = 500,000 rpm. The modal damping becomes zero at $\nu = 0.271$; and the corresponding value of modal stiffness is obtained from Fig. 3. The value of critical rotor mass is calculated by using Eq. 14. A similar procedure is used for calculating the threshold rotor mass for different speeds, as tabulated in Table 1.

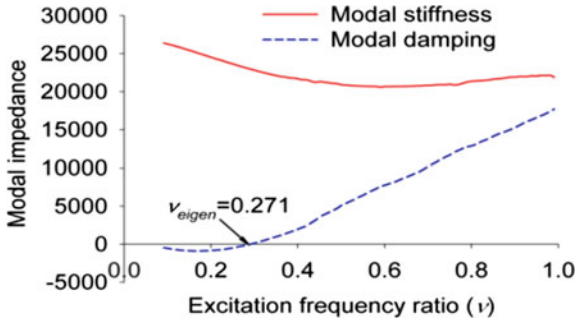


Fig. 3 Variation of modal impedance at $\varepsilon = 0.6$ and N = 500,000 rpm

Table 1 Comparison of m_{crit} with linear and nonlinear transient stability analysis

	Linear stability ($\varepsilon = 0.6$, $\alpha = 0.1$)	Nonlinear transient stability ($\varepsilon = 0.6$, $\alpha = 0.1$)
Speed (rpm)	m_{crit} (kg)	m_{crit} (kg)
100,000	8.1955E−4	0.006
200,000	3.4631E−4	0.005
300,000	1.8170E−4	0.005
400,000	1.5679E−4	0.004
500,000	1.1567E−4	0.006
600,000	8.9726E−5	0.004
700,000	4.9693E−5	0.001

The critical rotor mass m_r is also obtained by using the nonlinear transient stability method following the procedure given in Sect. 4. Figures 4 and 5 depicts the time domain rotor center responses in X and Y direction at $\varepsilon = 0.6$ and $N = 500{,}000$ rpm after perturbation for rotor masses $m_r = 6 \times 10^{-3}$ kg and 7×10^{-3} kg respectively. It is observed that the rotor returns to its equilibrium position after perturbation for the rotor mass $m_r = 6 \times 10^{-3}$ kg. Therefore, the rotor mass $m_r = 6 \times 10^{-3}$ kg qualifies as stable rotor mass. At rotor mass $m_r = 7$ 10^{-3} kg, the rotor does not return to its equilibrium position after perturbation. The time domain rotor center responses in X and Y direction show limit cycle behavior and therefore rotor mass $m_r = 7 \times 10^{-3}$ kg becomes the threshold value. Similarly, the critical rotor mass is obtained for other rotor speeds and recorded in Table 1. From Table 1, it is found that there is a significant difference between the values of m_{crit} predicted by the above two methods. This difference is due to the presence of an infinitesimal perturbation amplitude (for calculation of stiffness and damping coefficients) in the linear stability method while the nonlinear stability method involves nonlinear perturbation amplitude, unsteady nonlinear fluid film bearing reaction forces and unsteady foil deformation.

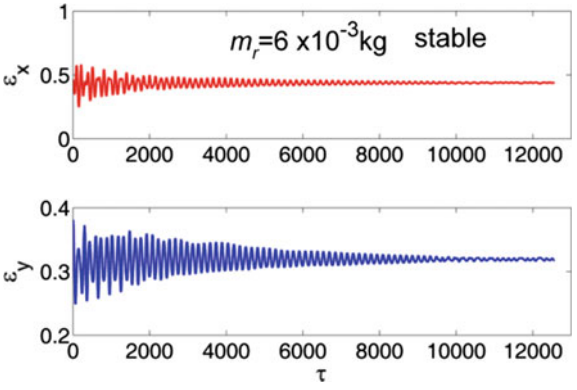


Fig. 4 Time domain responses of rotor center after perturbation in X and Y direction for $m_r = 0.006$ kg at $\varepsilon = 0.6$ and $N = 500{,}000$ rpm

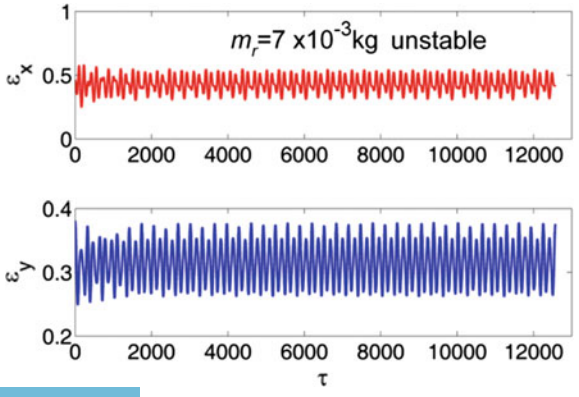


Fig. 5 Time domain responses of rotor center after perturbation in X and Y direction for $m_r = 0.007$ kg at $\varepsilon = 0.6$ and $N = 500{,}000$ rpm

Although there is a need of actual experimental validation, the results of the nonlinear stability method can be accepted because of its more realistic representation of the rotor bearing vibration problem.

6 Conclusions

The present work is an attempt to do a comparative study of linear and nonlinear transient stability analysis of meso scale rotor with gas foil journal bearings. A stability parameter, the threshold rotor mass, is obtained by both linear and nonlinear transient method. The comparison of threshold rotor mass is carried out. It is found that there is a significant difference between the values of threshold rotor mass predicted by linear and nonlinear transient stability analysis. This difference is due to the presence of an infinitesimal perturbation amplitude (for calculation of stiffness and damping coefficients) in the linear stability method. However, the nonlinear stability method involves nonlinear perturbation amplitude, unsteady nonlinear fluid film bearing reaction forces and unsteady foil deformation. Although there is a need for actual experimentation, the threshold rotor mass predicted by the nonlinear transient method may be a more accurate because of its explicit formulation.

References

1. Bhore SP, Darpe AK (2014) Rotordynamics of micro and mesoscopic turbomachinery: a review. J Vib Eng Technol 2(1):1–9
2. Vleugels P, Waumans T, Peirs J, Al-Bender F, Reynaerts D (2006) High-speed bearings for micro gas turbines stability analysis of foil bearings. IOP J Micromech Microeng 16:S282–S289
3. Salehi M, Heshmat H, Walton JF, Tomaszewski M (2007) Operation of a mesoscopic gas turbine simulator at speeds in excess of 700,000 rpm on foil bearings. ASME J Eng Gas Turbines Power 129:170–176
4. Lee Y-B, Park D-J, Kim C-H, Ryu K (2007) Rotordynamic characteristics of a micro turbo generator supported by gas foil bearings. IOP J Micromech Microeng 17:297–303
5. Kim D, Creary A, Chang SS, Kim JH (2009) Mesoscale foil gas bearings for palm sized turbomachinery: design, manufacturing and modeling. ASME J Eng Gas Turbines Power 131:1–10
6. San Andres L, Kim TH (2008) Forced nonlinear response of gas foil bearing supported rotors. Tribol Int 41:704–715
7. Kim D (2007) Parametric studies on static and dynamic performance of air foil bearings with different top foil geometries and bump stiffness distributions. ASME J Tribol 129(2):354–364
8. Bhore SP, Darpe AK (2013) Nonlinear dynamics of flexible rotor supported on the gas foil journal bearings. J Sound Vib 332:5135–5150
9. Heshmat H, Walowit JA, Pinkus O (1983) Analysis of gas lubricated foil journal bearings. ASME J Lubr Technol 105:647–655

10. Bhore SP, Darpe AK (2013) Investigations of characteristics of micro/meso scale gas foil journal bearings for 100-200 W class micro power system using 1st order slip velocity boundary condition and the effective viscosity model. J Microsyst Technol 19:509–523
11. Kim TH, San Andres L (2008) Heavily loaded gas foil bearings: a model anchored to test data. ASME J Eng Gas Turbines Power 130:1–8
12. Kim D, Lee S, Bryant MD, Ling FF (2004) Hydrodynamic performance of gas microbearings. J Tribol 126(2):711–718
13. Peng J-P, Carpino M (1993) Calculation of stiffness and damping coefficients for elastically supported gas foil bearings. ASME J Tribol 115(1):20–27

Predictive Models of Maintenance Needs for Power Distribution Wood Poles Using Machine Learning—A Conceptual Case Study

Alexandre Cesa, Carla Boehl and Kecheng Shen

Abstract The maintenance of a large network of power distribution wood poles is a challenging activity, in part due to the extensive geographical area in which the network is distributed. It is not uncommon that an operator has to maintain an infrastructure with hundreds of thousands of wood poles installed over an area bigger than many countries. Behind the challenge, there are various factors that drive the maintenance and renewal requirements of wood poles including their age, wood species, insect activity, soil composition, weather conditions and many others. As these factors highly interact with each other, it is not immediately obvious or simple to decide which poles to prioritise for the maintenance/renewal in lieu of others. Therefore, an improved process capability of decision making in this regard will have direct positive implications on the ownership cost, safety and level of service such as availability of the network. This paper discusses and illustrates how machine learning techniques could support decision making processes by predicting wood poles maintenance needs, i.e. by concomitantly assessing the multiple factors that drive the pole failures/integrity loss, through the use of diverse and disparate data sources such as geographical area features and scanned hand-written records of historical maintenance.

1 Introduction

The effective and efficient maintenance of a wood pole power distribution network is a complex topic. A few key challenges include the large geographical area where the wood poles are installed and the different variables (also called features

A. Cesa (✉) · C. Boehl · K. Shen
Curtin University, Perth, Australia
e-mail: 19107084@student.curtin.edu.au

C. Boehl
e-mail: carla.boehl@curtin.edu.au

K. Shen
e-mail: kechengshen@gmail.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_9

throughout the paper) that drive their economical and safe life. It is not uncommon a power distribution operator having under its management hundreds of thousands of wood poles—each one potentially requiring frequent inspection to assess its condition and maintenance needs.

A report issued by the Western Australia General Auditor in 2013 [1] highlights the topic's complexity and its relevance. It was found the practices that Western Power (operator) had followed to maintain its 600+ thousand wood poles across the state were not adequate, with direct adverse implications to the safety and level of service of the network. In this particular case, decades of underinvestment means an expenditure in excess of AU$1 Billion, above and beyond normally expected maintenance expenditures, would be required within the next decade to re-instate the safety of the network. In all, the report suggests that Western Power's Asset Management System was not adequate [1].

This paper discusses some conceptual and practical aspects of the analytical techniques that can support a better decision making process concerning the wood poles problem setting. It explores how machine learning methods could be used to handle large amounts of disparate data. In particular, the K-Nearest Neighbour method is used to demonstrate how geographical features of a large footprint wood poles network can be attained. In addition, an Artificial Neural Network method is used to attain features from hand written historical inspection sheets scans. Finally, the concept of how the disparate data sets could be analysed together and used to drive better informed decisions for maintaining this particular asset type are discussed.

2 Background and Motivation

The expenditure figures the report [1] suggests are fundamentally based on future maintenance requirements given an average wood pole expected life and the age distribution of the poles' population installed on this particular network. This "maximum age threshold" approach suggested by the report is likely to improve this particular poor situation observed but it would not be much cost and risk effective. In regards to the cost, it will over-maintain poles that do not yet require maintenance or replacement. It will also under-maintain poles with a less-than average life—which presents a safety risk to the community. For these reasons, it can be argued a more robust method of wood poles maintenance management is necessary.

More comprehensive approaches than solely using the age threshold criteria for wood poles replacement already exist. References [2–4] are examples, where documented standards and decision trees support determining maintenance needs of individual wood poles. Such documents are the result of many years of experience making it possible to achieve an acceptable level of asset integrity and performance. Often, these guidelines are a collection of technical considerations, practical experience and heuristics shortcuts [3].

Life of power distribution wood poles is affected by at least 34 different variables including soil type, weather condition, wood species, pole load and many others [5–7]. Mapping many of the influences that individual variables have on wood poles life could be easily achievable through life distribution such as Weibull and other simple modelling techniques like linear regression. These models would typically illustrate e.g. "wood pole life distributions by type of wood species" or "wood pole life distributions by geographical areas". One shortly realizes that these variables are hardly independent from each other. For instance, life by wood species is affected by fungi rot, which in turns depends on soil type and a "catalyst" like rain to grow. Fungi rot would also be affected by which chemical treatments the wood pole might have received and how long ago it happened [6, 7]. In all, this very complex structure of variables' dependency is what ultimately drive the lives of individual wood poles. This leads to the realization that single-variable or more generally low dimensionality models (i.e. models with few variables) as so far described could only be effective to some very limited extent.

In all, it can be said an ideal maintenance decision model would be one capable of predicting with a great deal of accuracy the life characteristics of each individual pole. The achievement of such granular predictive capability would enable one to truly optimise this asset system's performance, cost and risk profiles as a whole. There are different reasons why simple maintenance models as before discussed are suggested and ultimately adopted. These include the complex multivariate data analysis and the computational capability required to undertake it, the difficulty in retrieving and consolidating historical data stored in different formats and others. As time elapses and information technologies evolve, these challenges will become easier to deal with.

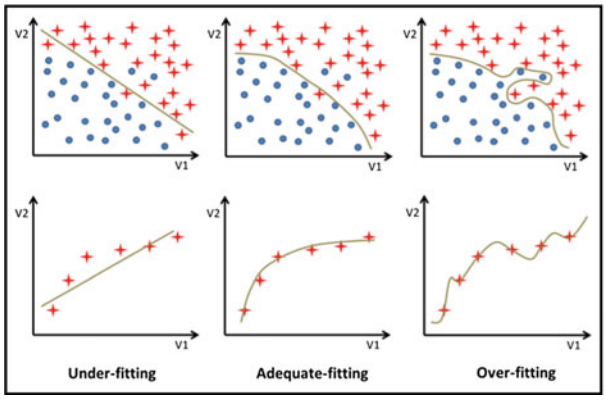
3 Literature Review

Machine learning (ML, also known as Statistical Learning) refers to a large set of tools for modelling and understanding complex data structures. With the ever increasing number of "big data" problems, ML has become a very hot field in many scientific areas (as well as other disciplines). Broadly speaking, supervised ML involves building a statistical model learnt from the data itself for predicting, or estimating, an output based on one or more inputs [8]. In other words, ML models differ from other modelling techniques that depend on prior knowledge of first principles and cause and effects. This means one can successfully build complex and yet accurate ML models without domain knowledge of the topic (though domain knowledge sometimes helps to increase model assertiveness).

There are different ML techniques; a few popular ones include Linear and Logistic Regression, Support Vector Machines (SVM), K-Nearest Neighbour (KNN) and Artificial Neural Networks (ANN) [9, 10]. Regardless of the technique selected, the idea of supervised learning involves using a dataset with known inputs and outputs to "train" a selected model hypothesis. By training the model we mean

iteratively conducting model parameterisation with the objective of minimising the model's output error: i.e. given e.g. inputs "A", "B" and "C" the model should return output "X" (and not "Y", for instance). It is quite easy for a multi-parameter model hypothesis to deliver a very low or zero-error output. Although this might sound promising, such a model would be of limited value to predict outputs from inputs not yet seen, meaning the model does not generalize well. Such a problem is related to the bias-variance trade-off and is illustrated in Fig. 1 by three model hypothesis fits to the same dataset: an under-fitting case (left, high bias), a reasonable model fit case (middle) and an over-fitting case (right, high variance). It can be observed in the under-fitting case that a linear boundary separation is not adequate as it is too simplistic. Alternatively, a highly "wiggly" model fits the training data well but would not likely generalize well to new unseen data. An adequate model hypothesis would be one in which the fit to the dataset minimizes both bias and variance simultaneously, i.e. the reducible error. As illustrated in both classification and regression examples in Fig. 1, in general there will be cases in which the model prediction (regardless of how good it is) would still yield what is known as an irreducible error. In the classification setting, this irreducible error means that in some instances, the model will miss-classify the correct output. This is easily illustrated through a Confusion Matrix as shown in Fig. 1 (right) for a binary classifier case.

There are different techniques to address the issue of model generalization. A very basic approach is to divide the dataset in "training dataset" (e.g. 70% of total size) and "validation/test" dataset (e.g. 30% of total size). This enables one to parameterise the model using both training and validation datasets and later, experimenting by means of the test dataset, the accuracy of the responses predicted by the model. This basic process is conducted iteratively, until the point when the model generalises sufficiently well to meet the user's expected predictive capability.



		Real Class	
		1	0
Predicted Class	1	TP	FP
	0	FN	TN

Fig. 1 Left: Different model hypothesis illustrating classification (top) and regression settings. Right: Confusion matrix—TP (true positive), TN (true negative), FP (false positive or type I error) and FN (false negative or type II error)

Additional iterative approaches include feature engineering, using different models (SVM, ANN, etc.), optimization techniques, model hyper-parameters and others. Resources [8, 11, 12] are gentle, self-contained and comprehensive introductions to the ML topic for the interested reader.

4 Methodology

For the purpose of model development, maintenance requirements of individual poles could be classified as Do nothing, Repair and Replace. A Repair typically involves doing (e.g.) some form of reinforcement of it the at the cost of $1,000 per pole while a Replacement costs approximately $9,000 [1]. Given the safety risks involved and large capital required to sustain level of service of this asset system, more accurate predictions of individual poles is obviously highly desirable.

Two practical uses of Machine Learning techniques are demonstrated. The first one illustrates how to attain geographical features from colour coded maps. A second method illustrates how to attain hand-written features from scanned inspection sheets. Finally, the overarching architecture concept of how these disparate models are collated for later individual poles' classification is illustrated and some relevant aspects of the task are discussed.

4.1 Attaining Geographical Features

Figure 2 illustrates how geographical features of Average Temperature, Average Rainfall, Soil Type and Termite Hazard were attained using Matlab (Ref. [13] contains a web-link to the code). As a first step, a network footprint "mask" is placed over the different feature maps in a process called image registration. Later, the KNN technique is used to discern the different colour levels of these maps, which are translated into different classes (e.g. three classes for the Soil type: Poor Drainage = 0, Average Drainage = 1 and Good Drainage = 2). Finally, feature maps are overlayed on each other and an arbitrary point is selected to show the different variables encoding and their actual physical meaning.

It is not surprising that the maintenance needs observed in particular areas may tend to cluster towards particular outcomes. However, outcomes might still vary widely within a small area and the "geographical map" approach can no longer be extended to successfully predict each individual pole's maintenance needs. Within the small area considered, obviously there will be poles with different ages, which received different levels of service over the years and other idiosyncrasies.

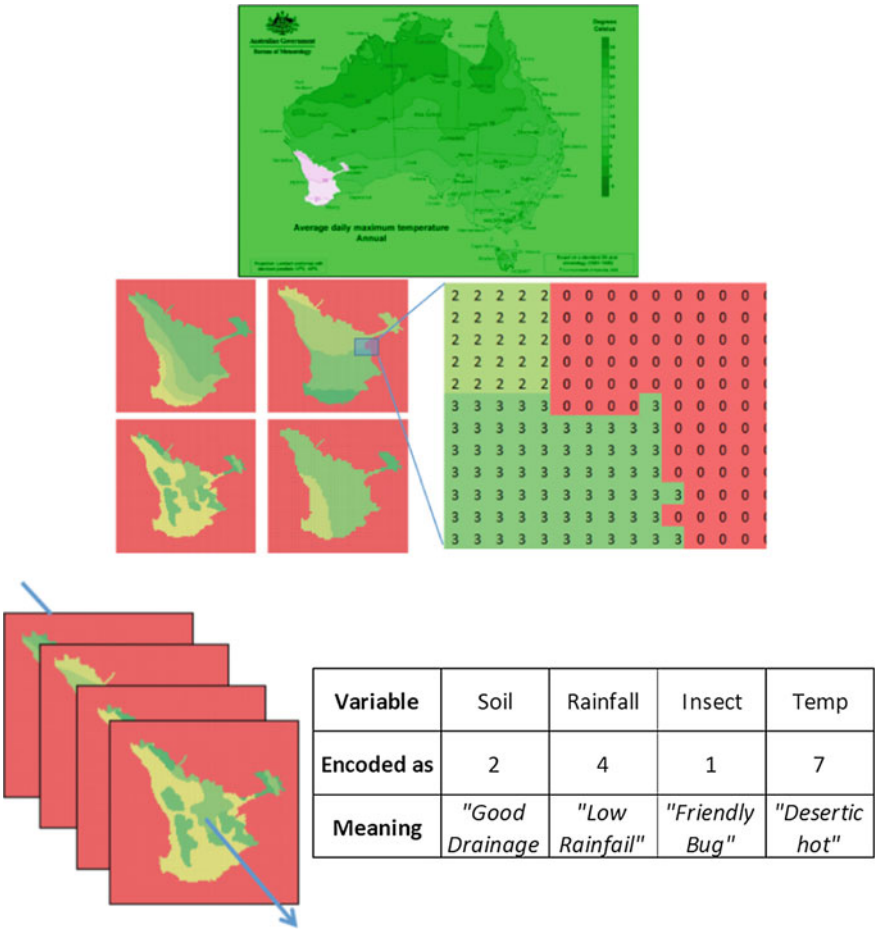


Fig. 2 Western power's footprint image registration in a particular geo-feature map; four feature maps with close-up detail, maps lined up for illustrating features of a particular location

4.2 Attaining Hand Written Features

Individual poles' inspection history is a data source that contains a wealth of information, likely recorded in different formats over the past decades: paper based and perhaps in electronic format more recently. Ignoring historical paper based inspection records would be naïve, so a method to retrieve this data is paramount. For this particular situation, the capability to recognize hand-written digits and letters would of great value. Yan LeCunn pioneers this task utilising the "MNIST dataset"—which is a collection of 70 thousand hand-written numbers from zero to nine [14]. The MNIST dataset is a popular dataset used by the ML community to benchmark algorithm performance in regards to the image recognition task. State of

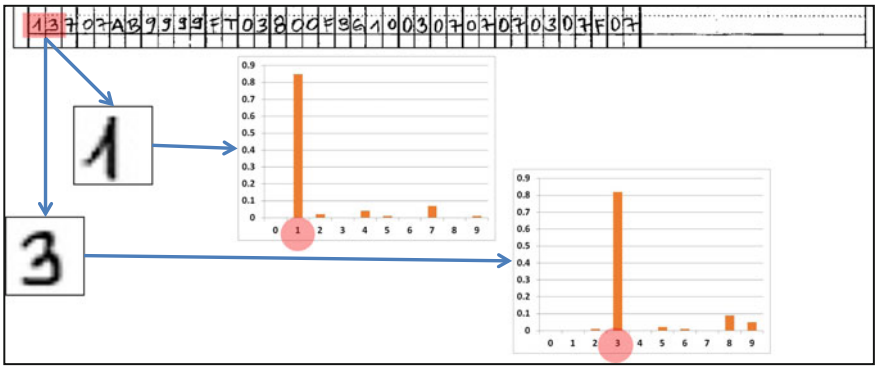


Fig. 3 Written digits scan from (e.g.) a wood pole inspection sheet and ANN correct probabilistic outputs for two example digits, "1" and "3" data encodings

the art algorithms currently reach 99.7% overall accuracy in correctly classifying hand written digits [12]—which in practice is as good as humans can be in the task of classifying these hand-written digits.

Figure 3 shows the bottom scan of a typical wood pole inspection sheet (see e.g. [4] as an example), with hand written records (encodings) regarding e.g. different treatments of the wood pole and their associated dates in between other featured records. The bottom of Fig. 3 shows two digits classification, "one" and "three" outputs from the classifier developed in Matlab [13]. It can be seen the ML ANN algorithm used performs well in the classification task, achieving approximately 95% overall accuracy for the MNIST benchmark.

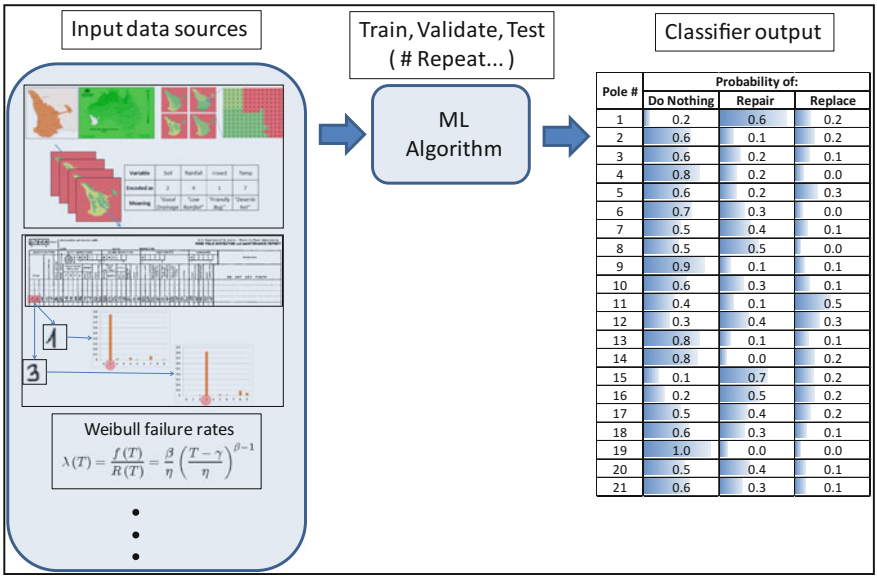
4.3 Final Classifier—Conceptual Architecture

The grouping of geo-features and individual poles historical inspection data records, as demonstrated, would very likely enable one achieving a sufficiently good predictive model of wood poles future maintenance needs. This is in part due to the fact the available dataset is relatively large and there are lots of examples to train a final classifier to properly classify the "Do nothing", "Repair" and "Replace" classes. In the particular case of Western Power network, there are more than 600 thousand wood poles with an average age of 45 years and, likely, millions of inspections have been completed over time. In the ML community, there is the popular saying that more data beats complex algorithms [15], which truly well applies in this case.

Nevertheless, it is relevant to note that the "Do nothing", "Repair" and "Replace" classes are not well balanced, meaning the classes proportions are not similar (i.e. each class does not comprise of ~33% of the dataset). Over the life of a pole, there will be far more events of "Do nothing" than "Repair" and similarly, more "Repair" events than "Replace" ones. Although the training data might be abundant indeed, it is often difficult to achieve good classification performance for all the classes on a heavily unbalanced dataset [16, 17]. ML algorithms trained with

unbalanced data will tend to misclassify the less representative classes, which is of concern if these are "Replace" cases that had been misclassified as (e.g.) "Do nothing".

One might rightfully attempt to improve the final model classifier assertiveness by (e.g.) increasing the number of explanatory variables or through feature engineering—which typically requires some domain knowledge. For instance, the training dataset that will be used to train the final classifier could be augmented by having other explanatory input variables not yet discussed. Potential candidates could be (e.g.) Weibull distribution Failure Rates regarding wood species and/or geographical locations. It can be argued that the inputs to the Weibull distributions suggested are somewhat reflected by the existing variables already being used—like the geo-feature maps. Such a case will lead to the problem of Collinearity, meaning two or more variables being considered are highly correlated and hence, not much additional information could be achieved by having these additional variables. The Collinearity issue can be addressed by performing dimensionality reduction through different techniques, like e.g. Principal Component Analysis—PCA [8]. The "correct" number of explanatory variables is dependent on the user expectations of accuracy and it is a trial process that trades-off different desirable aspects. In general, having too many (too few) explanatory variables increase (decrease) computational power requirements and might increase (decrease) the model assertiveness. The final classifier dataflow structure is conceptually illustrated in Fig. 4. It can be seen that outputs from previous ML models are used as inputs to another ML layer. Also, the required iterative nature of the process is mentioned and typical classifier probabilistic outputs are shown.



Pole #	Probability of:		
	Do Nothing	Repair	Replace
1	0.2	0.6	0.2
2	0.6	0.1	0.2
3	0.6	0.2	0.1
4	0.8	0.2	0.0
5	0.6	0.2	0.3
6	0.7	0.3	0.0
7	0.5	0.4	0.1
8	0.5	0.5	0.0
9	0.9	0.1	0.1
10	0.6	0.3	0.1
11	0.4	0.1	0.5
12	0.3	0.4	0.3
13	0.8	0.1	0.1
14	0.8	0.0	0.2
15	0.1	0.7	0.2
16	0.2	0.5	0.2
17	0.5	0.4	0.2
18	0.6	0.3	0.1
19	1.0	0.0	0.0
20	0.5	0.4	0.1
21	0.6	0.3	0.1

Fig. 4 Schematic dataflow of final classifier. Different inputs, iterative ML classifier algorithm process, classifier outputs for individual poles

In the wood pole classification problem, it is obvious the adverse risk of misclassification, in particular not identifying the poles that require Repair or Replacement i.e. False Negatives (Error Type II) cases. It is known that the failure to identify wood poles that require immediate or more frequent attention had resulted in catastrophe in the past [18]. Such a risk can be minimized by calibrating the decision boundary threshold between the three classification classes as shown in Fig. 4—which inevitably would lead to more False Positive (Type I Error) outcomes and increased operational costs. The fine-tuning of the classification model output should be an iterative process driven by the risk appetite the operator and/or stakeholders are prepared to accept in line with their existing Asset Management Policies, Strategies and Plans.

Generally, ML models need to be continuously "re-trained" as more data becomes available over the time. This is the case because it is reasonable to assume the model's explanatory variables might not be Identically Distributed—meaning that their behaviour (as well as the whole asset system behaviour) changes as time elapses. For instance, the quality of chemical treatments on wood poles conducted over the past decade is better than those conducted prior to this period or the inspections during the 1980s were not as comprehensive as after the 2000s. These changes over the time obviously affect the asset system performance and therefore, the predictive model needs to evolve accordingly.

There are many other nuances to the topic that are not addressed here but it suffices to state that the development of successful Machine Learning applications requires a substantial amount of "black art" that is hard to find in textbooks [15]. In all, it can be said one has to trial different approaches to achieve adequate model's performance as each case has its own peculiarities.

5 Conclusion

This paper offers an overview of how the topic of machine learning could support maintenance decisions for the case of a wood poles power distribution network asset system. It demonstrates two applications of ML methods to extract information from disparate raw datasets and conceptually, how these could be later used as inputs for a final predictive classifier. It also discusses in general the importance of considering multiple variables concomitantly, which is indeed paramount to many problems involving physical asset systems. The principles herewith discussed would benefit the decision making processes for various other asset systems too and potential developments in this regard are many. In the particular context of Engineering Asset Management, this includes e.g. predicting process plant upsets by continuously monitoring and processing process historian databases.

References

1. Western Australia Auditor General's Report (2013) Western power's management of its wood pole assets. Report 17, Nov 2013
2. Ausgrid, Network Standard NW000-S0098 (2015) NS145 Pole Inspection and treatment Procedures, Sept 2015
3. ActewAGL (2014) Asset specific plan. Poles, May 2014
4. USA Bureau of reclamations (1992) Wood poles maintenance. Facilities instructions, standards & Techniques, vol 4–6, Aug 1992
5. Cesa A, Shen K (2015) Use of multi-folds asset operational performance data to prioritise business improvement actions: challenges and opportunities. Presentation delivered at the WA Engineers Australia, 13 Aug 2015
6. Payhak S (2011) Leakage current in wooden structures used for power distribution. PhD thesis, Engineering School, RMIT University, Mar 2011
7. Rahman A (2003) Modelling inground decay of wood poles for optimal maintenance decisions. Queensland University of Technology, Master of Engineering Thesis, Jan 2003
8. James G, Witten D, Hastie T, Tibshirani R (2014) An introduction to statistical learning: with applications in R. Springer Publishing Company, Incorporated
9. Liu B (2007) Web data mining: exploring hyperlinks, contents, and usage data. Springer Science & Business Media
10. Friedman J, Hastie T, Tibshirani R (2001) The elements of statistical learning, vol 1. Springer series in statistics, Springer, Berlin
11. Ng A (2015) Machine learning from Stanford University by Coursera. https://www.coursera.org/learn/machine-learning
12. Nielsen MA (2015) Neural networks and deep learning. Determination Press
13. Cesa A (2017) Matlab code repository for the ML cases presented on this paper. https://github.com/alecesa/Digit-recognition-ANN
14. Le Cun Y, Boser B, Denker JS, Howard RE, Habbard W, Jackel LD, Henderson D (1990) Handwritten digit recognition with a back-propagation network. In: Touretzky DS (ed) Advances in neural information processing systems 2. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp 396–404
15. Domingos P (2012) A few useful things to know about machine learning. Commun ACM 55 (10):78–87. http://dx.doi.org/10.1145/2347736.2347755
16. He H, Garcia EA (2009) Learning from imbalanced data. IEEE Trans Knowl Data Eng 21 (9):1263–1284. https://doi.org/10.1109/tkde.2008.239, http://dx.doi.org/10.1109/TKDE.2008.239
17. Seiffert C, Khoshgoftaar TM, Van Hulse J et al (2007) Mining data with rare events: a case study. In: IEEE international conference on tools with artificial intelligence, Patras, Greece, 2007, vol 2, pp 132–139
18. Electrical Incident Report (2009) Power line fault and bush fire near cape Naturaliste lighthouse Dunsborough. Western Power, Perth, Western Australia, 7th Feb 2009. https://www.commerce.wa.gov.au/sites/default/files/atoms/files/dunsborough_fire_report_feb_2009.pdf

Successful Organisational Development of Asset Management Organisations

Jasper L. Coetzee and Solly Nkosi

Abstract Maintenance/Asset Management organisations are often not functioning at an optimal level. The reason for this, apart from the obvious possibility of poor management, is frequently to be found in the communication gap between those who manage the organisation and those who perform the actual upkeep of the equipment. This phenomenon is also known as the maintenance/asset management strategic gap (Coetzee in Maintenex conference, Johannesburg, [7]). This circumstance is created by the gap in tuition/training of these individuals, their difference in world view, difference in actual/perceived status, and internal organisational politics. But, whatever the reason(s), the fact is that this gap exists and causes difficulty in driving through positive change, and often results in a poor organisational culture. Mostly, this situation leads to lower than optimal availability and reliability of the organisation's production system. This then leads to relatively low production and high maintenance cost, with the consequence of lower than possible profits being achieved. This paper discusses a methodology for rectifying this situation. This method, although being developed and taught from the late '90s, and parts of it being tested in the early 2000s, has only relatively recently been afforded the opportunity of being tested fully at a production organisation, DMS Powders, in South Africa (Coetzee in Maintenance Forum, [5]). This paper expounds the method, and adds a short case study of the process and the results achieved.

J. L. Coetzee (✉)
University of Pretoria, Pretoria, South Africa
e-mail: jasper.coetzee@up.ac.za

S. Nkosi
DMS Powders, Meyerton, South Africa
e-mail: solly.nkosi@dmspowders.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_10

1 Introduction

Most Asset Management/Maintenance Departments in industrial organisations struggle with achieving excellent results, due to the complexity of the function. This was highlighted when the commission of enquiry into the poor results/high cost of maintenance in Britain in the 60s suggested widening the scope of the function to the full equipment lifecycle. Their suggestion of incorporating a Terotechnological Department in such organisations did not gain traction then, but was the spark that eventually led to the creation of ISO 55000.

It is important to note that the upkeep of equipment can only be successful if the total equipment management and maintenance function functions in total harmony with the objectives to be achieved through the equipment. This is not at all easy to achieve.

The complexity of the manufacturing systems and the techniques employed in manufacturing/production organisations requires an increasingly sophisticated approach to the management of the asset management function. At the same time there is a world-wide trend towards a widening in the gap between maintenance management and maintenance operations, which leads to the maintenance function not performing at the level which is required for long term business success.

There is thus an urgent need for bridging this strategic gap by strong maintenance strategic leadership, which promotes top to bottom team forming in the business to achieve results that are commensurate with the highly developed technological, procedural, and systems base of the Asset Management function.

The solution to this is often sought through moving the maintenance organisation to world-class on a step-by-step basis [1, 2, 12]. This paper suggests and illustrates the results obtained using an alternative stepped approach. This is effected by recognising that the organisation, like the people of which it consist, develops more naturally through a paced learning process than through a step-by-step technique by technique implementation.

2 The Asset Management Strategic Gap

The functioning of the Asset Management organisation can be understood as the combination of a few control cycles. In particular, this includes a strategic cycle, a tactical cycle, and a work control cycle in the model known as the Maintenance Cycle model [8]. This is shown in Fig. 1.

Following from the Maintenance Cycle model, Fig. 2 shows a diagrammatic representation of the order in which the main components of the maintenance cycle model should be implemented. It particularly identifies the dependence of the main components on each another. The bottom-most component (the execution of maintenance work) is the end result that should be achieved by the maintenance organisation. But this (bottom-most component) is dependent on each of the components above it.

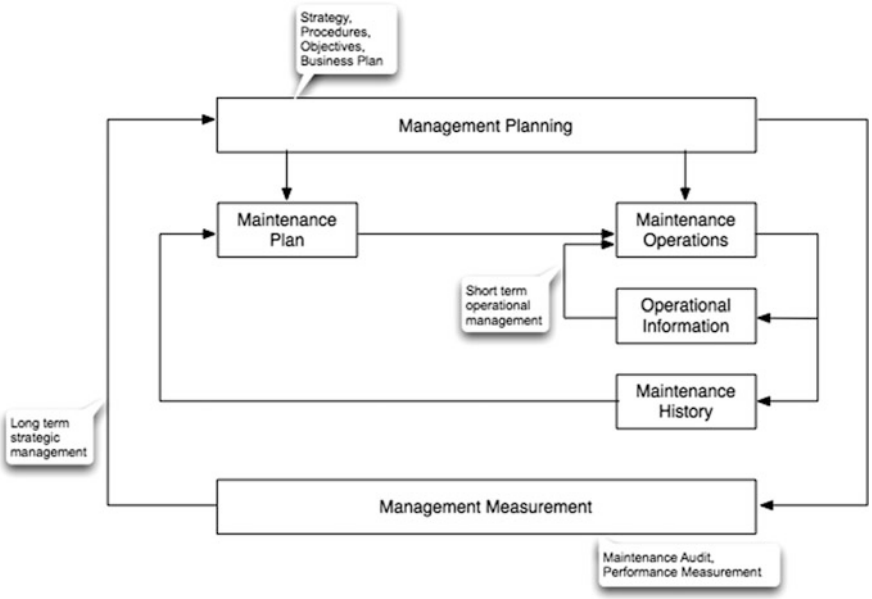


Fig. 1 The maintenance cycle [8]

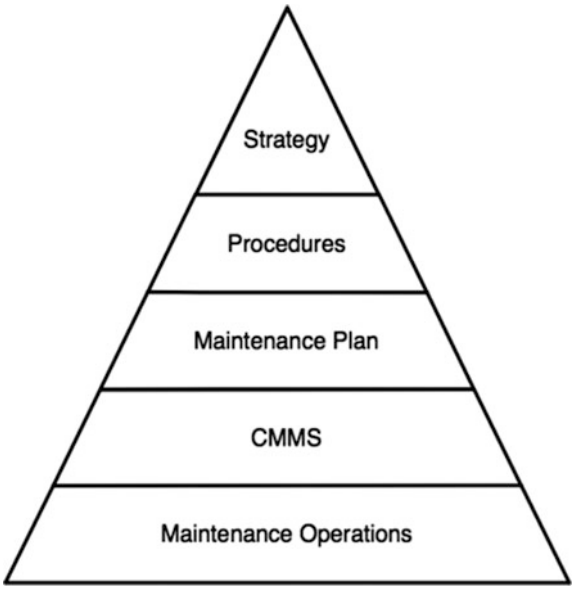


Fig. 2 Strategy implementation triangle [4]

In particular, the most important (and foundational) component is the maintenance strategy, which dictates what procedures should exist and to which the maintenance plan is sub-ordinate, and so forth. The maintenance plan is then driven through the CMMS to identify the work to be done by Maintenance Operations.

Coetzee [7] pointed out that the maintenance organisation is an organism of which the various parts must function in full harmony towards the achievement of a maximum contribution to the success of the business. He also stated that such harmony cannot be achieved by only implementing highly sophisticated (and localised) solutions to problems experienced in sub-parts of this organism. Examples of such localised solutions are RCM, or a new CMMS.

Nevertheless, it is very tempting to address only one or two of the levels in the triangle, which are perceived to be under-developed in some or other way. But this ignores the holistic nature of the maintenance organisation, and inevitably does not result in the benefits sought after.

If the Asset Management organisation understands these principles, all this might work well apart from one commonly experienced problem: the Strategic Gap. If one investigates the reasons for non-success of Asset Management organisations, typically through an audit, one almost invariably realises that there is something fundamentally wrong in the way in which the organisation communicates up and down the hierarchy, in the fabric of cohesion of the perceived strategic direction of the organisation. This is the Strategic Gap [7], which is shown in Fig. 3.

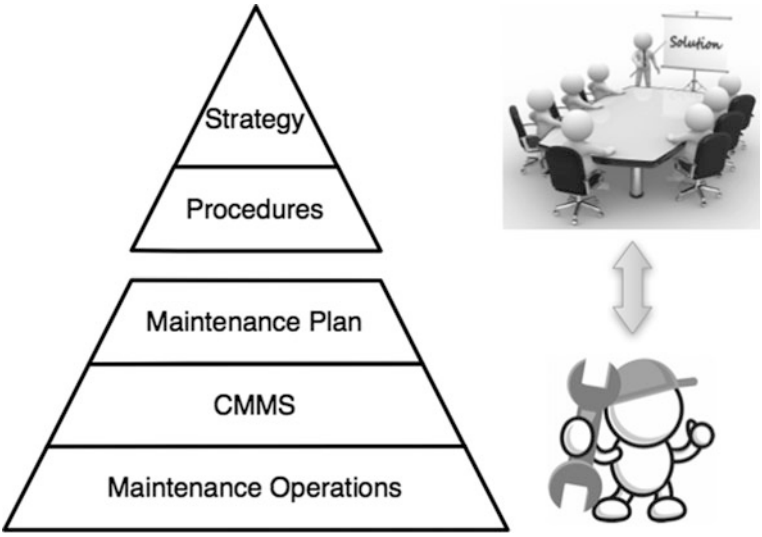


Fig. 3 Strategic gap [7]

This strategic gap shows itself in the organisation through the following symptoms, which gives a clue regarding the source of the gap:

Managers' view	Supervisors/Workers' view
Supervisors/Workers are not open for their ideas	Management are in an ivory tower—totally divorced from reality
Want to implement new concepts	Want to get on with the job
Arrange meetings to effect change, but without much result	Go with the flow, but keep on doing the job as they know best

3 Implications of Traditional Maintenance Practice

The leaders in the field of Asset Management are engineers of one or other discipline. They understand the technological background of that specific discipline well, but are typically no experts in understanding the failure mechanisms of the equipment under their care. Furthermore, they are mostly also not well equipped for their role as the leader of people. All this lead to the typical Asset Management/Maintenance organisation at best delivering under-optimised results. This in turn causes the production organisation not achieving the expected profitability.

A major problem is that many of these leaders do not recognise their lack of understanding of these issues. They are thus blind leaders of the blind, as is so often evident when performing audits.

4 Attempts at Improvement

Many consulting organisations realise that the issue of achieving excellence in Asset Management is to be addressed bit by bit. The Maintenance Function (and even more so for a fully developed Asset Management function) is an intricate function, having to apply technology driven solutions through people, using elaborate procedures and logistics. You just cannot change a poorly performing Asset Management function to a World Class one with a few strokes of a magic wand.

Therefore, the logical answer is to divide the task into manageable parts, and implement these parts one by one. However, the problem in all this is which part should go first, which second, and so on, to achieve success.

It would thus seem as if the prevalent thinking in organisational development in becoming World Class involves addressing one by one facet at a time.

One example of such an approach is that of the Physical Asset Management Practice of Coopers & Lybrand (later PricewaterhouseCoopers) under the leadership of Campbell [2]. This is a four stage approach, starting by developing leadership in the organisation, followed by developing control measures, then

continuous improvement through the application of RCM and TPM, and finally by taking quantum leaps using Process Reengineering.

After the death of John Campbell, this approach of PwC was modified to a three step model [1], where the first step Develop Leadership is followed by putting in place a variety of System Essentials, followed by the last step Choosing Excellence using RCM, Continuous Improvement, and Evidence Based Asset Management (EBAM).

Another popular approach is to start with Basic Maintenance, and then progress through Proactive Maintenance towards Strategic Management of the function, and then on to World Class. One well-known organisation practicing this type of approach is SAMI (Strategic Asset Management, Inc.), with the following 5 stage development path [12]:

Stage 1: Planned Maintenance
Stage 2: Proactive Maintenance
Stage 3: Organisational Excellence
Stage 4: Engineered Reliability
Stage 5: Performance Culture

All these approaches can possibly lead to success, depending on exactly how it is applied. However, because success is dependent on how well one gets the total maintenance team aligned and activated, the amount of energy that goes into persuading all individuals to agree with the change of direction and supporting the team effort wholeheartedly can be immense.

Whichever method is used, success can only be achieved through a holistic approach [8]. The total Asset Management organisation is an organism consisting of people with individual opinions, implementing a myriad of technical solutions/processes, and utilising many different management methods/procedures/systems. Such system must find an equilibrium state, with agreement and cooperation between all the role players.

5 Measuring the Outcome of Improvement Drives

One of the important tools in improvement drives is a proper measuring instrument. This mostly takes the form of an audit of some or other kind. While such audit can be devised from different perspectives on the Asset Management function, Table 1 gives a short synopsis of the auditing scheme used for the process reported on in this paper.

Table 1 Auditing scheme employed

Main measurement area	Sub-measures
Results achieved	Availability of plant/equipment, reliability, operability, quality, equipment condition, safety, cost, client satisfaction
Strategic excellence	Asset management strategy, procedures, objectives, business plan, performance measurement, auditing
Tactical excellence	Maintenance plan, maintenance planning, maintenance supervision, quality of execution, operational purposefulness
Systems	Systems driven management, systemised processes, CMMS, management reporting, effective use of reporting
Asset management logistics	Work place design, supply of spares, rotables management, maintenance equipment, tools, manuals, fault tracing support, document management, classes of personnel employed
Asset management technology	Condition monitoring, test facilities, fault tracing/diagnostics, machining facilities, servicing facilities
People centredness	People's commitment, training relevance/adequacy, experience relevance/adequacy, training matrix per post type, depth of technical expertise, alignment of personal and business objectives, team functioning
Leadership	Leadership programmes, managerial acumen, management focus, management control
Asset management improvement	Furthering of learning, furthering of innovation, learning through failure, use of RCFA, problem solving competence
Asset management culture	Personnel attitudes, motivation, confidence in business processes, responsible and transparent communication, cooperative relationships, trust, rework, low personnel turnover

6 Components of the Process

The idea of a learning organisation was coined by Senge [14] in 1990. The learning organisation is one "where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured …" [14]. The learning meant are what Luthans and Youssef [10] call *Tacit Knowledge*. It is not the knowledge that you acquire in the formal education system, or even the experience that you have before joining the organisation, but it is the knowledge that the people of the organisation build themselves as part of the social structure of the organisation. It is the knowledge built by the teams in the organisation as they interact with each other. The process of building tacit knowledge is slow, but it defines the culture of the organisation and its success if it is led in the right direction.

The core of the phase by phase methodology put forward in this paper was originally published in 2005 [6]. It is based on the principle that an organisation consists of people, and people need to be led slowly towards building new tacit knowledge that defines a new winning culture for the organisation. Technical and

systems development is seen as part of this process of the development of tacit knowledge and not as separate "technical only" developments. The point is that the total success of the organisation depends on people, so they must be led to believe in the organisation's goals and its processes. And this is a slow process. To achieve this, two aspects need to be driven, using a phase-by-phase approach, namely the content and the depth at which the content is developed and implemented in the organisation.

The content elements (leftmost two columns of Table 2) were defined to be:

1. Analysis: The improvement process is dependent on regular measurement and evaluation. Each phase of the process starts and ends with an audit (see Table 1). In parallel with this a SWOT analysis is performed where the personnel of the Asset Management organisation evaluates themselves. These analyses are used to focus the whole process on the most critical aspects identified by the outside auditors, and the personnel of the organisation.
2. Strategic Level aspects. These include the Asset Management Strategy, Management Procedures, Objectives, Business Plan, Performance Measurement process, and the Audit instrument. These aspects all follow from the outer (strategic) loop of the Maintenance Cycle, shown in Fig. 1.
3. Tactical Level aspects. These include the Maintenance Plan, the CMMS, Supervision, and Planning aspects. These all follow from the inner (tactical) loop of the Maintenance Cycle, shown in Fig. 1.
4. Business Focus aspects. The business focus includes the Maintenance Focus, Management Focus, and Business Culture.
5. Training.

The depth of application is increased during the three phases of the use of the methodology. These are shown as the three levels in Table 2.

The three phases identified in Table 2, which are each nominally applied over a period of one year, are:

1. Basic: set the scene for the later phases, lie the foundations. See the descriptions per the various content aspects in this column.
2. Intermediary: build on the successes of the basic phase, by implementing a "60%" solution.
3. Advanced: refine to a full level of Maintenance Excellence—goal is 80% of World Class.

The result is developed cumulatively as illustrated in Fig. 4. Each phase addresses all 5 the content elements listed above, but at progressively increased intensity, until the total result on the right is achieved.

Two important aspects which are central to the successful development of the Asset Management organisation, per the scheme above, regards the closing of the strategic gap, and building a new organisational paradigm involving all the people in the organisation.

Table 2 Phased approach to Asset Management Improvement projects

Aspect	Strategic element	Level		
		Basic	Intermediary	Advanced
Analysis	Audit	Main areas of shortcomings	More deeply entrenched 'sins'	Show 'societal' problems
	SWOT Analysis	First identification	Refine SWOT	'Know yourself' (without excuses)
Strategic level	Strategy (or policy)	Philosophy, mission, vision, basic elements	80% Strategy	Final Strategy
	Management procedures		20% of procedures (based on Pareto)	Detailed procedures
	Objectives		Basic annual objectives	Detail annual objectives
	Business plan		Basic plan with budget, capital plan, facilities plan	Five year Business Plan, detail annual plan, zero based
	Performance measurement		Identify, measure '20%' indices	Integrated Performance Measurement instrument
	Audit instrument		Basic instrument, use annually	Detailed instrument, use annually
Tactical level	Maintenance plan	ID critical MSI's, start history collection	FMEA/FMECA for MSI's, ID critical Failure Modes	Detailed RCM for critical Failure Modes
	CMMS	Increase width of use of present CMMS	Wide/deep use of CMMS	Changes to the CMMS. New tasks from RCM
	Supervision	Supervisory training	Assimilation of supervisory training	High interaction with planners and production
	Planning	Planning training	Optimise use of planning principles	Advanced planning principles, improve interactivity
Business focus	Maintenance focus	Basic care, focus on MSI's	CBM, focus on Failure Modes (FM's)	Proactive maintenance, including CBM, UBM, CI
	Management focus	Manage strategy elements, focus on MSI's	Manage objectives, business plan, focus on FM's	Manage strategically, focus on Excellence
	Business culture	Culture assessment, develop team culture	Team forming, find solutions to problems	Reduction of variability; Top to bottom team-forming

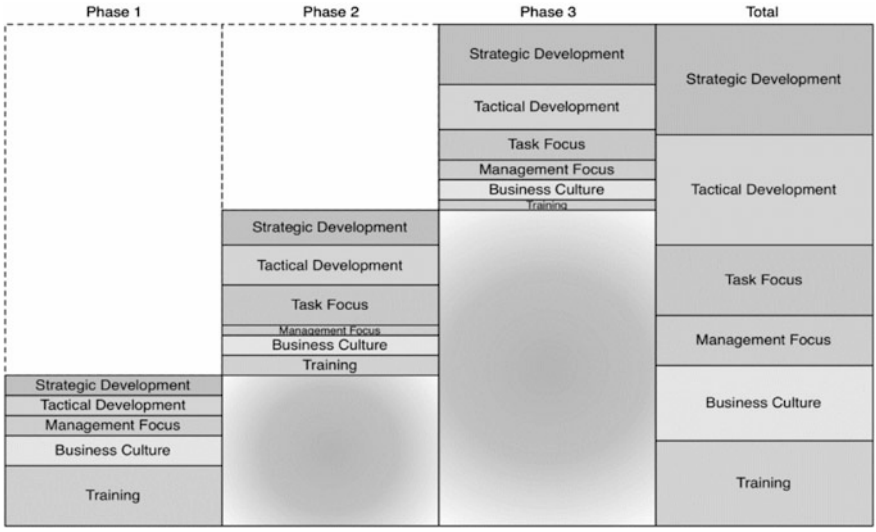


Fig. 4 Phased development during Asset Management Improvement

The solution is to take a cross section (the strategically most dominant one) of the maintenance organisation and address all these aspects (the ones listed in Table 2) simultaneously. The idea is to take a slice from the organisation (engineer, supervisor(s), planner(s), artisans) and build the new business culture in that slice, grappling with some of the most critical issues in the organisation. In the application at DMS Powders, the whole maintenance organisation was segmented into 4 such slices, and each group during the first phase tasked to analyse, using Root Cause Analysis, one of the critical management weaknesses they identified in the SWOT analysis.

This process of inter-level therapy is based on the principle of Top to Bottom Team Forming [7], which brings all levels together in a single team to discuss mutually important issues and find solutions for them. In this process the engineer finds that artisans have good inputs and can add much to make his/her own ideas even better. Also, the artisan finds that the engineer does have good, practical ideas that are worth pursuing. This process also brings about a relationship between leader and follower, which is conducive to the engineer being able to lead the organisation from the front [9], with the followers willingly following.

The result of this all is that the strategic gap is removed, the communication from top to bottom and bottom to top is improved substantially, and a new, common organisational paradigm (culture) is built, which materially improves the functioning and profitability of the organisation. An added benefit is that it contributed to DMS Powders' objective of using effective teams and team synergy towards long term sustainability [11].

7 Successful Application Example

DMS Powders in Meyerton, Gauteng, South Africa, produces and markets ferrosilicon powders for application in dense media separation (DMS) technology. Dense Media Separation is a process involving the suspension of high-density powder in water to form a dense media, which is used to separate heavy mineral particles from lighter non-mineral particles in a sink-float process.

The company outsourced its maintenance functions over a number of years, which led to the maintenance function progressively falling back into pure fire-fighting. Furthermore, a very poor business culture, as well as very poor maintenance outcomes resulted, with a commensurate decline in business profits.

M-Tech Consulting Engineers, a company specialising in Asset Management Strategic Consulting, were approached to implement a maintenance improvement project to rebuild the lost organisational capability, and improve the business culture of the maintenance department [3].

The project ran over a five year period (middle 2011 to middle 2016), longer than the nominal three year period, mainly due to DMS organisational constraints.

This resulted in a huge improvement in the audit outcomes of the department (from 31% of World Class before the project started to 48% at the end of the first phase, 61% at the end of the second phase, and 75% at the end of phase 3). The second last statistic was corroborated by an independent external auditor, while the last was the result of a totally independent ISO 55000 audit.

DMS Powders experienced a 16% increase in income over this period (adjusted for the Production Price Index, based on 2011 = 100%). All of this cannot be attributed to the Maintenance Improvement Project alone, but it played a very significant part in the improvement achieved.

The results of the Maintenance Improvement Project also included a reduction in production resource consumption [13]. A research project at the University of Pretoria, submitted as a Master's Thesis, used DMS Powder's actual data to prove a reduction in the use of raw material, and including power, water, and production time inputs as the maintenance situation improved.

This, in the opinion of the authors, conclusively proved the experiential concepts as put forward by various authors (notably [1, 2, 4, 6, 9]) in Sects. 2 and 4 above.

References

1. Campbell JD, Reyes-Picknell JV (2015) Uptime: strategies for excellence in maintenance management, 3rd edn. Productivity Press
2. Campbell JD (1995) Uptime: strategies for excellence in maintenance management. Productivity Press
3. Coetzee GM (2016) Case study: maintenance improvement project reduces resource usage. Vector, Nov/Dec 2016
4. Coetzee JL (1999) A holistic approach to the maintenance "problem". J Qual Maint Eng 5(3)

5. Coetzee JL (2015) Maintenance improvement process—a phased approach to reach a World Class Maintenance standard. Maintenance Forum
6. Coetzee JL (2005) Maintenance strategic management—a prerequisite for business survival. IMEC 2005, Toronto
7. Coetzee JL (2000) Maintenance success using a holistic approach. In: Maintenex Conference, Johannesburg
8. Coetzee JL (2004) Maintenance. Maintenance Publishers
9. Coetzee JL (2006) The strategic role of leadership for maintenance executives. PMME 2006, Melbourne, Australia
10. Luthans F, Youssef CM (2004) Human, social, and now positive psychological capital management: investing in people for competitive advantage. Organ Dyn 33(2)
11. Nkosi S (2015) Triple bottom line legacy. Australia
12. Peterson SB. The future of asset management. ReliablePlant.com
13. Rahiman MA, Coetzee JL (2014) Resource consumption as a maintenance performance metric. In: World congress on engineering asset management
14. Senge PM (1990) The fifth discipline: the art and practice of the learning organization. Doubleday

Combining Reliability Assessment with Maintenance Performance Analysis Using GAMM

Adolfo Crespo Márquez, Luis Barbera Mártinez, K. A. H. Kobbacy, Antonio Sola Rosique, Antonio Guillén López, Antonio De la Fuente Carmona and Asier Erguido

Abstract GAMM (Graphical Analysis for Maintenance Management) is a method that supports decision-making in maintenance management through the visualization and graphical analysis of reliability data. One of the most important features of GAMM is that fosters the combination of reliability assessment with the current asset performance analysis. As a basis for reliability analysis, the GAMM method uses a nonparametric estimator of the reliability function using historical data, sometimes in very limited amounts. For successful results, experience and knowledge in maintenance management are strictly necessary. At the same time, experience shows that the method may become really amicable for managers due to clear visualization and by using a set of basic rules. The analysis can be made per category of assets, per asset, per maintenance skill or per failure mode. In this paper we present certain rules to be applied when using GAMM with the intention to show the capabilities of the method and the possibilities that offers for practical maintenance and asset performance analysis.

A. Crespo Márquez (✉) · L. Barbera Mártinez · A. Sola Rosique · A. Guillén López · A. De la Fuente Carmona
Department Industrial Management, School of Engineering, University of Seville, Seville, Spain
e-mail: adolfo@us.es

A. Sola Rosique
e-mail: asrasrasr@telefonica.net

A. Guillén López
e-mail: ajguillen@us.es

K. A. H. Kobbacy
Scientific Chair for Operation and Maintenance Technology, College of Engineering, Taibah University, Medina, Saudi Arabia

A. Erguido
IK4-Ikerlan Technology Research Centre, Operation and Maintenance Area, Mondragón, Spain
e-mail: aerguido@ikerlan.es

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_11

1 Introduction

The importance of the maintenance function management has grown in the past decade as a result of the continuous advancement of industrial organizations towards more efficient operations and improved productivity [1]. The basic concept leading maintenance engineering is the maintenance management process continuous improvement by incorporating knowledge, intelligence and analysis. Maintenance engineering supports decision-making and fosters maintenance effectiveness, efficiency and operational result [2].

The GAMM (Graphical Analysis for Maintenance Management) method is a graphical tool supported by statistical analysis of collected historical data, related to the sequence of past equipment technical revisions [3–5]. The method provides visual information allowing managers to recognize interesting equipment behavior patterns and supporting them to make management decisions accordingly. In order to do so, GAMM collects valuable information such as: number of technical revisions performed, type (corrective or preventive), duration, quality of the maintenance service, reliability of the equipment before the service, and/or equipment failure's variability over a period of time. Moreover, GAMM provides a visual representation from a complete or partial historical record of the performed maintenance work, showing different patterns of analysis and presenting useful information for decision making and problem solving (The reader is addressed to other referred papers published by authors of the method in 2013, included in the references, those are papers providing examples and case studies using this methodology for specific maintenance management problems).

In order to facilitate and support maintenance decision making for asset management, this contribution proposes a set of basic rules that maintenance engineers and/or managers may consider when using GAMM for preventive maintenance program evaluation. These rules help and speed up the process of understanding the implications of results presented in the two GAMM method graphs: Graphic 1 and Graphic 2 (See Figs. 1 and 2). The set of rules provides easy access to certain variables patterns showing useful information for maintenance management and decision making in the short, medium and long term. Therefore, we can consider this paper as a complementary research of the GAMM method, as Barberá et al. published it in 2013.

2 A Brief Review of the GAMM Method

The Graphical Analysis for Maintenance Management (GAMM) is a quantitative and qualitative analysis method that intends to support decision making in maintenance management by using the logic of ad hoc designed dispersion diagrams and by conveniently processing equipment reliability and maintainability data [6]. Input data for GAMM relates to the history of equipment maintenance interventions: the

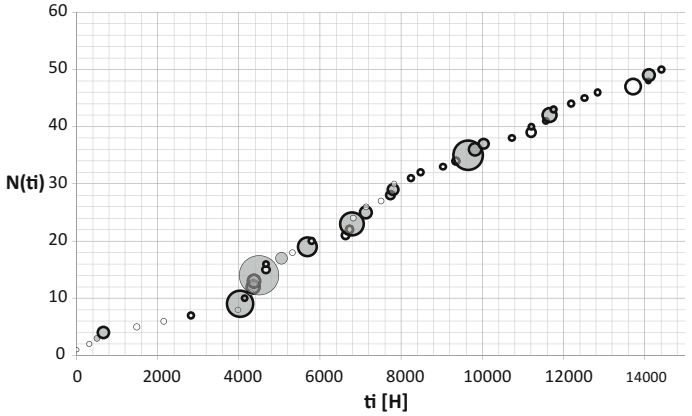


Fig. 1 Example of the graphs generated by GAMM (Graphic 1), that shows the number of interventions over time

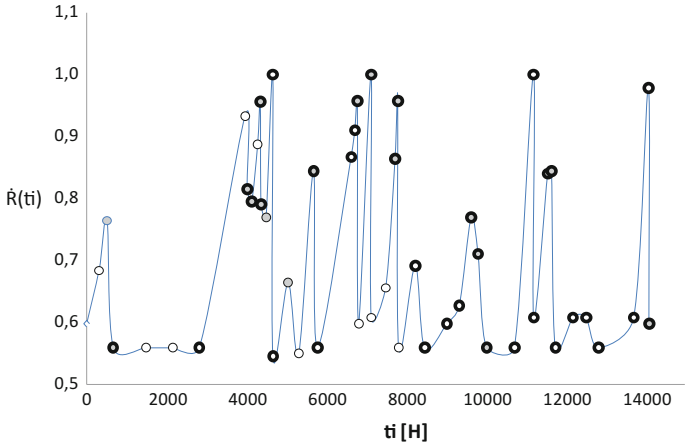


Fig. 2 Example of the graph generated by GAMM (Graphic 2), that shows estimated reliability at the moment of the different interventions

operating time when they took place, the type of maintenance intervention (preventive/corrective), the intervention duration time and the state of equipment/system during the intervention. Using this data, GAMM first plots a detailed accumulated histogram of interventions (in graphic 1, where N(t) is the number of interventions). Then, the equipment reliability function is estimated using Nelson-Aalen nonparametric method's algorithms. Finally, this function is evaluated just in the moment before each maintenance intervention and then plotted (in Graphic 2). Graphic display uses a color code for the type of maintenance and a line code for the equipment status (in both graphics). The estimation method can

consider both the historic full data and the censored historical data. The combination of these variables with the sequence of interventions graphics display (Graphic 1 and 2) generates enormous synergies with regard to the information provided by the diagram, thus establishing new sources of analysis.

GAMM provides useful information regarding systems' reliability and maintainability analysis. This broadens the spectrum display for maintenance management, adding new evaluation parameters and graphically determining possible areas of improvement. The graphics of GAMM illustrate aspects such as:

- Trends in the behavior of the interventions.
- Preventive maintenance (PM) program deviation: By plotting the type of intervention accomplished over time we can appreciate how far is real PM schedule from the one scheduled in the CMMS. Therefore, we can track and monitor attainment to pre-established PM frequencies.
- Quality of operations and/or quality of the PM program: a graphic display of interventions by type (preventive or corrective) provides also a valuable data sequence for the recognition of certain behavior patterns concerning systems operation and maintenance. For instance, by observing corrective interventions within preventive ones, high sequences of corrective short term stoppages or accumulation of corrective interventions immediately after a PM one, may help to identify the existence of problems when operating the equipment or with the quality of the maintenance activities, respectively.
- Reliability function: GAMM shows the reliability of the equipment/system just before the maintenance activity. This can be worthy additional information to identify the existence of corrective task during operating periods when we expect high levels of system reliability (indicating abnormal premature system failure).
- Efficiency and quality in the implementation of interventions: the duration of each intervention is represented by the size of the bubble in GAMM (Graphic 1). This variable identifies those interventions lasting longer than the expected execution time, questioning whether our work scheduling and preparation was correct, or just if this was a consequence of other factors out of our control.
- Impact on production: The variable *Det* is used to capture the status of the equipment/system when maintenance is executed. The bubbles with hard shaded border represent interventions in which the equipment must stop. This fact will be affecting their individual availability and possibly impacting overall system/plant availability.

In summary, GAMM is a quantitative and also qualitative analysis, which intends to support decision-making in maintenance management [3–5], it may help to improve dependability and to reduce unavailability costs of the systems as well as other costs associated with unforeseen maintenance interventions

3 Developing Rules to Apply GAMM Results

The following rules, included in Table 1, provide easy access to envision certain variables patterns showing useful information for maintenance management and decision making in the short, medium and long term. Then, rule-based reasoning can be applied for a more consistent GAMM implementation. The rules have been developed by looking for patterns in different data sets from different equipment and considering all aspect of the GAMM method, such as: relative numbers of CO

Table 1 Rules to apply GAMM method results

Rule 1	**If** the Graphic 1 shows a line of $N(t_i)$ (number of interventions) versus t (time) which has approximately constant slope and $R(t_i)$ ranging between $R(t_i)$ max = 1.00 and $R(t_i)$ min = 0.70, **then** you have an acceptable[R1.1] maintenance program. **If** the line of $N(t_i)$ vs t has not a constant slope or $R(t_i)$ has sudden or large fluctuations, **then** the maintenance program is not optimal
	R1.1: regular number of maintenance events over an increasing operating time which means that a certain average reliability over time is guaranteed thanks to your maintenance activities
Rule 2	**If** number of PM[a] interventions >80% of the total interventions ($N(t_i)$), **then** it could be considered as an acceptable[R2.1] level of PM. **If** number of PM interventions >95%, **then** increase PM interval could be considered[R2.2]. **If** number of PM interventions <80%, **then** it could be necessary to evaluate the quality and frequency of the PM interventions. Decrease PM interval could be considered[R2.3]
	R2.1: according to many industries best practices. However we should check the effects of existing failures and how they are distributed just in case they may become critical (size of the circle) **R2.2**: Number of PM interventions >95% could also be a consequence of excessive maintenance to some non-critical failure modes **R2.3**: Number of preventive interventions below 80% are scenarios showing room for potential improvement of PM policies
Rule 3	**If** there are CO[b] interventions carried out in periods of high reliability ($R(t_i) > 0.85$) or shortly after a PM intervention (less time than the average duration between PM interventions), **then** the PM task before that CO intervention could be poor (problems in the execution of PM intervention) or the failure could be the result of improper equipment operation, among others, and it is necessary to evaluate it
Rule 4	**If** there is a logic/constant pattern of PM durations[R4.1] (downtimes, according to the tasks carried out on each intervention), **then** the PM program duration shows consistent maintainability. **If** there is a high variability of PM durations, **then** deviations in execution time due to external factors (lack of tools, waiting for parts, lack of staff training, accessibility problems, etc.) should be checked
	R4.1: However, in some cases there are two or more different patterns of PM duration due to, e.g., overhauls
Rule 5	**If** the temporal pattern of time between PM interventions is not constant, **then** the maintenance program is probably not well scheduled or its fulfillment should be reviewed

[a]Preventive maintenance
[b]Corrective maintenance

and PM events, range of PM intervals, intervention durations and reliability before each intervention, among others.

The development process of the GAMM rules is shown in Fig. 3.

This set of rules has been developed in order to get an acceptable performance of GAMM method, from different industrial environment. Therefore, these rules should be considered as a support for applying GAMM properly. Nevertheless, some numerical limits such as reliability range or number of PM interventions should be redefined considering each particular industrial context (e.g. aeronautical sector). It is common in rule-based expert systems that rules may require subjective judgement to be made. For example rules containing conditions such as "sudden or large fluctuations" in rule 1 and "high variability" in rule 4 will require establishing an arbitrary level to help quantifying such fluctuations or variation. This subjective judgement is a common feature of expert systems [7]. Such qualitative judgement may lead to overlap between the rules leading to multiple decisions for same case. Typically an expert system must be tested using large-scale data to establish the extent of overlapping.

4 Validating the Rules

Validation aims at substantiating that GAMM rule-based reasoning performs with an acceptable level of accuracy. There are different validation methods in the literature. A review of qualitative and quantitative validation methods of expert systems can be found in O'Keefe et al. [8]. In this study the validation of GAMM rule base will adapt the typical approach adopted by applying test cases through a system and comparing results against known results or expert opinion. To that end, the results of selected cases where the GAMM rules are applied will be compared with the recommendations generated by HIMOS (an intelligent decision support system—IDSS—developed for decision support, within the industrial maintenance

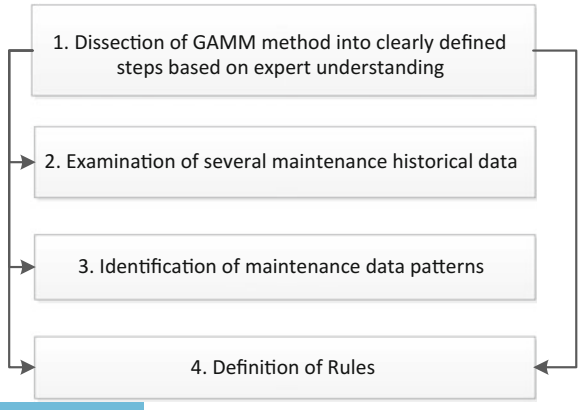


Fig. 3 Development of GAMM rules

field, in the University of Salford [1]). A total of 5 sets of historical maintenance data sets of equipment used in petrochemical industry. These five data sets were analyzed by an expert panel consisting of five experts in maintenance modeling as mentioned in [1] and at the same time they were uploaded and processed by HIMOS with the aim of comparing their recommendations of appropriate maintenance policies for each data set. In what follows the recommendation provided by the GAMM defined rules in (Table 2) will be compared with that of the published HIMOS for same data sets (see [1]. In Figs. 4, 5, 6, 7, and 8) below the graphical output from GAMM method for the five data sets are shown. The 5 data sets are published in [1] as data sets 1, 3, 5, 6 and 7. These figures provide confirmation of the GAMM Rule Base recommendations. In Table 2 below the outcome of applying GAMM rule base to each data set is compared with the validated output from HIMOS.

Table 2 Comparison HIMOS versus GAMM rules recommendations for 5 data sets

Data set 1	
HIMOS	Increase PM interval from 135 to 176 days (CBR-Geometric I model)
GAMM rules	R(ti) has sudden or large fluctuations: maintenance program is not optimal. Number of PM interventions >95%: increase PM interval could be considered. There is a logic/constant pattern of PM durations (downtimes, according to the tasks carried out on each intervention): PM program duration shows consistent maintainability. The temporal pattern of time between PM interventions is not constant: maintenance program should be reviewed
Data set 2	
HIMOS	Increase PM interval from 177 to 403 days (CBR-RPOW model)
GAMM rules	The line of N(ti) vs t has not a constant slope and R(ti) has sudden or large fluctuations: maintenance program is not optimal. Number of PM interventions >95%: increase PM interval could be considered. There is a logic/constant pattern of PM durations (downtimes, according to the tasks carried out on each intervention): PM program duration shows consistent maintainability. The temporal pattern of time between PM interventions is not constant: maintenance program should be reviewed
Data set 3	
HIMOS	Increase PM interval from 159 to 223 days for the first sequence but gradually decrease to 86 days (tenth sequence) (NHPPScoSpm model)
GAMM rules	R(ti) has sudden or large fluctuations: maintenance program is not optimal. Number of PM interventions >95%: increase PM interval could be considered. There are CO interventions carried out in periods of high reliability (R (ti) > 0.85) and also shortly after a PM intervention (less time than the average duration between PM interventions): PM interventions before that CO interventions could be poor (problems in the execution of PM intervention) or the failure could be result of improper equipment operation and it is necessary to evaluate them. There is a logic/constant pattern of PM durations (downtimes, according to the tasks carried out on each intervention): PM program duration shows consistent maintainability. The temporal pattern of time between PM interventions is not constant: maintenance program should be reviewed

(continued)

Table 2 (continued)

Data set 4	
HIMOS	Increase PM interval from 108 to 140 (first sequence)–117 (tenth sequence) days (RPDWSco)
GAMM rules	R(ti) has sudden or large fluctuations: maintenance program is not optimal. Number of PM interventions > 95%: increase PM interval could be considered. There are CO interventions carried out shortly after a PM intervention (less time than the average duration between PM interventions): PM task before that CO intervention could be poor (problems in the execution of PM intervention) or the failure could be the result of, improper operation of equipment and it is necessary to evaluate them. There is a logic/constant pattern of PM durations (downtimes): PM program duration shows consistent maintainability. The temporal pattern of time between PM interventions is not constant: maintenance program should be reviewed
Data set 5	
HIMOS	Decrease PM interval from 103 to 21 days (Delay-time Model, Weibull distribution, manual selection mode)
GAMM rules	Graphic 1 shows a line of N(ti) (number of interventions) vs t (time) which has partially constant slope and R(ti) has sudden or large fluctuations: maintenance program is not optimal. Number of PM interventions <80%: it could be necessary to evaluate the quality and frequency of the PM interventions. Decrease PM interval could be considered. There are CO interventions carried out in periods of high reliability (R(ti) > 0.85) and shortly after a PM intervention (less time than the average duration between PM interventions): PM task before that CO intervention could be poor (problems in the execution of PM intervention) or the failure could be the result of improper operation of equipment and it is necessary to evaluate them. There is a high variability of PM durations: deviations in execution time due to external factors (lack of tools, waiting for parts, lack of staff training, accessibility problems, etc.) should be checked. The temporal pattern of time between PM interventions is not constant: maintenance program should be reviewed

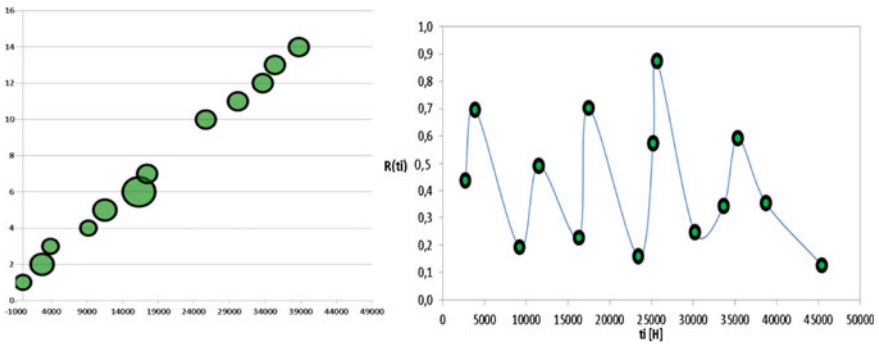


Fig. 4 GAMM graphics using data set 1

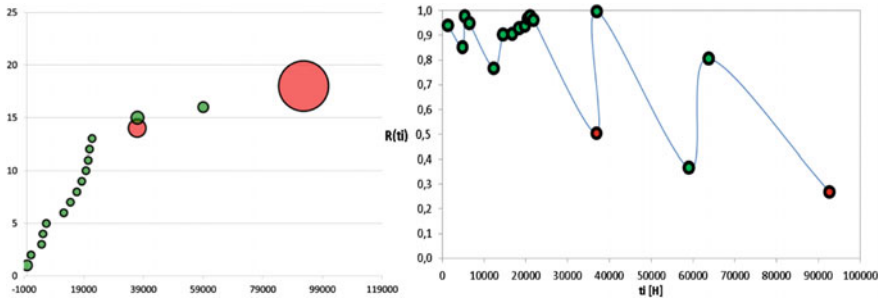


Fig. 5 GAMM graphics using data set 2

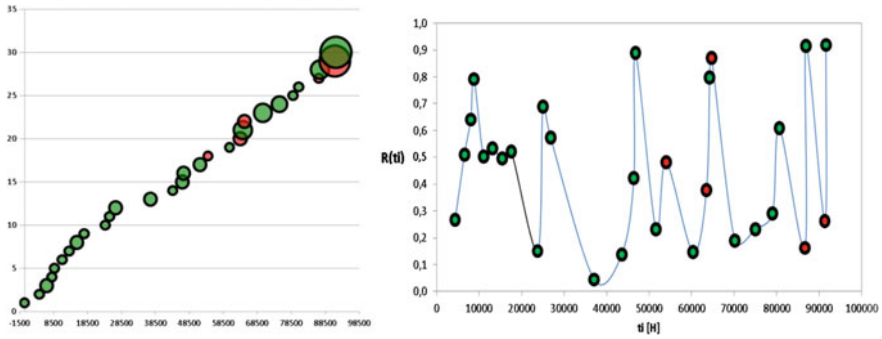


Fig. 6 GAMM graphics using data set 3

For each data set we could appreciate GAMM rules and HIMOS recommendations that were not contradictory but complementary. It was also interesting to see how GAMM rules replace the expertise usually provided by a knowledge-based system (KBS), or the 'intelligent' DSS (IDSS) in the knowledge subsystem.

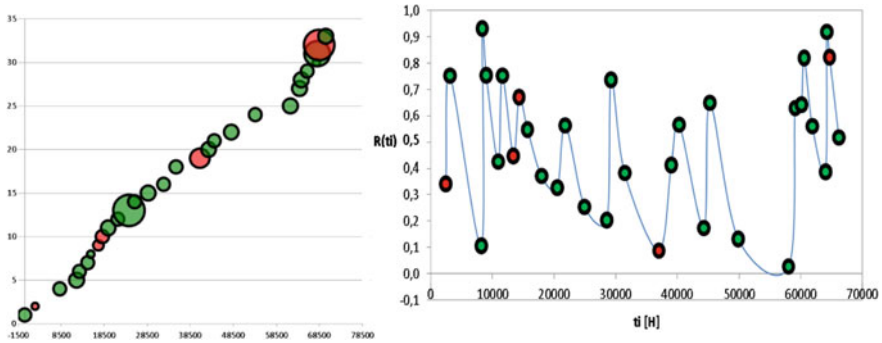


Fig. 7 GAMM graphics using data set 4

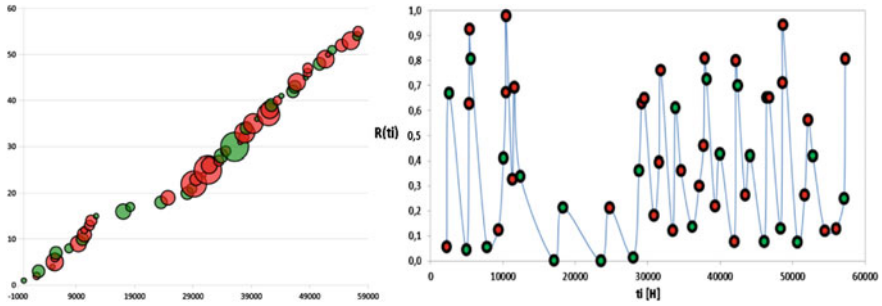


Fig. 8 GAMM graphics using data set 5

5 Conclusions

The aim of this contribution is to present and validate a proposed rule-based reasoning to support maintenance management decision making when using GAMM, the Graphical Analysis for Maintenance Management method. The development of the proposed GAMM rules have been presented and discussed. For the validation of the proposed interpretation rules, GAMM method was applied to analyze a total of five sets of maintenance historical data extracted from the petrochemical industry, and previously analyzed by HIMOS. The application of GAMM (together with methods like HIMOS) could be a support tool for operational maintenance management of short and medium term, providing very useful information, clear and easy to interpret which can be checked with one another in order to determine the optimal maintenance policy (frequency of preventive maintenance). This broadens the spectrum display for maintenance management, acquiring new evaluation parameters and determining possible areas of improvement.

References

1. Kobbacy KAH, Jeon J (2001) The development of a hybrid intelligent maintenance optimisation system (HIMOS). J Oper Res Soc 52:762–778
2. Barberá L, Crespo A, Viveros P, Stegmaier R (2012) Advanced model for maintenance management in a continuous improvement cycle: integration into the business strategy. Int J Syst Assur Eng Manag 3(1):47–63 (Jan–Mar 2012)
3. Barberá L, Crespo A, Viveros P, Arata A (2013) The graphical analysis for maintenance management method: A quantitative graphical analysis to support maintenance management decision making. J Qual Reliab Eng Int 29(77–87):5
4. Barberá L, Crespo A, Viveros P, Stegmaier R (2013) A case study of GAMM (Graphical analysis for maintenance management) in the mining industry. Reliab Eng Syst Saf (RESS) 121:113–120
5. Barberá L, Crespo A, Viveros P, Stegmaier R (2013) A case study of GAMM (Graphical analysis for maintenance management) applied to water pumps in a sewage treatment plant, chile. J Qual Reliab Eng Int

6. Surucu B, Sazak HS (2013) Graphical methods for reliability data. In Wiley encyclopedia of operations research and management science, pp 1–11. https://doi.org/10.1002/9780470400531.eorms1062. Copyright © 2010 John Wiley & Sons, Inc
7. Kobbacy K, Proudlove AH, Harper MA (1995) Towards an intelligent maintenance optimisation system. J Oper Res Soc 46:831–853
8. O'Keefe RM, Balci O, Smith EP (1987) Validating expert system performance. IEEE Expert Winter:81–87

Condition Monitoring of Rotating Machinery with Acoustic Emission: A British–Australian Collaboration

Davide Crivelli, Simon Hutt, Alastair Clarke, Pietro Borghesani, Zhongxiao Peng and Robert Randall

Abstract Industries such as transport and energy generation are aiming to create cleaner, lighter, more reliable and safer technology. Condition monitoring of rotating machinery is an established way of reducing maintenance costs and associated downtime. Whereas vibration based condition monitoring has been validated in literature and in industrial applications, acoustic emission (AE) technologies are a relatively new and unexplored solution for machine diagnostics. Being based on the passive recording of ultrasonic stress waves, their frequency range gives direct access to phenomena such as friction in gear teeth sliding, bearing rolling contacts and crack formation and propagation. However, the complexity of AE signals generated in multiple machine components requires a better understanding of their link with tribological phenomena. To further knowledge in this area, a team of researchers from Cardiff University, Queensland University of Technology and University of New South Wales are conducting a joint research activity which includes: (i) the use of a twin-disk test-rig to reproduce controlled rolling-sliding contact conditions typical of gear contacts, (ii) the analysis of AE data using advanced cyclostationary signal processing and (iii) the establishment

D. Crivelli (✉) · S. Hutt · A. Clarke
School of Engineering, Cardiff University, The Parade, Cardiff CF24 3AA, UK
e-mail: crivellid@cardiff.ac.uk

S. Hutt
e-mail: hutts@cardiff.ac.uk

A. Clarke
e-mail: clarkea7@cardiff.ac.uk

P. Borghesani
Queensland University of Technology, Brisbane, QLD, Australia
e-mail: p.borghesani@unsw.edu.au

Z. Peng · R. Randall
University of New South Wales, Sydney, NSW 2052, Australia
e-mail: z.peng@unsw.edu.au

R. Randall
e-mail: b.randall@unsw.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_12

of a relationship between tribological conditions and AE signal characteristics. This paper outlines this project, discusses its preliminary results and introduces future extensions of this research to key industrial applications.

1 Introduction

The development of detection and diagnostic methods for gear or bearing faults or surface distress in power transmission systems is an active research area [8, 9, 20], with wide applicability ranging from wind turbine gearboxes to aerospace power transmissions. The potential results of gear or bearing failure are serious, from increased asset downtime and maintenance expenditure to catastrophic failure with life-threatening consequences. For example, there are currently widespread problems with gearbox failures in wind turbines, where a typical gearbox life is around 5 years with replacement costing hundreds of thousands of pounds [15, 17]. Thus, new monitoring techniques which can identify incipient failures are in great demand. AE monitoring is such a technique which is used widely in other applications and which offers significant advantages in terms of early fault detection and diagnosis when compared to other techniques [11]. It is arguably the most sensitive NDT technique available, and relies upon the detection of stress waves which propagate through the solid material as it undergoes strain. When compared with the reasonably mature application of AE to structural monitoring, the use of AE to monitor rotating machinery in general, and gears and bearings in particular, is still at the developmental stage, particularly when applications such as high speed, heavily loaded aerospace or wind turbine transmissions are considered. Researchers have investigated the AE from spur gears [12, 16, 18] and from rolling element bearings [2, 6, 19] but there remains much confusion surrounding the sources of acoustic emission in concentrated elastohydrodynamic (EHL) contacts within gears and bearings. Further work is required not only to understand the source mechanisms and propagation but also to provide an automated technique for identifying damage in these types of applications. It is well known that surface failure (micropitting) usually originates at prominent asperity features, and investigations have shown that it may be possible for AE to provide a measure of the level of asperity contact for meshing gears under a range of operating conditions [16] and for rolling element bearings [6]. However, these AE investigations often operate under conditions which are not wholly representative of typical aerospace or wind turbine design practice or do not link closely with corresponding tribological research into contact conditions (particularly under mixed lubrication conditions where asperity contact is likely).

This paper presents two case studies detailing initial work as part of a British-Australian collaboration aimed at improving understanding of the fundamental link between tribological conditions and the generation of AE for lubricated contacts, and also to apply novel analysis techniques to improve the ability of AE-based technologies to detect faults within mechanical systems.

2 Analysis Techniques

The AE wavestream data recorded during the tests outlined in this paper have been analysed using a number of techniques, outlined below:

2.1 Frequency Content/Binned Analysis

The wavestreams were analysed in terms of their frequency content using amplitude FFT algorithms within the Matlab environment. The analysis of successive wavestreams allowed the evolution of frequency content with time to be examined. Furthermore, to allow for slight fluctuations in operating speed, the frequency content was analysed in frequency "bins" or ranges, summing content within each range, to evaluate the particular frequency ranges of interest. A full description of these techniques may be found in Cockerill et al. [5].

2.2 Chebyshev Moments

A reconstruction of the wavelet decomposition of a signal, in particular, can be used as a form of time-frequency transform, where each wavelet level is more sensitive to certain frequencies within a signal.

The Chebyshev moments calculation procedure from Crivelli et al. [7] is as follows:

1. sample a discrete waveform with N points;
2. compute a discrete wavelet transform using M detail levels (Daubechies 10);
3. reconstruct and rectify the wavelet details into a N $\times$ M matrix;
4. compute the Chebyshev moments of the N $\times$ M matrix up to the desired degree.

Steps 2–3 produce a virtual image (matrix) of the time domain signal, where each row represents a wavelet detail level (approximately a frequency band). As the degree of the Chebyshev moment increases, more detail about the representation of the signal will be carried in the representation. These moments can be used as signal descriptors and further used for comparing datasets, as shown by Sebastian et al. [13]. From previous tests, a polynomial degree of 5 (25 moments) is sufficient to carry enough information while maintaining the computational effort low enough to be usable in real time applications.

3 Case Studies

3.1 QUT Bearing Rig

The data analysed in this section was collected on a bearing prognostic test-rig located in the laboratories of the Queensland University of Technology (QUT), Brisbane (Australia). The test rig comprises a variable-speed electric motor, driving the main shaft where a test bearing (type 6806) is installed and loaded radially by a screw mechanism and a spring. The bearing is installed in healthy conditions and loaded at approximately 1.5 times its dynamic load rating in order to accelerate the degradation process. Degradation is obtained by repeating a duty cycle which includes a series of constant speed intervals with different rotating speeds (in the range 210–1300 rpm), with one full duty cycle repeating every 960 s. The run-to-failure test is considered complete when RMS vibration levels reach 24.5 g.

Chebyshev descriptors were extracted from the AE data for each revolution of the bearing test rig and different speeds. Figure 1 shows the individual moment values at 500 rpm on an exponentially weighed moving average (EWMA) control chart. The EWMA control chart is used for monitoring when the process goes "out of bounds". It assigns an exponentially weighted average to all prior data so that it becomes less sensitive to small drifts in the process (running in, periodic variations …). This results in violations when the drift is quick, i.e. when damage is present. As such, when analysed, the values indicate that signals are consistent in the first part of the test but as expected tend to deviate towards the end.

Figure 2 shows the number of EWMA control chart violations (number of parameters out of control) at different speeds, with control boundaries set at 2σ. The variations become apparent with Chebyshev moments much earlier in the test and at slower speeds when compared with traditional statistical moments (mean, standard deviation, skewness and kurtosis). It is also clear that, especially at runs closer to the failure of the bearing, the Chebyshev moments-based analysis can pick up variations in the signals at speeds where statistical moments are not showing any violation. It is also apparent that some Chebyshev moments appear to be highly correlated. This is likely due to the fact that two nearby moments can be sensitive to similar effects in relatively long signals, due to the polynomial order they represent. However, further investigation is required to confirm this assumption.

3.2 Cardiff Variable Lambda Rig

In order to investigate in a fundamental way the relationship between the generation of AE and conditions within elastohydrodynamic (EHL) rolling/sliding contacts, such as those between the teeth of heavily loaded gears or in the roller/raceway contact in a rolling element bearing, a power-recirculating twin disc test rig was developed. This rig can be seen in Fig. 3 and is described more fully by Clarke

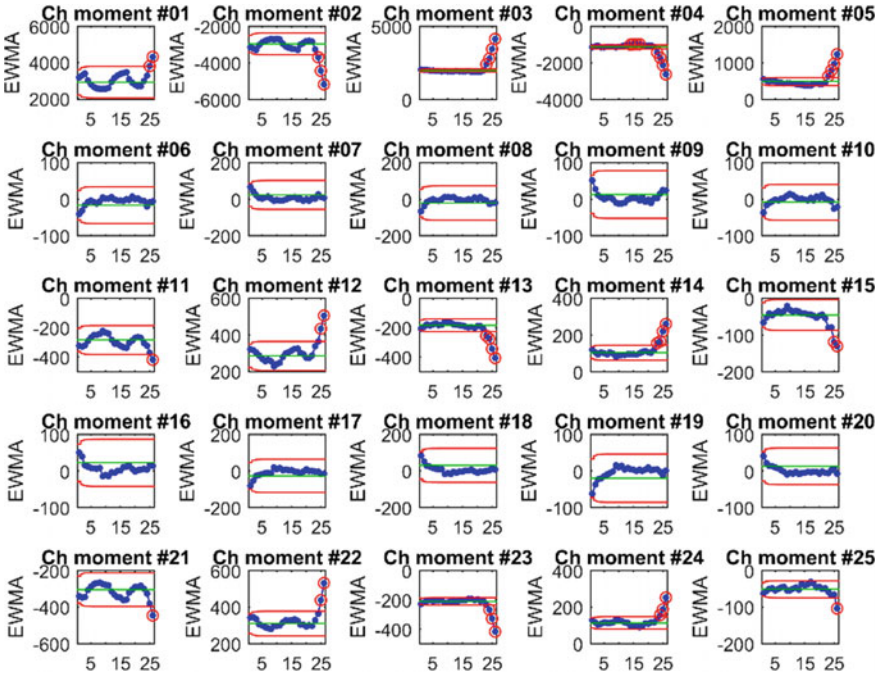


Fig. 1 Individual Chebyshev moment values @500 rpm plotted against test run: EWMA control charts. Violations (red circles) when values fall out of bounds (red lines). Some moments are more sensitive than others to changes in signal time/frequency content

et al. [3, 4]. It is capable of operating well into the mixed lubrication regime where load is carried by both direct asperity contact and by a partial hydrodynamic film.

In this rig, power is recirculated between the EHL contact between the test disks, and a gear pair, meaning that the drive motor only has to overcome frictional and other losses in the system. The test disks (made to typical gear material and hardness specifications) are 76.2 mm in diameter and have a 304.8 mm crown radius, which results in an elliptical EHL point contact with a nominal 4:1 aspect ratio, with the major axis parallel to the shaft. They are finished using an axial grinding technique which generates the crown, and gives a surface with similar roughness directionality to that of ground gear teeth in relation to the motion of the surfaces.

The shafts on which the disks are mounted are gear connected, giving a slide/roll ratio in this work of 0.5. The fast shaft speed is adjustable between 200 and 3000 rpm, and the hydraulically applied load can be varied to generate Hertzian contact pressures of up to 2.1 GPa. The contact is lubricated using a naval gear oil, supplied at a controlled temperature. The rig is fitted with an acoustic emission sensor mounted on the slower disc, connected to the AE system via slip rings. Raw AE wavestreams were recorded at 5 MHz sampling frequency using a Mistras AE system with a configuration similar to that of Cockerill et al. [5].

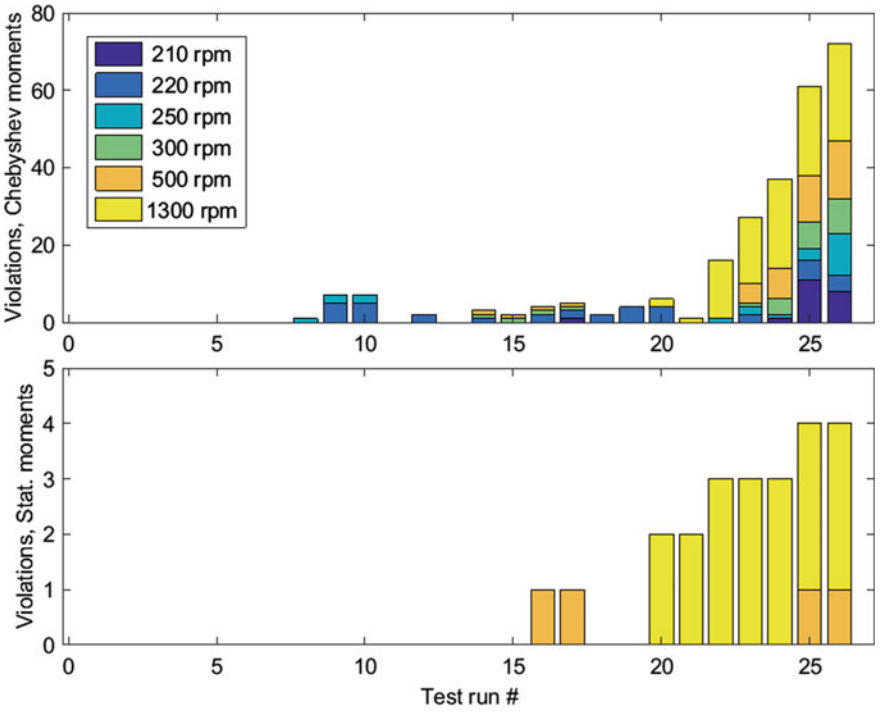


Fig. 2 Number of violations at different RPM during the bearing run to failure: Chebyshev moment violations (above) are more sensitive at low speed and provide an earlier and more gradual warning when compared to statistical moment violations (below)

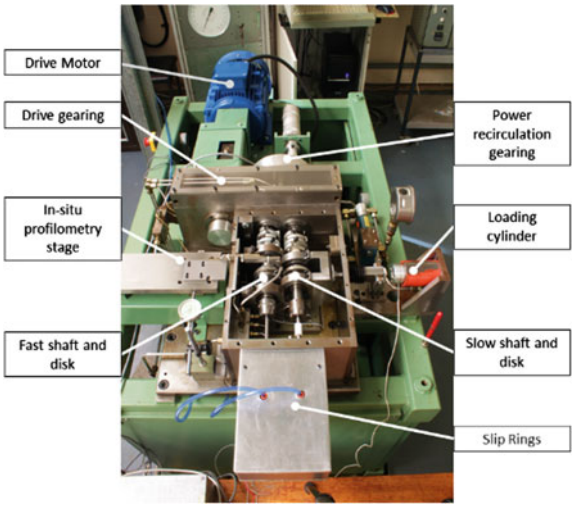


Fig. 3 Twin disc test rig for investigation of mixed lubrication conditions

In order to quantify the contact conditions, the concept of the lambda ratio (Λ) is used. This quantifies the combined effects of load, speed and temperature on mean metallic contact levels, and is defined as:

$$\Lambda = \frac{h}{\sqrt{R_{q1}^2 + R_{q2}^2}} \tag{1}$$

where h is the calculated smooth-surface film thickness, and R_{q1} and R_{q2} are the root mean square surface roughnesses for the two disk surfaces. The smooth-surface film thickness is calculated using the well-known formula given by Chittenden et al. [1] and the effects of temperature on lubricant viscosity (and therefore on film thickness) are taken into account by using the disk temperatures (measured using embedded thermocouples) to calculate lubricant viscosity. It is widely accepted [3, 4, 10, 14] that mixed lubrication occurs over a range of Λ between 0.1 and 2, with direct asperity contact carrying an increasing proportion of the total contact load as Λ decreases.

Initially, the test disks were run-in using the procedure described in Clarke et al. [3, 4], and for the tests reported here the surface roughness of the disks was stable. In order to measure the AE generated by the contact at a range of Λ ratios, the disks were initially run at a fast disk speed of 300 rpm, with load applied to generate a maximum contact pressure of 1.2 MPa. Shear heating of the lubricant within the contact caused the temperature of the test disks to rise (leading to a reduction in Λ), with AE wavestreams being frequently recorded during this temperature rise. Subsequently, the oil supply temperature was increased in order to extend the range of Λ ratios experienced further into the mixed lubrication regime, and the process repeated at fast disk speeds of 500, 1000, 1500 and 2000 rpm. Following the tests, wavestreams were examined for their frequency content by using a binned FFT. Frequencies between 150 and 300 kHz were found to be particularly sensitive to contact conditions. Figure 4 clearly demonstrates that there is a link between AE and the level of mixed lubrication. For any of the five speeds, as Λ falls, the level of AE increases, clearly demonstrating the interaction of asperities on the contacting surfaces to be a source of AE.

As these interactions occur more frequently (and are more heavily loaded) as Λ falls, so the AE activity increases. The effect is more pronounced at the higher speeds, which may be explained when one considers that at higher speeds, the overall levels of energy within the system increase, so it is reasonable to expect that any AE signal will be higher as the test speed and overall energy levels are increased. The results of these tests clearly demonstrate the potential of AE to be used as a tool to monitor contact conditions within realistic, heavily loaded elastohydrodynamic contacts operating within the mixed lubrication regime.

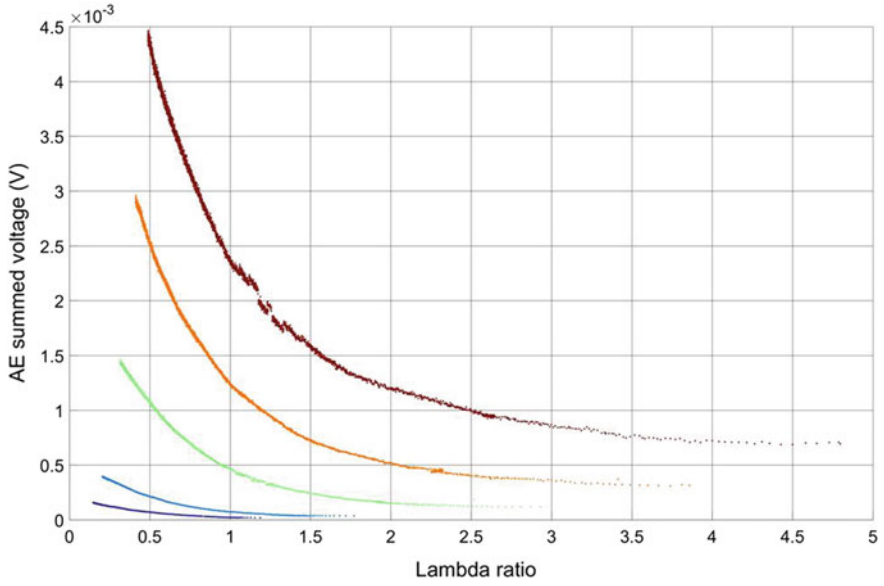


Fig. 4 Total AE voltage between 150 and 300 kHz versus Lambda Ratio (Λ). Each line represents data collected at a particular fast disk rotational speed—from bottom to top these are 300, 500, 1000, 1500 and 2000 rpm

4 Discussion and Planned Work

The results clearly demonstrate the sensitivity of AE to tribological conditions both at a component level (in the bearing tests) and at a contact level (in the disk machine mixed lubrication tests). There is clearly scope for using AE to develop sensitive monitoring techniques which can not only detect component-level failure but can provide insight into the status or health of the lubrication mechanism between surfaces. Many failure mechanisms (such as micropitting) which are prevalent in highly loaded power transmission systems are initiated at the surface roughness level, and the results shown here give confidence that AE may be able to provide early insight into incipient failures of this nature. In order to develop this research programme, further tests are planned at the bearing level (on two rigs at different universities in order to provide some "blind" test data for analysis techniques), at the surface level (tribometer tests with dry and lubricated surfaces of various roughness), on simulated gear tooth contacts (disc machine tests to investigate further the AE response under mixed lubrication conditions and during micropitting initiation) and on gear/bearing systems under various operating conditions. These activities should provide useful data and insight and allow the evaluation of the effectiveness and viability of AE as a sensitive condition monitoring technique for rotating machinery. Further research will involve developing models for the mechanism of AE generation within these tests.

5 Conclusions

At this initial stage of the collaborative project, it can be concluded that:

- AE is a sensitive tool capable of detecting the failure of components such as rolling element bearings.
- Advanced techniques such as Chebyshev moments may be used to enhance the sensitivity of AE monitoring.
- AE is sensitive to conditions within lubricated rolling/sliding contacts operating under mixed lubrication conditions.
- In particular, AE between 150 and 300 kHz rises as the levels of asperity interaction/direct metallic contact rise as indicated by a reduction in lambda ratio.
- Overall levels of AE within the contact rise as the speeds of the surfaces are increased, due to an overall increase in energy within the system.
- Further work is planned to investigate these phenomena in more detail.

References

1. Chittenden R, Dowson D, Dunn J, Taylor C (1985) A theoretical analysis of the isothermal elastohydrodynamic lubrication of concentrated contacts I. Proc R Soc Lond 397(A):245–269
2. Choudhury A, Tandon N (2000) Application of acoustic emission technique for the detection of defects in rolling element bearings. Tribol Int 33:39–45
3. Clarke A, Weeks IJJ, Evans HP, Snidle RW (2016) An investigation into mixed lubrication conditions using electrical contact resistance techniques. Tribol Int 93:709–716. https://doi.org/10.1016/j.triboint.2014.10.010
4. Clarke A, Weeks IJJ, Snidle RW, Evans HP (2016) Running-in and micropitting behaviour of steel surfaces under mixed lubrication conditions. Tribol Int 101:59–68. https://doi.org/10.1016/j.triboint.2016.03.007
5. Cockerill A, Clarke A, Pullin R, Bradshaw T, Cole P, Holford K (2016) Determination of rolling element bearing condition via acoustic emission. Proc Inst Mech Eng Part J: J Eng Tribol 230(11):1377–1388. https://doi.org/10.1177/1350650116638612
6. Couturier J, Mba D (2008) Operational bearing parameters and acoustic emission generation. J Vibr Acoust 130:24502. https://doi.org/10.1115/1.2776339
7. Crivelli D, McCrory J, Miccoli S, Pullin R, Clarke A (2017) Gear tooth root fatigue test monitoring with continuous acoustic emission: advanced signal processing techniques for detection of incipient failure. Structural Health Monit. https://doi.org/10.1177/1475921717700567
8. Decker HJ (2002) NASA-TM2002-211491: Gear crack detection using tooth analysis
9. Decker HJ (2002) NASA-TM2002-211492: Crack detection of aerospace quality gears
10. Guangteng G, Spikes HA (1996) An experimental study of film thickness in the mixed lubrication regime. In: Dowson D (ed) Elastohydrodynamics 96: fundamentals and applications in lubrication and traction. Elsevier, Amsterdam, pp 159–166. https://doi.org/10.1016/s0167-8922(08)70445-0
11. Loutas TH, Sotiriades G, Kalaitzoglou I, Kostopoulos V (2009) Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements. Appl Acoust 70(9):1148–1159. https://doi.org/10.1016/j.apacoust.2009.04.007

12. Raja Hamzah RI, Mba D (2007) Acoustic emission and specific film thickness for operating spur gears. J Tribol 129(4):860–867. https://doi.org/10.1115/1.2769732
13. Sebastian C, Patterson E, Ostberg D (2011) Comparison of numerical and experimental strain measurements of a composite panel using image decomposition. Appl Mech Mater 70:63–68. https://doi.org/10.4028/www.scientific.net/AMM.70.63
14. Sharif K, Evans H, Snidle R (2012) Modelling of elastohydrodynamic lubrication and fatigue of rough surfaces: the effect of lambda ratio. Proc Inst Mech Engineers Part J: J Eng Tribol 226(12):1039–1050. https://doi.org/10.1177/1350650112458220
15. Spinato F, Tavner PJ, van Bussel GJW, Koutoulakos E (2009) Reliability of wind turbine subassemblies. IET Renew Power Gener 3(4):387–401. https://doi.org/10.1049/iet-rpg:20080060
16. Tan CK, Mba D (2005) Identification of the acoustic emission source during a comparative study on diagnosis of a spur gearbox. Tribol Int 38(5):469–480. https://doi.org/10.1016/j.triboint.2004.10.007
17. Tavner PJ, Xiang J, Spinato F (2007) Reliability analysis for wind turbines. Wind Energy 10(1):1–18
18. Vicuña CM (2014) Effects of operating conditions on the Acoustic Emissions (AE) from planetary gearboxes. Appl Acoust 77:150–158. https://doi.org/10.1016/j.apacoust.2013.04.017
19. Williams T, Ribadeneira X, Billington S, Kurfess T (2001) Rolling element bearing diagnostics in run-to-failure lifetime testing. Mech Syst Signal Process 15(5):979–993. https://doi.org/10.1006/mssp.2001.1418
20. Wong AK (2001) Vibration-based helicopter gearbox health monitoring—an overview of the research program in DSTO. In: Proceedings of defence science and technology organisation international conference on health and usage monitoring, pp 1–12

Semi-analytical Approach to Vibrations Induced by Oscillator Moving on a Beam Supported by a Finite Depth Foundation

Zuzana Dimitrovová

Abstract Structures subject to moving loads have several in rail, road and bridge engineering. When the velocity of the moving system approaches the critical velocity, then the induced vibrations are significantly augmented and safety and stability of the structure as well as of the moving system are compromised. The classical models predict the critical velocity much higher than the one observed in reality, because the wave propagation is restricted to the beam structure. But if the beam is supported by elastic continuum, then the waves can be dominant in the foundation and the interaction with the beam cannot be overlooked. This contribution analyses the critical velocity of an oscillator moving on a beam supported by a foundation of finite depth by semi-analytical methods.

1 Introduction

The objective of this contribution is to fill the gap in semi-analytical solutions related to wave propagation induced by moving loads. An analysis of the critical velocity of an oscillator moving on a beam supported by a foundation of finite depth is presented. Such analysis is important for transport engineering, as structures subjected to moving loads have several applications in rail, road and bridge engineering. The classical model where the beam structure is supported by massless linear springs is still used by several companies, due to its simplicity. Nevertheless, in this model the wave propagation is restricted to the beam structure and no dynamic interaction with the foundation can be accessed. Models considering beams placed on elastic continua of finite depth provide better access to the interaction mechanism. Nevertheless, as the inertial effects are important in the

Z. Dimitrovová (✉)
Departamento de Engenharia Civil, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisbon, Portugal
e-mail: zdim@fct.unl.pt

Z. Dimitrovová
IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_13

supporting structure, they are also important is the moving system. This means that the moving system should not be approximated by moving forces, as it is commonly done, but at least by a moving oscillator.

There are already several works dedicated to analytical and semi-analytical solutions, mainly concerned with stability issues of the moving oscillator, [1, 2]. In this paper different approach is presented directly linked to the issue of the critical velocity. At first, the problem of a moving force travelling over a beam supported by a finite depth elastic continuum in two-dimensions will be reviewed. Then, the semi-analytical solution of the mass moving on a beam supported by elastic springs will be extended to moving oscillator and validated by software LS-DYNA. Finally, the foundation will be substituted by frequency dependent springs, which is an acceptable approximation of the elastic finite depth continuum, and conclusions on the critical velocity will be drawn.

2 Finite Depth Elastic Continuum

It is assumed that the load P is traversing an infinite beam supported by a foundation of finite depth H, by a constant velocity v, as depicted on Fig. 1. It is further assumed that: (i) the beam obeys linear elastic Euler–Bernoulli theory; (ii) the foundation is represented by a linear elastic homogeneous continuum of finite width b under plane strain condition; (iii) the beam may be subjected to a normal force N acting on its axis (considered positive when inducing compression); (iv) gravitational effects on the beam and on the foundation are neglected.

It was derived in [3] that if horizontal displacements are omitted, but shear deformations are accounted for in the vertical dynamic equilibrium of the foundation, then the critical velocity ratio α_{cr} defined as $\alpha_{cr} = V_{cr}/v_{cr}$, where V_{cr} is the new value of the critical velocity and v_{cr} is the classical value of the critical velocity of the load passing a beam on Winkler's foundation, is governed by an approximate formula

$$\alpha_{cr} = \left(\sqrt{1-\eta_N} - \vartheta_s\right)\sqrt{\frac{2}{2+M^{2+\sqrt{\vartheta_s}}}} + \vartheta_s \tag{1}$$

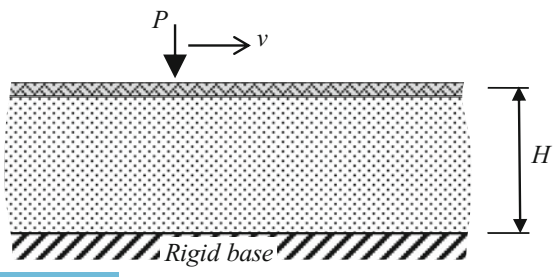


Fig. 1 Infinite beam on an elastic foundation of finite depth subjected to a moving load

where, ϑ_s is the shear ratio defined as $\vartheta_s = v_s/v_{cr}$ with v_s being the shear-wave velocity, and M is the mass ratio defined as the square root of the foundation mass to the beam mass $M = \sqrt{\rho bH/m}$ with ρ being the soil density and m the mass per length of the beam. Thus, for a lower mass ratio, the critical velocity approaches the classical value $v_{cr,N}$ with the effect of the normal force specified by the ratio η_N; and for a higher mass ratio, it approaches the velocity of propagation of shear waves in the foundation, which is the lowest wave-velocity of propagation related to the model adopted, because the Rayleigh waves cannot be developed if horizontal displacements are not properly introduced.

In the full two-dimensional model, both dynamic equilibrium equations are considered and thus horizontal displacements are not omitted. The system of governing equations read

$$\mu\Delta\mathbf{u} + (\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}) = \rho\mathbf{u}_{,tt} \tag{2}$$

$$EIw_{,xxxx} + Nw_{,xx} + c_b w_{,t} + mw_{,tt} + p_f = P\delta(x-vt) \tag{3}$$

where $\mathbf{u} = (u_x, u_z)$ is the vector of the displacement field in the foundation, λ and μ are Lame's constants of the elastic continuum, EI and c_b stand for the bending stiffness and the coefficient of viscous damping of the beam, p_f is the foundation pressure and t is the time. In this paper, derivatives will be designated by the corresponding variable symbol in the subscript position, preceded by a comma. Moreover, ∇ is the gradient and Δ is the Laplace operators applied on spatial variables x, z. The unknown beam deflection $w(x,t)$, displacement u_z, spatial coordinate z and force P are assumed positive when acting downward. Spatial coordinate x is positive to the right, the load travels from the left to the right and finally, δ is the Dirac delta function.

The solution method exploits displacement potentials and moving coordinate. Then, for the steady-state solution time dependent terms can be neglected and after the Fourier transform, analytical solution can be obtained in the Fourier space. The additional interface condition is ambiguous and can be written in form of zero horizontal displacement (ZHD), zero shear stress (ZSS) or some combination of these two in form of horizontal interface spring. The critical velocity can be identified in a semi-analytical way by identification of double poles on the real axis of the Fourier variable of the beam displacement image. Approximate formula for the critical velocity ratio is given as a sum of two parts, one is an addition to the previous Eq. 1 and the other one is an adaption Eq. 1

$$\alpha_{cr,add} = a_1\vartheta_j\left(\sqrt{\frac{2}{2+M^{2+\sqrt{\vartheta_j}}}} - \sqrt{\frac{2}{2+a_2\left(\vartheta_j+a_N\right)M^{2+\sqrt{\vartheta_j}-a_N}}}\right) \tag{4}$$

$$\alpha_{cr,prev} = \left(\sqrt{1-\eta_N} - a_3\vartheta_j\right)\sqrt{\frac{2}{2+M^{2+\sqrt{\vartheta_j}}}} + a_3\vartheta_j \tag{5}$$

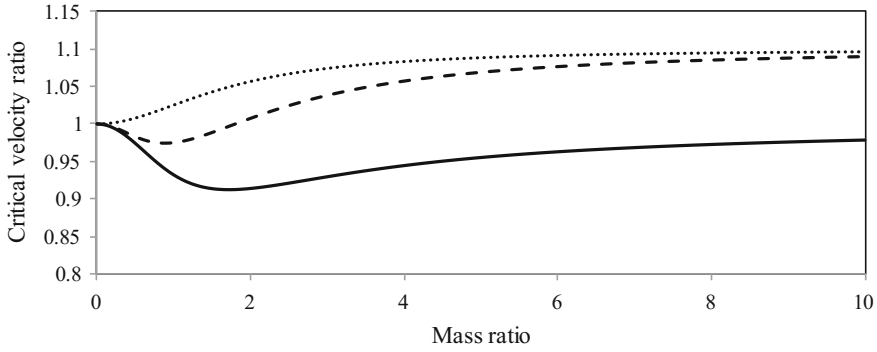


Fig. 2 Critical velocity ratio: simplified model (dotted), ZHD condition (dashed), ZSS condition (full), $\vartheta_s = 1.1$

where $a_N = 2a\eta_N\left(\vartheta_j - \eta_N\right)^2$ with $a = 1$ or 10 when $\vartheta_j < \eta_N$. In addition the subscript j is related to the velocity of propagation of shear waves, thus $j = s$, $a_1 = 0.3$, $a_2 = 0.4$, $a_3 = 0.99$ for ZHD; and for ZSS it is related to the velocity of propagation of Rayleigh waves, thus $j = R$, $a_1 = 0.5$, $a_2 = 0.4$ and $a_3 = 0.98$, so $\vartheta_R = v_R/v_{cr}$, where v_R is the velocity of propagation of Rayleigh waves. Results of the critical velocity are slightly affected by this additional interface condition, as can be seen in Fig. 2.

In Fig. 3 the influence of the normal force is shown for $\eta_N = 0.5$, simplified approach and ZSS condition are placed together in one graph for $\vartheta_s \in \langle 0.5; 1.5\rangle$.

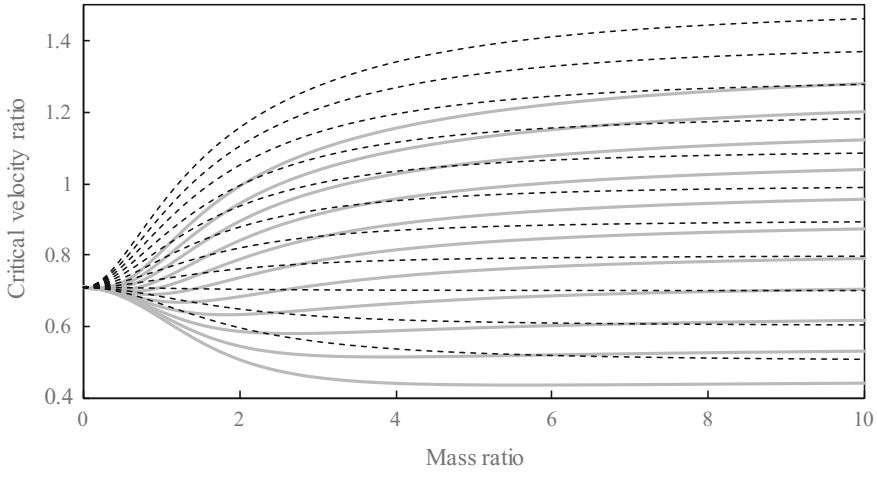


Fig. 3 Critical velocity ratio: simplified model (dashed), ZSS condition (full grey), $\vartheta_s \in \langle 0.5; 1.5\rangle$, for each case the curves are starting from the bottom

It is seen, that there are some differences between the results of the critical velocity ratio related to ZHD and ZSS reflected especially in the asymptotic value of the critical velocity, which is v_s and v_R for the two conditions, respectively. On the other hand the asymptotic value of the simplified model and ZHD is approximately the same.

3 Massless Springs

In order to extend the previous analysis to the moving oscillator, finite beams on elastic springs foundation will be considered first. The reason is that, as shown in [4], the results on long finite beams provide very good approximation to the results related to infinite ones, so long as the mass is added to the moving force. Then the semi-analytical technique for solution of the moving mass problem by eigenvalues expansion can be extended to the moving oscillator and finally to the frequency dependent foundation. Last step would be to consider the complete two-dimensional continuum.

Thus, let a uniform motion of a constant mass and a vertical force with harmonic component along a horizontal infinite beam posted on a two-parameter visco-elastic foundation be assumed. Besides the previous simplifying assumptions, it is assumed that the mass is always in contact with the beam and its horizontal position is determined by its velocity. The problem at hand is depicted in Fig. 4.

The equation of motion for the unknown vertical displacement reads

$$EIw_{,xxxx} + \left(N - k_p\right)w_{,xx} + mw_{,tt} + c_b w_{,t} + kw = p(x,t) \tag{6}$$

with the loading term being

$$p(x,t) = \left(P + P_0 \sin\left(\omega_f t + \varphi_f\right) - Mw_{0,tt}(t)\right)\delta(x - vt) \tag{7}$$

Because this analysis will be used only for representation of the infinite situation, boundary conditions can be selected in the most convenient way. Assuming that

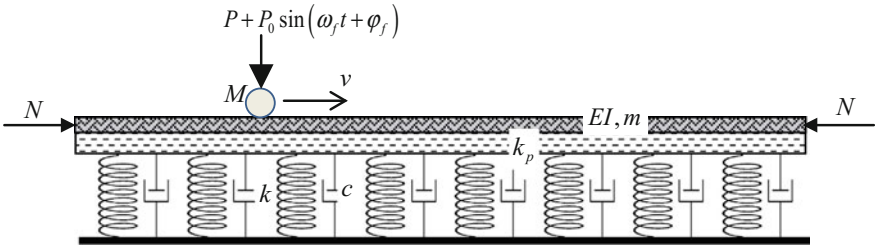


Fig. 4 Finite beam on a visco-elastic two-parameter foundation subjected to a moving load and a normal force

only n modes will be used, a compact matrix form for the generalized coordinate calculation can be presented as

$$\mathbf{M}(t)\cdot\mathbf{q}_{,tt}(t)+\mathbf{C}(t)\cdot\mathbf{q}_{,t}(t)+\mathbf{K}(t)\cdot\mathbf{q}(t)=\mathbf{p}(t) \tag{8}$$

where square $n\times n$ matrices $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are not approximations resulting from some discretization of the problem, but are defined by vibration modes in their exact analytical form

$$M_{jk}=\delta_{jk}+Mw_j(vt)w_k(vt) \tag{9}$$

$$C_{jk}=\delta_{jk}\frac{c}{m}+2Mvw_j(vt)w_{k,x}(vt) \tag{10}$$

$$K_{jk}=\delta_{jk}\omega_j^2+Mv^2w_j(vt)w_{k,xx}(vt) \tag{11}$$

$$p_j=\left(P+P_0\sin\left(\omega_f t+\varphi_f\right)\right)w_j(vt) \tag{12}$$

These terms can be reordered, keeping the diagonal part of the matrices and introducing an additional variable λ that will join the necessary terms

$$\boldsymbol{M}(t)=\begin{bmatrix}M_{jk}^D & -w_j(vt)\\ Mw_k(vt) & 1\end{bmatrix},\boldsymbol{C}(t)=\begin{bmatrix}C_{jk}^D & 0\\ 2Mvw_{k,x}(vt) & 0\end{bmatrix} \tag{13}$$

$$\boldsymbol{K}(t)=\begin{bmatrix}K_{jk}^D & 0\\ Mv^2w_{k,xx}(vt) & 0\end{bmatrix} \tag{14}$$

Thus there is only one coupled equation and the second time derivative of the additional variable equals the contact force. If the oscillator is added, then the loading term from Eq. 7 must be completed by

$$(-k_{os}w_0(t)+k_{os}w_{os}(t))\delta(x-vt) \tag{15}$$

and additional equation read

$$M_{os}w_{os,tt}(t)+k_{os}w_0(t)-k_{os}w_{os}(t)=P_{os} \tag{16}$$

where the subscript "os" designates the quantities related to the oscillator. For typical values related to railways applications and very soft foundation, it is seen that mass induced frequency is superposed with the oscillator frequency (Fig. 5).

Therefore, the method for the critical velocity determination follows the identification of the resonant case. The mass induced vibration can be determined by semi-analytical methods presented in [4]. The typical graph of the real part of the induced frequencies is shown in Fig. 6.

Then according to the natural frequency of the oscillator, the velocity of the moving system is determined.

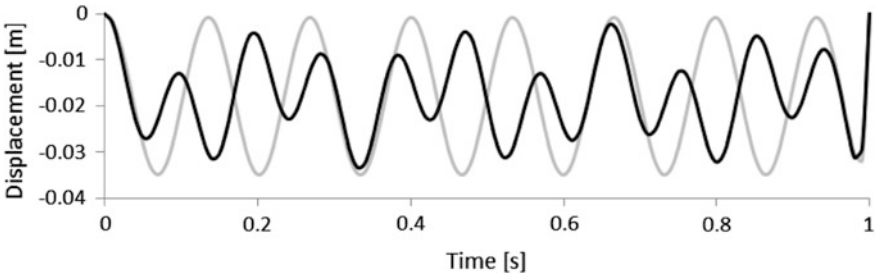


Fig. 5 Beam deflection under the load: moving oscillator (black) moving mass (grey)

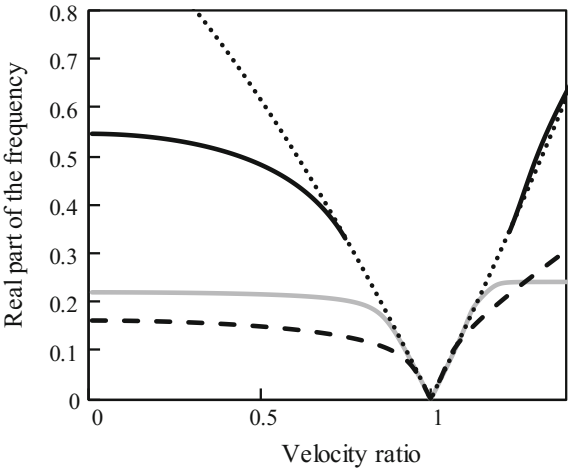


Fig. 6 Real part of the dimensionless induced frequencies: two-mass oscillator (black and grey), moving mass (dashed) and cutting frequency (dotted)

4 Conclusions

In this paper, a technique for the semi-analytical determination of the critical velocity of the moving oscillator is proposed. Based on the characteristics of the oscillator, such a critical velocity can be lower that the critical velocity of the moving force and also lower than the velocity inducing instability of the moving mass.

Acknowledgements This work was supported by FCT, through IDMEC, under LAETA, project UID/EMS/50022/2013.

References

1. Metrikine AV, Popp K (1999) Instability of vibrations of an oscillator moving along a beam on an elastic half-space. Eur J Mech A-Solid 18:331–349
2. Metrikine AV, Verichev SN, Blaauwendraad J (2005) Stability of a two-mass oscillator moving on a beam supported by a visco-elastic half-space. Int J Solids Struct 42:1187–1207
3. Dimitrovová Z (2016) Critical velocity of a uniformly moving load on a beam supported by a finite depth foundation. J Sound Vib 366:325–342
4. Dimitrovová Z (2017) New semi-analytical solution for a uniformly moving mass on a beam on a two-parameter visco-elastic foundation. Int J Mech Sci 127:142–162

Bearing Defect Detection Using Envelope Extraction for Dimension Reduction

Fang Duan, Michael Corsar, Linghao Zhou and David Mba

Abstract Fault detection of the rolling element bearing (REB) has been the subject of extensive research because of its detrimental influence on the reliability of machines. Vibration-based condition monitoring is one of the commonly used methods. In most cases, vibration signals are attenuated and contaminated resulting from background noise and complex structure. Independent component analysis (ICA) has been proved to be an effective method to separate bearing defect related feature from background noise. However, it is a prerequisite that the number of observations has to be larger than that of sources. The requirement cannot be satisfied in helicopter main gearbox (MGB) bearing condition monitoring because it is not possible to install more sensors than vibration sources considering the complexity of the MGB. Hence, this paper investigates the feasibility of using envelope extraction to reduce signal dimension. The experiment was conducted on a MGB operating under different load level and input speed. The results show that bearing defect related feature was observed by combing envelope extraction and the ICA method.

F. Duan (✉) · L. Zhou
School of Engineering, London South Bank University, SE1 0AA London, UK
e-mail: duanf@lsbu.ac.uk

L. Zhou
e-mail: zhoul7@lsbu.ac.uk

M. Corsar
School of Engineering, Cranfield University, MK43 0AL Cranfield, UK
e-mail: m.r.corsar@cranfield.ac.uk

D. Mba
Faculty of Technology, De Montfort University, LE1 9BH Leicester, UK

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_14

1 Introduction

Rolling element bearings (REB) are an essential and critical part of rotating machinery. During the service time, faults occurring in bearings may lead to serious damage and fatal breakdown. Hence, REB faults detection and diagnosis in the early stages of damage is necessary to prevent malfunctioning and failure of the whole device. Once bearing defect occurs, a series of impulses are generated every time when a running roller passes over the surface of bearing flaws. As a consequence, the vibration signature of the damaged bearing consists of an exponentially decaying sinusoid having the structure resonance frequency [1]. The amplitude and periodicity of the impulses are related to the structure of bearing, load, flaw location, operating speed and so on. Failures are often preceded by changes in the normal vibration of the system. Therefore, it is feasible to examine the health condition of REBs by monitoring the vibration signals.

Vibration sensors (typically accelerometers) are one of the commonly used sensors in monitoring bearing defects. Since vibrations sensors are normally attached on the case of machine, faults related features are usually immersed in background noise, which makes them difficult to be detected. Over the last few decades, great effects have been made to develop advanced signal processing methods to extract features from signals. Consequently, bearing fault detection methods have been derived from simple calculation of kurtosis and root mean square (RMS) values and Fourier transformation to more sophisticated schemes, such as spectral kurtosis [2], fast kurtogram [3], empirical mode decomposition [4] and independent component analysis (ICA) [5].

Since most gearbox includes multiple vibration sources, the observed vibration signals are the mixtures of different vibration sources. The ICA algorithm has been employed to decompose vibration sources to extract fault related features [5–7]. As one of the general requirements of ICA algorithm, the number of observations should be larger than the number of sources. However, it is not feasible to have more vibration sensors than the vibration sources in the complex MGB scenario. Guo et al. [6] proposed using envelope analysis to reduce dimension prior of using ICA algorithm. The proposed method was validated in a small gearbox with two gears.

So far, most studies were based on simple machinery structures, which have limited number of bearings and gears. In this research, bearing defect detection has been extended from a single bearing to a complex full size helicopter main gearbox (MGB). The MGB consists of several gears and bearings to convert the engine power from high speed and low torque to low speed and high torque to drive the main rotor blades to generate lift. A defect was seeded on one of the planetary gear bearing outer race of the second stage epicyclic reduction gear module. There are multiple vibration sources in the MGB because of the complexity of MGB. Hence, the dimension of recorded vibration signals was firstly reduced by envelope extraction. Then, FastICA algorithm [8] was applied to decompose vibration sources. The experiment was conducted under different rotating speed and load

level. The analysis results proved the efficacy of the combination of the envelope extraction and FastICA in detecting bearing fault in complex environment.

2 FastICA and Envelope Extraction

2.1 FastICA

The FastICA algorithm is a fast and robust fixed-point algorithms for ICA analysis [8]. In our previous study, the FastICA algorithm was employed to separate the multichannel signals into the mutually independent components in the similar scenario [9]. The recorded vibration signals from m accelerometers were regarded as observations $\mathbf{x} = [x_1, \ldots, x_m]$. The FastICA algorithm was utilized to find a mixing matrix A to separate the source signals $\mathbf{s} = [s_1, \ldots, s_i]$. The general ICA model can be expressed as

$$\boldsymbol{x} = A\boldsymbol{s} \tag{1}$$

or equivalent expression of

$$\boldsymbol{s} = A^{-1}\boldsymbol{x} \tag{2}$$

ICA algorithm seeks a matrix W, which has a good approximation of A^{-1}. The FastICA algorithm is based on a fixed-point iteration scheme for finding a maximum of the non-Gaussianity of $\mathbf{w}^T\mathbf{x}$, where $\mathbf{w}$ is the row vector of matrix W. $\mathbf{w}$ can be calculated by using approximate Newton iteration

$$\boldsymbol{w}^+ = \boldsymbol{w} - \mu\left[\mathrm{E}\left\{\mathbf{x}g\left(\boldsymbol{w}^T\mathbf{x}\right)\right\} - \beta\boldsymbol{w}\right] / \left[\mathrm{E}\left\{g'\left(\boldsymbol{w}^T\mathbf{x}\right)\right\} - \beta\right] \tag{3}$$

$$\boldsymbol{w}^* = \boldsymbol{w}^+ / ||\boldsymbol{w}^+|| \tag{4}$$

where $\boldsymbol{w}^*$ and $\boldsymbol{w}^T$ denote the new value and transpose of $\mathbf{w}$, respectively; μ is a step size parameter that may change with the iteration count, $\mathrm{E}\{\cdot\}$ is the mathematical expectation; $g'(\cdot)$ is the derivative of contrast function $G(\cdot)$; $\beta = \mathrm{E}\{\boldsymbol{w}^T\mathbf{x}g(\boldsymbol{w}^T\mathbf{x})\}$. The selection criteria of G and recommended contrast functions can be found in the literature [8]. The algorithm firstly generates $W(0)$ with random elements. Then, the new matrix $W(k)$ is calculated using Eqs. 3 and 4. The algorithm iterates until the value of $I - |W^T(k-1)W(k)|$ is less than a threshold ε or the iteration number k is larger than a given value M.

In that study, the MGB was operated at the high input speed of 16,000 rpm and high loading level of 180 kW. The defect related feature was clearly revealed in one of the independent components spectrum separated by FastICA [9]. However, the FastICA cannot separate the recorded vibration signals when MGB run at low speed and low loading level due to the limited number of vibration sensors. In general, the

number of observations should be larger than the number of sources to have a good separation. However, it is not possible to have more accelerometers than vibration sources in MGB. As an alternative, envelop analysis was employed in this paper to reduce signal dimension to compensate of the limited number of sensors. The envelope analysis method was briefly introduced in the next section.

2.2 Envelope Analysis

Envelope analysis a commonly used method to obtain the bearing defect harmonics from the envelope signal spectrum analysis. The spectral kurtosis (SK) is one of the well-known envelope analysis techniques to detect and characterise transients in a signal. It computes a kurtosis at given frequency resolution in order to discover the presence of hidden non-stationeries and their corresponding frequency bands. The high SK value indicates that the strong impulsive component occurs. The SK can be defined as the fourth-order normalized cumulant

$$k_x(f) = \frac{\left\langle |H(n,f)|^4 \right\rangle}{\left\langle |H(n,f)|^2 \right\rangle^2} - 2, \tag{5}$$

where $H(n, f)$ is the complex envelope of $x(n)$ at frequency f.

Among newly developed envelope extraction methods, the fast kurtogram algorithm, proposed by Antoni [3], is able to compute the kurtogram over finely samples. The high efficiency of the algorithm makes it more suitable for on-line faults diagnosis. In order to increase the calculation efficiency of SK, the fast kurtogram employed binary tree algorithm to split frequency bands. Figure 1 illustrate the kurtogram representation at nodes $\{f_i; \Delta f_k\}$ of the $\{f; \Delta f\}$ plane, which is compounded of $2^K - 1$ kurtosis values.

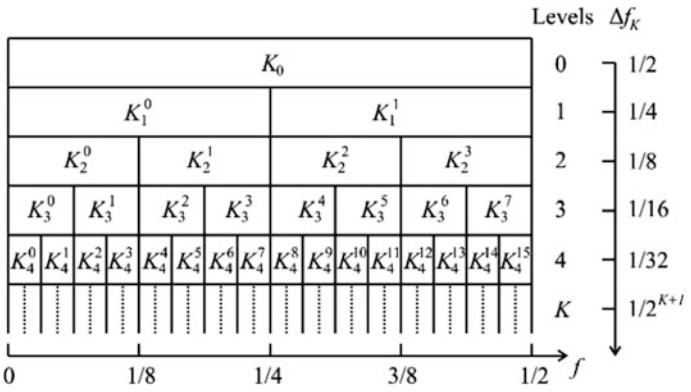


Fig. 1 Sketch of the binary frequency and frequency resolution plane of the fast kurtogram

3 Experiment Rig and Results Analysis

3.1 Experiment Rig

The experiment rig is shown in Fig. 2. The selected CS29 Category "A" SA330 Puma helicopter MGB was mounted on a platform. Flanges at both sides of the platform were designed to support the absorption dynamometer. The absorption dynamometer used air pressure to generate a clamping force between the rotating drive plates driven by the output shaft of the MGB. The level of resistance is proportional to the desired loading on the MGB. Air pressure and flowing water, which was used to remove heat generated by the frictional torque, were delivered to the absorption dynamometer using flexible external tubes.

The MGB consists of five reduction gear modules (RGMs), forward (Fwd) RGMs and after (Aft), left hand (LH) and right hand (RH) RGMs, main RGM and 2-stage epicyclic (Epi) RGM. In the 2-stage Epi RGM, the first and second stage contains 8 and 9 planets gears, respectively. The defect was seeded on one of the planetary gear bearing outer race of the second stage Epi RGM, as shown in Fig. 1b, c. The vibration signals were measured by two PCB triaxial accelerometers (356A32/NC, 100 mV/g). Two accelerometers were located at the case of Epi RGM and Fwd RGM, respectively. The vibration signals were captured by a NI cDAQ-9188XT CompactDAQ data acquisition with sampling frequency of 25.6 kHz.

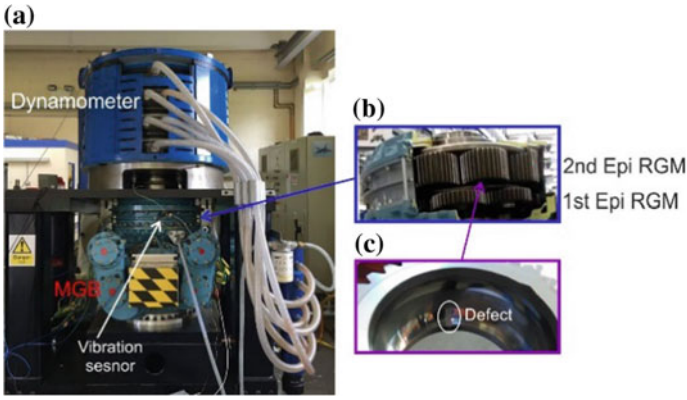


Fig. 2 **a** Experimental rig; **b** 2-stage Epi RGM; **c** Seeded defect at the bearing outer race

3.2 Bearing Defect Detection Methodology

In this study, the experiment was conducted under different loading level and input speed, which were tabulated in Table 1. The seeded bearing defect was simulated by machining a rectangular slot across the bearing outer race with length, width and depth of 10, 4 and 3 mm, respectively. Once bearing outer race defect occurs, the ball impacts on the defect every time when it passes the defect on the outer race. The impact generated vibration can be regarded as a defect signature for fault diagnosis. The defect signature so called bearing outer race defect (ORD) frequency f_{ORD} can be calculated by using

$$f_{ORD} = \frac{N}{2}\frac{S}{60}\left(1 - \frac{d}{D}\cos\alpha\right), \tag{6}$$

where $N = 13$ is the number of rollers; S is planet gear speed of the second stage Epi RGM. $d = 12.5$ mm is the diameter of roller; $D = 63.65$ mm is the pitch diameter and $\alpha = 0$ is the nominal contact angle. The calculated f_{ORD} is equal to 59.88 and 68.55 Hz when the input speed is 14,000 and 16,000 rpm, respectively.

Although the ORD frequency is independent of loading level, the low speed and low loading level made the f_{ORD} difficult to be detected using the FastICA scheme. Therefore, the fast kurtogram algorithm was firstly employed to reduce the dimension of signal.

3.3 Validation and Analysis

Two triaxial accelerometers generate a total six observations. The worst scenario of low input speed of 1400 rpm and low loading level of 100 kW was utilized to illustrate the effectiveness of the combination of two methods. Firstly, the envelopes of seven observations were extracted by the fast kurtogram algorithm. Figure 3a, b show the original vibration signal waveforms of the sensor 1 in tri-axis and their corresponding envelope extracted using the fast kurtogram algorithm. The optimal level and filtering band were set to 7 and 150 Hz, respectively to keep the same length of the envelopes. Although each signal has different optimal level and filtering band, these two values were chosen based on the most common values of

Table 1 Experimental conditions for bearing outer race fault

Input speed (rpm)	Loading level (kW)	BPFO (Hz)
14,000	100	59.88
14,000	180	59.88
16,000	100	68.55
16,000	180	68.55

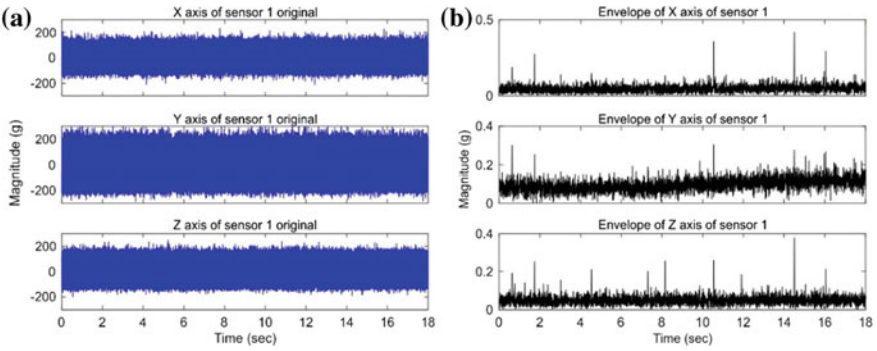


Fig. 3 **a** Original waveforms of sensor 1 in tri-axis; **b** envelope of sensor 1 waveforms in tri-axis

each signal. Similarly, Fig. 4a, b show the original vibration signal waveforms of the sensor 2 in tri-axis and their corresponding envelope.

Finally, the FastICA was utilized to separate vibration sources from 6 mixtures. The separated independent components were transferred from time domain to frequency. The spectrum of these independent components was shown in Fig. 5. Although there was still some mixing frequency components presented in the separation, the defect related feature f_{ORD} of 59.88 Hz was clearly observed in the spectrum of the independent component 4 and 6.

The same procedure was applied to the rest operating condition. The defect related feature f_{ORD} of 59.88 Hz were observed at the spectrum of one of the independent component when the MGB operated at 14,000 rpm input speed, as shown in Fig. 6. Likewise, the defect related feature f_{ORD} of 68.55 Hz were observed at the spectrum of one of the independent component when the MGB operated at 16,000 rpm input speed.

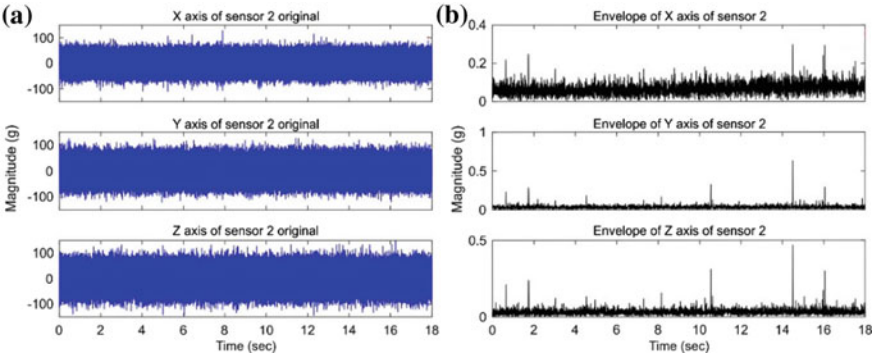


Fig. 4 **a** Original waveforms of sensor 2 in tri-axis; **b** envelope of sensor 2 waveforms in tri-axis

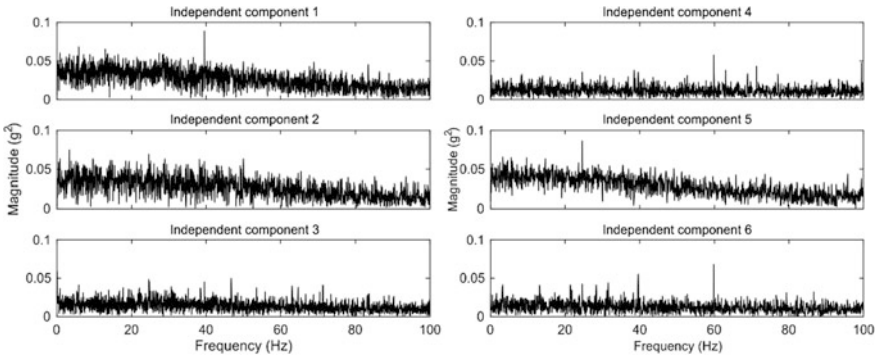


Fig. 5 Spectrum of independent components separated by FastICA

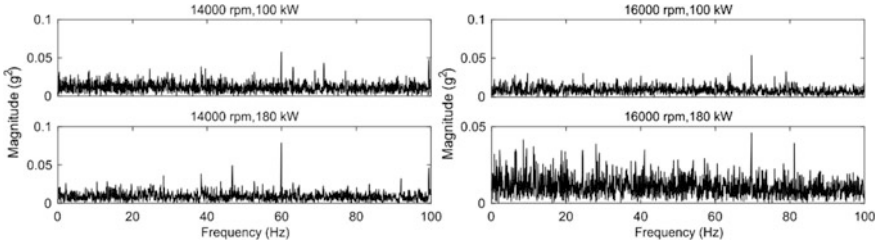


Fig. 6 Independent components including f_{ORD} under different input speed and loading level

4 Conclusion

Since the vibration sensors are usually mounted on the case of helicopter MGB, vibration signal are attenuated and contaminated resulting from strong background noise and multipath transmission. Under these circumstances, it is difficult to extract faults related features from recorded vibration signals. This paper investigates the possibility of combining the fast kurtogram algorithm and the FastICA algorithm to extract faults related features from recorded vibration signals. The fast kurtogram algorithm was firstly utilized to reduced signal dimension to eliminate the constraint of limited number of sensors. Then, the FastICA algorithm was employed to separate vibration sources. The experiment was conducted under different loading level and input speed. In all cases, the defect related feature f_{ORD} was clearly observed at the spectrum of one of the independent component. These results prove the efficacy of the combination of the fast kurtogram algorithm and the FastICA algorithm.

References

1. Stalin S (2014) Fault diagnosis and automatic classification of roller bearings using time-domain features and artificial neural network. Int J Sci Res 3:842–851
2. Wang Y, Liang M (2011) An adaptive SK technique and its application for fault detection of rolling element bearings. Mech Syst Signal Process 25:1750–1764
3. Antoni J (2007) Fast computation of the kurtogram for the detection of transient faults. Mech Syst Signal Process 21:108–124
4. Guo W, Tse PW, Djordjevich A (2012) Faulty bearing signal recovery from large noise using a hybrid method based on spectral kurtosis and ensemble empirical mode decomposition. Measurement 45:1308–1322
5. Ye H, Yang S, Yang J, Ji H (2006) Vibration sources identification with independent component analysis. In 2006 6th World Congress on Intelligent Control and Automation, pp 5814–5818, 0–0 0 2006
6. Guo Y, Na J, Li B, Fung R-F (2014) Envelope extraction based dimension reduction for independent component analysis in fault diagnosis of rolling element bearing. J Sound Vib 333:2983–2994
7. Roan MJ, Erling JG, Sibul LH (2002) A new, non-linear, adaptive, blind source separation approach to gear tooth failure detection and analysis. Mech Syst Signal Process 16:719–740
8. Hyvarinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw 10:626–634
9. Duan F, Corsar M, Zhou L, Mba D (2017) Using independent component analysis scheme for helicopter main gearbox bearing defect identification. In 2017 Prognostics and System Health Management Conference, Accepted, to be published 2017, pp 1–5

Forecast Model for Optimization of the Massive Forming Machine OEE

Markus Ecker and Markus Hellfeier

Abstract The EMuDig 4.0 project target is to link all relevant systems and sensors inside a massive forming company with influencing systems from outside as basis for a smart factory. Data and information extracted from integrated sensors and systems in connection with new methods for analysis and algorithms shall be used to optimize the OEE of massive forming machines. The primary target is a quick and direct information to indicate machine irregularities as soon as they appear. The available information not only allows efficiency improvement of single process steps. It supports the optimization of the whole value-added chain.

1 Introduction

The massive forming industry is affected by changing market and product requirements. Increasing demands on quality and new challenges, for instance due to lightweight design and electric mobility, will be the major affect for the massive forming industry in the future or already today. The necessity to produce the different products with optimized OEE is essential in order to stay competitive in the worldwide massive forming industry. This requires more effective engineering for parts and processes, greater efficiency, widening of the process limits, global networking of supply chains or increased flexibility through smaller lot sizes. One of the central questions facing the engineering in the massive forming industry (design of the product properties, processes and plants) here will be to what extent the planning and design process, the efficiency and the stability of the whole value-added chain can be further increased through digital transformation and

M. Ecker (✉) · M. Hellfeier
SMS Group GmbH, Mönchengladbach, Germany
e-mail: markus.ecker@sms-group.com

M. Hellfeier
e-mail: markus.hellfeier@sms-group.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_15

broader know-how bases. It is already recognisable today that the customers of the massive forming industry see digital networking as offering great opportunities for meeting their future challenges. They will therefore very soon start demanding the strategic approach in the sense of Industry 4.0 also from their suppliers. Studies show that large companies are already working far more intensively on Industry 4.0 solutions.

2 Overall Project Description

The research project EMuDig 4.0 is to provide the massive forming industry with an improved Overall Equipment Effectiveness (OEE) through the integration of digital networking of the whole value-added chain. Scientific institutions are to be interconnected with industry partners to put creative ideas and research results into practical application. The research project is funded by the Federal Ministry of Economic Affairs and Energy (BMWi). The EMuDig 4.0 project gives an important impulse for further development of Industry 4.0 within the massive forming industry domain.

The activities under the research project are subdivided into the following subprojects to form a consortium of four academic and three industrial partners:

Subproject 1, **Generation of raw material**:
OTTO FUCHS KG (process chain aluminium)
Hirschvogel Umformtechnik GmbH (process chain steel)
Subproject 2, **Forming process**:
University of Stuttgart, Institute of Metal Forming Technology
Subproject 3, **Production plants**:
SMS group GmbH
Subproject 4, **Production tools**:
South Westphalia University of Applied Sciences, Laboratory for Metal Forming (LFM)
Subproject 5, **Logistics process**:
University of Stuttgart, Institute of Industrial Automation and Software Engineering
Subproject 6, **Factory cloud and data analytics**:
TU Dresden University, Centre for Information Services and High Performance Computing (ZIH)

The involvement of the associated partner “Industry Association of Massive Forming” will ensure that the results of this research project are distributed to the companies of the German massive forming industry.

2.1 Problem Definition

The massive forming process already today comprises a challenging value-added chain:

- Raw material production
- Tool making
- Production of the raw part
- Subsequent heat treatment
- Machining and further finishing to create the part ready for installation

With a view to the challenges of tomorrow already described, we have to see that we do not have a sufficient command of the complex interaction of the numerous parameters and their various interrelationships with the current approach. The challenges of massive forming today can be summarised as follows:

With the state-of-the-art in massive forming, the process stability is not described by process parameters. The process stability will measure with a great time lag at the end of the forming and/or heat treatment process using random sample tests of product properties (e.g. cpk value). The product diversification cannot be attributed to individual process stations, so that a clear identification of the causes is achieved only very seldom. Open-loop or closed-loop interventions in the production process are only possible with a significant delay, if at all. The limited process stability reduces the plant efficiency which leads to considerable financial losses with the capital-intensive forming plants with their hourly rates of up to EUR 1,600.

The **lack of correlation** between the process and the product properties prevents a strategic overall optimisation of the process with the goal of attaining a sustainable boost in efficiency.

Information on the tools is essentially limited to the production volume achieved that can **vary by up to 100%** without any identifiable reasons.

Machines and plants are repaired in the event of standstill and/or undergo preventive maintenance at fixed intervals. **There is a lack of the necessary data** as well as suitable forecast models for load-oriented maintenance cycles.

The lack of networking between the individual links of the value-added chain and the part properties prevents an overall strategic optimisation of the process with the goal of attaining a sustainable boost in efficiency.

The massive forming process is generally a mass production process in which several hundred up to several thousand parts are produced and logistically tracked/identified in a production lot; single part tracking and identification is employed today only in exceptional cases (e.g. aerospace industry) and then at great expense. It is therefore not possible at the present time to **correlate parameters** of the value-added chain with the property of a part.

2.2 Project Aim and Methodological Approach

This project aims to show that, and in how far, OEE and cost effectiveness of two different value-added chains for massive forming of steel with large batch sizes and aluminium with small batch sizes can be improved through digital technologies and control concepts.

The envisaged solution involves adapting and integrating available digital technologies in the engineering of the whole value-added chain to the needs of massive forming. In order to achieve this, a methodological approach is to be developed, tested and evaluated as part of a joint research project using two demonstrator and one model application. The data from complex production facilities can be recorded, saved, processed and analysed along the value-added chain in order to boost the overall plant efficiency through interdisciplinary cooperation between plant owner, plant manufacturer, cloud operator, software producer and science.

With the aim of achieving fundamental, reliable and transferable results, the following elements of the value-added chain are essential for the efficiency of a process and are taken into consideration in the planned demonstrator/model application:

The raw material production, with the aim to interlink the semi-finished product (starting material) technologically and logistically with the further processing process chain ("predictive material property"). This will be achieved with a unique "digital fingerprint" with respect to semi-finished production and property data.

The production process of forming and heat treatment, with the goal of achieving reliable control of all the product properties ("predictive quality") and reducing the scrap rate. This is attained through close links with the raw material production, the tool and plant condition, the process engineering and involving an active control system for influencing the part properties.

The production plant, with the goal of obtaining a load-dependent prediction of plant conditions and corresponding preventive service and maintenance measures ("predictive maintenance") through close links with the forming process, with maintenance and with spare parts procurement logistics.

The production tools, with the aim of predicting the tool condition/wear in relation to the "tool life history" in order to enable corresponding preventive measures. For example adapting the lubrication/spraying condition to the prevailing situation or initiating the planning of a tool change ("predictive tool management") through close links with tool engineering, logistical control of tool making and the incorporation of 3D technologies for in-process assessment of tool states.

The logistics process, with the goal of logistical control and traceability down to the "smallest possible sub-quantities", ideally single part tracking of a complete production lot ("smart logistics").

The factory cloud and data analytics, with the aim to select and further development of methods to implement a system for analysis of the efficiency of the whole value-chain ("big data").

The aspired result of the presented subproject "Production Plant" in Chap. 3 is a forecast model allowing load depending forecasts of the plant condition and according maintenance measures (predictive maintenance). The following milestones have to be achieved to realize this result:

The identification of relevant reasons for machine downtimes and development of a specific sensor concept. Target is collection and quality control of OEE-relevant machine data.

Investigation of correlations between massive forming machine data and performance failure aiming to develop the specific forecast model.

Identification of predictive maintenance measures to improve OEE of massive forming machines.

Realization of an onsite demonstrator with a massive forming application including development of process relevant interfaces and linked with hard- and software for data analysis.

3 Subproject "Production Plant"

3.1 Introduction

The production plants in the massive forming industry manufacture a wide range of products and materials of steel, aluminium, nickel and titanium. Product geometry and product material represent the main influencing factors on the load collective to which the massive forming machines are exposed. This results in a specific wear situation for the respective machine. In addition, the load on the plants and the quality and quantity of their maintenance are further factors influencing the specific wear situation.

Determination of the individual wear situation of a plant in massive forming and the derivation of predictions for the availability-critical wear states are major problems facing the massive forming industry. **Valid statements on the wear allowance** of a machine are consequently difficult to make.

3.2 Current Status/State of the Art Technology

Condition monitoring is currently a widespread concept for describing the wear condition and diagnosing the wear allowance for a wide range of machines and plants as basis for predictive maintenance. These will be accomplished using for instance temperature, vibration and torque sensors (Fig. 1 Predictive maintenance principle [1]).

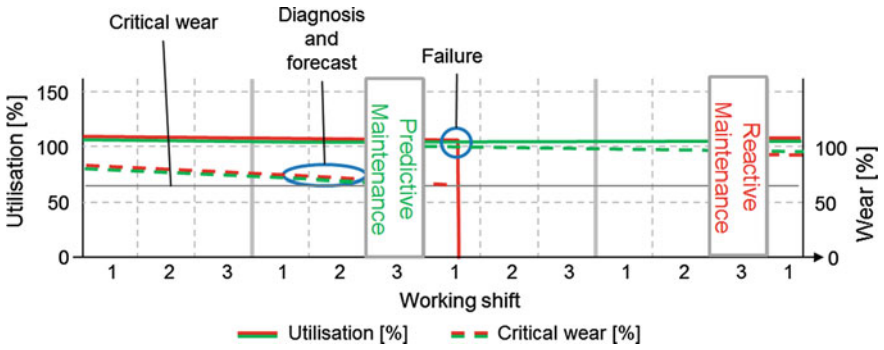


Fig. 1 Predictive maintenance principle [1]

For massive forming machines, these diagnoses are far more difficult to make. Many of the parameters used for conventional condition monitoring can be affected or superimposed by ambient and process influences, e.g.:

Vibrations in and around the machine that are superimposed e.g. by vibrations due to bearing wear during the forming process on mechanical closed-die forging presses.

Particularly during hot forming, temperature measurements can also be influenced by the heated workpiece.

Soiling of the plants caused by separating agents or scale.

These exemplary ambient and process influences result in a significant limitation on the validity of the condition monitoring. As a result, optimum use is not made of the wear allowances of critical components and assemblies, with the following negative impacts on the overall plant efficiency:

Downtimes
Quality issues
Maintenance schedules too short or too long

3.3 Target

The target for this subproject is to improve the overall equipment effectiveness (OEE) by using predictive maintenance for the production plants. For the two demonstrator plants, predictive models are to be developed in cooperation between Engineering and operative experts from plant manufacturers and plant operators that enable predictions of availability-limiting events to be made. These are then to be used to improve the availability of the plants.

Networking of the various sub-projects and the associated exchange of information, e.g. on the tool condition and the part quality, should then help to enhance the validity of the predictive models.

3.4 Innovation

The main innovation of the sub-project does not lie in the area of the measurement data acquisition. This lies in the development of a specific predictive model for the predictive maintenance of massive forming machines that includes the largest possible number of condition parameters.

The results of this sub-project are as follows:

1. **Holistic measurement concept** for early indication of part and assembly damage for massive forming machines in two representative production lines in the plants of the application partners.
2. **Prediction model** derived from the measurement concept for forecasting unscheduled standstills, e.g. on the basis of remaining wear allowances.
3. **Optimised maintenance measures** for massive forming machines to sustain the plant availability (e.g. time, scope, optimum use of wear allowance, etc.).

3.5 Practical Application

During the preliminary work on predictive maintenance outside the field of massive forming, not only innovative measuring methods such as video motion amplification, but also classic measuring methods were used.

Torques
Vibrations
Temperature
Forces
Oil flow, particles and viscosity,

For a closed-die forging press, the set-up for data collection and analysis planned for this project is shown in Fig. 2.

The SMS Genius Condition Monitoring system is integrated into the forging press for this purpose. This receives values out of the PLC system, signals from sensors already installed and additional installed sensors according to the demands are defined for the early indicators for the failure of availability-critical components and assemblies. The Genius CM System can be linked to an MES/ ERP system on site.

Working principle of Genius Condition Monitoring

The Genius CM is based on an independent software solution that has been successfully invented by the SMS group. It all starts with Genius CM recording your plant's process data and sensor signals. This information goes into the analysis of the actual component status. Out of this comes a perfect picture of the process-dependent conditions which the system can clearly display.

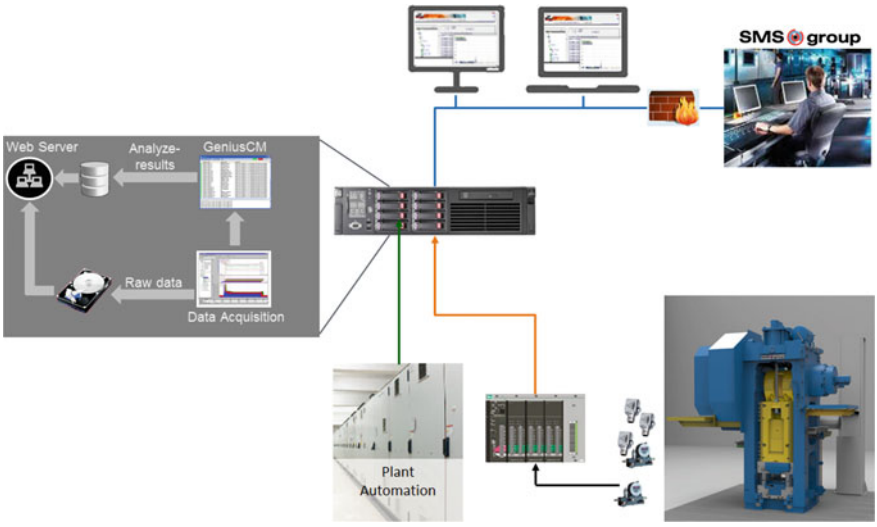


Fig. 2 Example setup for an closed-die forging press [2]

Unique to Genius CM is its modular design so the user can tailor the configuration that he needs. There are field units he can use for online data capture on site. That provides for scalability and problem-free expansion. The system monitors a large number of critical or at-risk components simultaneously.

Condition Evaluation

After collection the data of all different kind of signals the correct conclusions have to be made. For reasons that some damages can occur very quickly, an online evaluation which represents the current status of the components in real time is essential. In addition, the process- oriented evaluation offers the possibility to record dependencies and effects between production parameters and system state (Figs. 3 and 4).

In order to develop a specific predictive model for the predictive maintenance of a massive forming machine, the implementation model shown in Fig. 5 will be applied.

3.6 Conclusion

A forecast model for optimization of the massive forming machine OEE has a beneficial impact for a sustainable asset management and to improve the competitiveness not only in the area of the massive forming industry.

The major challenge of this project is to link all relevant influencing systems inside and outside of a massive forming company as basis for a smart factory. Due

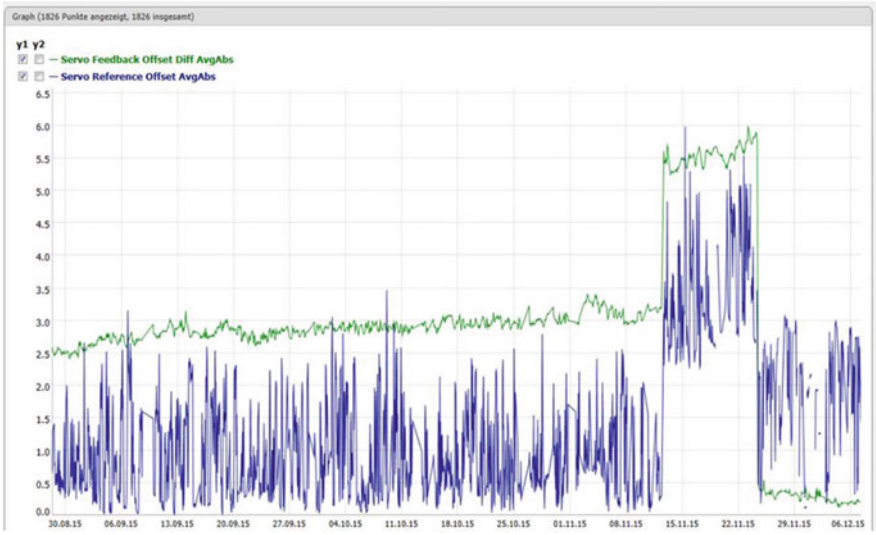


Fig. 3 Online evaluation [2]

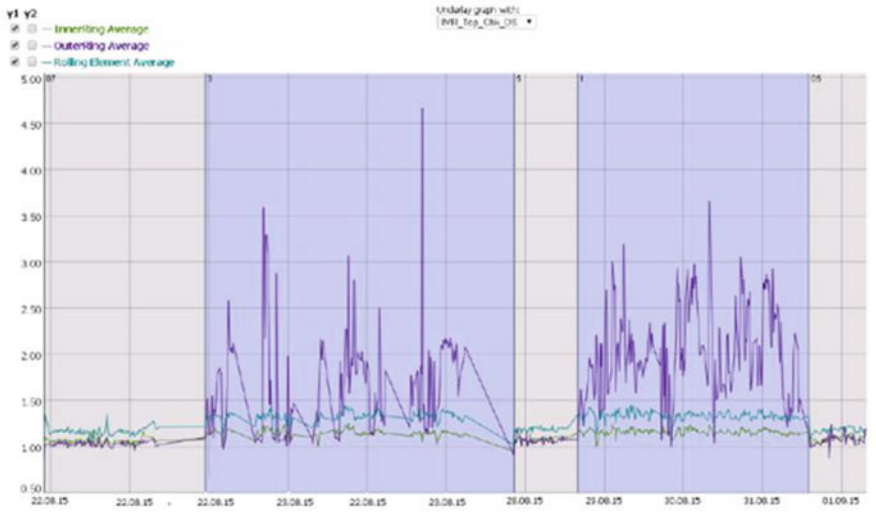


Fig. 4 Process oriented evaluation [2]

to the integration of digital networking of the whole value-added chain a big amount of data will be collected and analysed. The data from complex production facilities can be recorded, saved, processed and analysed in order to boost the

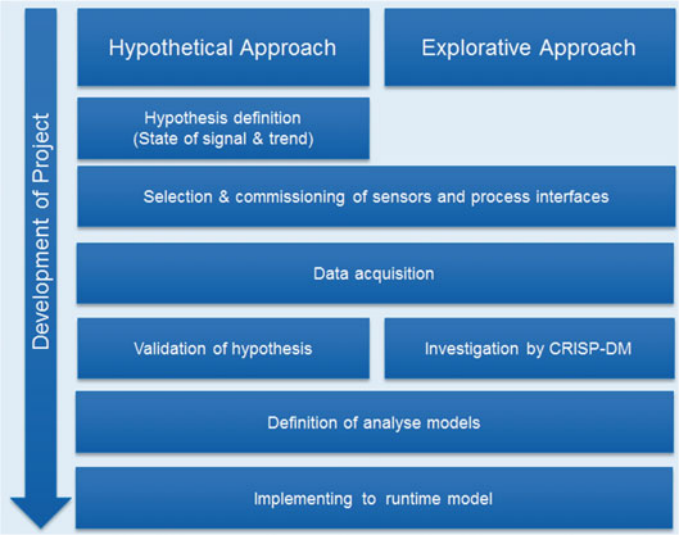


Fig. 5 Implementation model [2]

overall plant efficiency through interdisciplinary cooperation between plant owner, plant manufacturer, cloud operator, software producer and science.

During the early project phase it became obvious that the bilateral interaction of the involved scientific institutions and industry partners as well as a clearly structured appliance is of paramount importance for a successful project.

References

1. Stüer P (2015) RWTH Publications, [Online] Available at: https://publications.rwth-aachen.de/record/465189?ln=de [Haettu 15 05 2017]
2. SMS group GmbH (2016) Düsseldorf: SMS group GmbH

Quantitative Bowtie Risk Model: An Agile Tool in the Utility Toolkit

Daniel Falzon

Abstract A major challenge faced by asset-intensive organisations is understanding the influence individual assets have on their parent facilities, and further, understanding the level of influence individual assets have on an organisation meeting its strategic goals. The need for an agile decision-making tool which provided this insight was identified and a Quantitative Bowtie Risk Model (QBRM) was developed. Probabilistic data such as condition rating, as-new failure frequency and barrier effectiveness are entered as modelling inputs. The quantitative likelihood and consequence outputs are then calculated and presented on the Bowtie graphic, which are mapped to the organisation's Corporate Risk Heat Map. Risk-dollars are used to demonstrate annualised fiscal risk exposure, providing cost-benefit optimisation of mitigation options. This allows the analyst to explore proposed mitigation options by altering inputs to represent the proposed options and then compare against the base case bowtie. The QBRM allows the analyst to explore TOTEX asset decisions, optimising risk, cost and performance across a combination of options. The QBRM also provides data confidence levels to the analyst by using qualitative and quantitative ratios to strengthen business case justification and helps identify knowledge areas needing improvement. This paper demonstrates the value of the QBRM through a real case study centred on a water pumping station.

1 Introduction

As more and more organisations seek to align their asset management practices to ISO 55000 [1], it is imperative that they can demonstrate data-led decision-making which shows a clear line-of-sight between their asset interventions and the beneficiaries.

D. Falzon (✉)
Lead Asset Management Planner, SA Water, Australia
e-mail: daniel.crystal@bigpond.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_16

Brown and Humphrey state that in its most general sense, asset management is a business approach designed to align the management of asset-related spending to corporate goals. The objective is to make all infrastructure-related decisions according to a single set of stakeholder-driven criteria. The payoff is a set of spending decisions capable of delivering the greatest stakeholder value from the investment dollars available [2].

It is important for asset-intensive organisations to be able to quantify what influence assets have on the facility within which they are housed, and ultimately, the corporation's risk profile. Further, as facilities mature, safeguards may become less effective or degraded meaning risks may increase [3]. This insight into the risk profile of individual assets and the subsequent impact on the corporate risk position is essential for organisations in managing their assets in the most prudent and efficient manner. Nordgard and Solum note that in electrical distribution organisations, companies are developing strategies for maintenance and reinvestments, where the emphasis on cost effectiveness is balanced with other important dimensions of risk [4].

Evaluation of the author's employer, a large bulk water supplier, showed there was no risk methodology in use that quantitatively demonstrated the link between asset risk, facility risk and corporate risk. Further, there was no measurement of quantitative versus qualitative data sources to demonstrate decision-making robustness, or reporting of proactive versus reactive cost investment. Following an investigation into the various risk assessment methods available, a gap was identified in that a risk methodology was needed which provided the analyst with the ability to simulate asset interventions relating to changes in condition rating and system interventions using maximum acceptable outage times. The impact of changes at both the facility and system level needed to align with the organisation's specific risk categories and simulate real improvements to the organisation's risk position. The tool needed to present complex inputs and outputs on a single graphic and in such a way that it could be understood by a range of stakeholders with varying asset knowledge. To meet these requirements a Quantitative Bowtie Risk Model (QBRM) was developed. Because of their graphical nature, the biggest advantage of bowtie diagrams is the ease to understanding risk management by upper management and operations groups [5].

This paper firstly outlines the concept of the bowtie method and its advantages. Application of the concept as it aligns to an organisation's corporate risk matrix and risk categories is discussed. The proactive modelling input methodology for threat identification, condition rating and barrier effectiveness is described, followed by reactive modelling methods including consequence identification, consequence barrier effectiveness and risk-dollars. Finally, modelling outputs including costs and benefits of simulated asset interventions are presented.

2 Methods and Approach

The bowtie method is a risk assessment methodology commonly used in process safety assessment. This is primarily because it is an ideal way of representing qualitative and quantitative causes and effects that surround the event under analysis [6]. Probability bowties are used in managing the risks of major hazard facilities, being particularly useful where HAZOP teams have industry and plant-specific experience of the frequency of incidents and the effectiveness of control measures [7]. It is also used to determine adequate levels of incident barriers in assessment of safety cases [8]. The bowtie method has been used in other utilities, assisting in the decision-making and prioritisation of water pipe renewal [9], and risk-based decision making in the oil and gas sector [10]. Research in dynamic bowtie modelling has been undertaken [11] which has sought to overcome the static nature of the bowtie method. In considering the objectives referred to in Sect. 1, the bowtie method was recognised as being the most likely to facilitate meeting the challenges in the one tool.

The approach used to develop the QBRM was to add the parameter of Maximum Acceptable Outage (MAO) (refer Fig. 1) and assign it to the top event and present the model inputs and outputs on a bowtie graphic aligning the consequences to the organisation's corporate risk categories (refer Fig. 2). In addition to the MAO, condition-based failure likelihoods were added to the individual threats, to simulate potential future asset improvement or deterioration. Values are assigned to modelling parameters such as failure frequency and barrier effectiveness. The quantitative likelihood and consequence outputs are then calculated and presented on the Bowtie graphic.

Risk-dollars are presented to demonstrate annualised fiscal risk exposure and provide cost-benefit optimisation of mitigation options. Conditional event probabilities are polarised against various categories (e.g. most probable event, longest outage event) and automatically mapped to the organisation's corporate risk matrix where risks are evaluated against the corporate risk appetite for the relevant risk category.

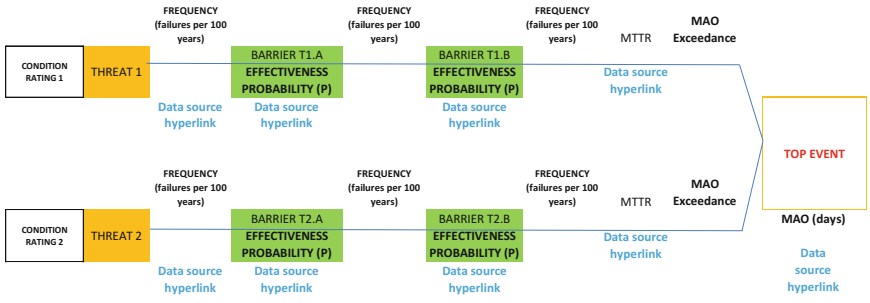


Fig. 1 Simplified QBRM proactive side

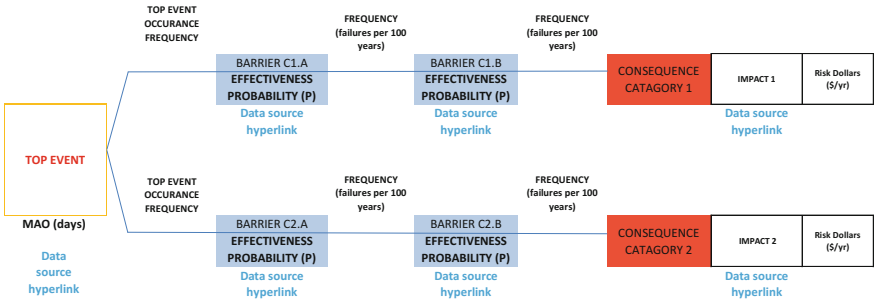


Fig. 2 Simplified QBRM reactive side

Once a top event is selected for analysis (e.g. pumping station failure) and all relevant parameters are entered, a "base-case" bowtie is produced. The base-case bowtie will highlight the threats which are of concern. The base-case is then presented to the various key stakeholders who are invited to propose the various mitigation options. Potential "future-state" bowties are then developed that represent the proposed mitigation options (or combination of options). Options that simulate a proposed future state can include "hard" interventions such as capital replacement or refurbishment, or "soft" interventions such as the installation of sensors or increased redundancy (i.e. barrier improvement). Each future-state bowtie is then compared against the base-case bowtie and assessed in terms of their net present value versus their commensurate reduction in risk.

3 Modelling Inputs

Causes that are identified as having a mean-time-to-recover (MTTR) greater than the MAO for the top event are assigned as threats, along with the relevant as-new (i.e. infant mortality) threat occurrence frequency. The source from which the data is collected must be recorded in the data entry table as a hyperlink. Failure to record where the data has been sourced will show as an alert on the graphic. The data source must also be assigned a qualitative or quantitative descriptor for future assessment of model strength.

Following specification of the relevant threats, a condition rating is assigned to the assets that represent the threats. The condition rating failure frequency (F_c) overrides the as-new threat occurrence frequency resulting in a condition-modified occurrence frequency in accordance with the organisation's agreed asset condition rating matrix (refer Table 1).

If the failure frequency based on the condition rating matrix has a lower return interval than the as-new failure frequency, then the as-new failure frequency is used. Condition rating assessments are noted in terms of their quantitative or qualitative nature, e.g. ultrasonic thickness testing on a discharge pipe is quantitative, however a purely visual assessment would be qualitative.

Table 1 Condition rating matrix

As New	Initial signs of deterioration	Satisfactory	Poor	Urgent action
Expected failure well above 25 years	Expected failure within 25 years	Expected failure within 10 years	Expected failure within 5 years	Expected failure within 12 months
$F_c = F_{new}$	$F_c = 0.04$	$F_c = 0.1$	$F_c = 0.2$	$F_c = 1$

Following the assignment of condition ratings to the various asset threats, consideration of threat barriers and their relevant effectiveness probabilities are entered in the model (refer Eq. 1). An equivalent process is followed to define any secondary threat barriers and the resulting secondary residual threat frequency (refer Eq. 2). Costs are assigned to each primary and secondary barrier and summed to give a total proactive cost.

The top event frequency (refer Eq. 3) represents the total likelihood of any threat materialising. Upon consideration of the relevant threats to the top event, consequence categories are assigned in line with the organisation's corporate risk matrix, e.g. water security, water quality, financial, environmental WHS, etc., and risk impacts are also assigned to each category. The top event frequency for a given risk category is calculated by summing the frequencies applicable to the relevant risk category (refer Eq. 4).

To account for the variable effectiveness consequence barriers had in mitigating the various threats contributing to a consequence category, a weighted effectiveness is used and presented on the bowtie (refer Eq. 5).

4 Case Study: Raw Water Pumping Station

To demonstrate the real value of the QBRM, a base-case bowtie model was developed for the Raw Water Pumping Station (RWPS) at Tailem Bend in South Australia. The top event was defined as the "loss of required throughput". Workshops were held with key stakeholders so that threat and consequence data could be collected and verified. The MAO was set at 0.521 days (based on available downstream storage) and the threats, conditions, proactive barriers and MAO exceedances were identified as shown in Table 2.

The pump station is designed such that in the event of an outage, a mobile diesel pump can be connected to supply limited water. This is the only reactive measure that can be taken; however, this measure is not effective against the threat of main isolation valve failure, due to the mobile pump discharge tie-in point being upstream of the valve. Consequently, the mobile pump had an overall weighted effectiveness of 9.92%. The consequences were aligned with the applicable consequence categories on the corporate risk heat map, namely Water Supply Security and Financial (ref. Fig. 3).

Table 2 Threat, barrier and MAO exceedance identification

Threats	Condition	Primary barrier	Secondary barrier	MTTR (days)	MAO exceedance (days)
Pump	Satisfactory	Redundant pump set	Redundant pump	3	2.48
Motor	Satisfactory	Redundant pump set	Redundant pump set	3	2.48
Switchboard	Initial signs	Electromechanical relay	–	5	4.48
Main isolation valve	Poor	None	–	7	6.48
Internal pipework	Satisfactory	NDT monitoring	–	14	13.48

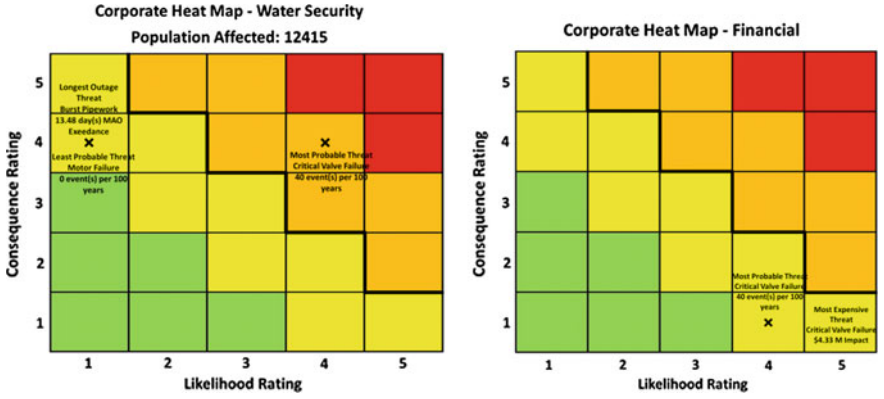


Fig. 3 Corporate risk heat map outputs of base-case bowtie

It was identified from the model that the main isolation valve was the principal threat driving the high risk (Water Security) at the facility due its poor condition, lack of proactive barriers, and ineffective reactive barrier.

Following consideration of the principal threat, four mitigation options were identified aimed at treating the risk of major valve failure including a system intervention downstream of the facility (Option 1a), simulated by increasing the MAO. Replacement of the valve was simulated by toggling the condition rating. The comparative results of each intervention bowtie are presented in Table 3.

The cost associated with each threat was modelled using net present value financial evaluation over a 30-year period. Consequence costs included lost revenue, collateral damage, and water carting.

Following evaluation of the model outputs it was clear that the optimised decision was to implement Option 1d. It was evaluated that whilst Option 1b had a lower

Table 3 Summary of water security base-case and intervention bowties

	Base case	Option 1a	Option 1b	Option 1c	Option 1d
Description	Current state OPEX	Increase system storage by 9 ML CAPEX	Valve replacement (avoid failure) CAPEX	Use pipe spool in lieu of valve (accept failure) OPEX	Combination of option 1b and option 1c CAPEX
MAO (days)	0.521	0.8	0.521	0.521	0.521
MAO exceedance for failure mode (days)	6.48	6.2	6.48	0.48	0.48
Top risk event frequency (p.a.)	0.454	0.454	0.054	0.454	0.054
Corporate failure residual frequency (p.a.)	0.409	0.409	0.010	0.209	0.009
Risk dollars ($k/year)	(1733)	(1658)	(3.6)	(128)	(0.27)
NPV ($k)	(3385)	(3239)	(100)	(343)	(115)
Likelihood rating	4	4	1	4	1
Consequence rating	4	4	4	1	1
Corporate risk rating (security of supply)	16	16	4	4	1

Table 4 Vulnerability ratios and data source quality

Vulnerability ratios		Source data quality	
Threat-barrier quick ratio	0.83	Quantitative references	18
Consequence-barrier quick ratio	2.00	Qualitative references	14
Proactive-reactive cost ratio	2.95	Quantitative-qualitative ratio	1.29

life-cycle cost, investing in Option 1d at an increased cost of $15 k provided a reduction in risk by 3 orders of magnitude, and therefore a greater cost-benefit ratio.

The model provided qualitative versus quantitative ratios (refer Table 4) for the data used in the model, so that confidence levels could be used to strengthen the business case and identify knowledge areas needing improvement.

The QBRM enabled a data-led, effective and targeted asset management decision to be made, resulting in the addition of Option 1d to the delivery team's 4-year program of work.

5 Conclusions

Modern asset-intensive organisations must be able to demonstrate prudency in their asset decision-making. The QBRM revealed the most critical assets within the pumping station facility and allowed for the optimisation of capital and non-capital interventions.

This modelling approach allowed the analyst to simulate the effectiveness of threat and consequence barriers, facilitating the exploration of trade-offs between hard and soft asset interventions. By using risk-dollars, it is possible to quickly evaluate the costs and benefits across various asset interventions. The automatic reporting of key modelling ratios such as threat-barrier, consequence-barrier, and qualitative-quantitative data sources provides the asset manager with confidence levels relating to the development of business cases. The QBRM is simple and agile in its application and provides a clear representation of individual asset influence on corporate risk categories that can be communicated to stakeholders with varying asset knowledge. It proved to be an effective tool for evaluating the links between asset risk and corporate risk.

6 Equations

$$F_{tp} = F_c(1 - P_{tp}) \tag{1}$$

$$F_{ts} = F_{tp}(1 - P_{ts}) \tag{2}$$

$$F_t = \sum_{i=1}^{N} F_{ts,i} \tag{3}$$

$$F_{tr} = \sum_{i=1}^{N} F_{tsr,i} \tag{4}$$

$$E_w = \frac{\sum_{i=1}^{N} \left(F_{ts,i} \times P_{cp,i}\right)}{\sum_{j=1}^{N} F_{ts,j}} \tag{5}$$

where:

F_{tp} primary residual threat frequency
F_c condition rating failure frequency
P_{tp} primary threat barrier effectiveness probability
F_{ts} secondary residual threat frequency
P_{ts} effectiveness probability of secondary threat barrier
F_t top event frequency
N total number of threats applicable to the risk category
F_{tr} top event frequency for a given risk category
F_{tsr} secondary residual threat frequency by risk category
E_w weighted effectiveness
P_{cp} effectiveness probability of primary consequence barrier
$F_{ts,i}$ secondary residual threat frequency for a given category

Acknowledgements The author wishes to acknowledge Kiyoon Kim, Luke Dix and Angus Paton of SA Water Corporation.

References

1. International Organization for Standardization (2014) ISO 55000 asset management—overview, principles and terminology, s.l.:s.n.
2. Brown R, Humphrey B (2005) Asset management for transmission and distribution. IEEE Power Energ Mag 3(3):39–45
3. Emery DS (2014) Operational risk using bowtie methodology. Edinburgh, UK, IChemE
4. Nordgård DE, Solum G (2009) Experiences using quantitative risk assessment in distribution system asset management. Prague, CIRED
5. Saud YE, Israni K, Goddard J (2013) Bow-tie diagrams in downstream hazard identification and risk assessment. Process Saf Progr 33(1):26–35
6. Ouache R, Adham AA (2014) Reliability quantitative risk assessment in engineering system using fuzzy bow-tie. Int J Curr Eng Technol 4(2):1117–1123
7. Cockshott JE (2005) Probability Bow-ties: a transparent risk management tool. Trans IChemE, Part B, Process Safety and Environmental Protection 83(B4):307–316
8. De Dianous V, Fievez C (2006) ARAMIS project: a more explicit demonstration of risk control through the use of bow–tie diagrams and the evaluation of safety barrier performance. J Hazard Mat Osa/vuosikerta 130:220–233

9. Park J, Koo M, Kim J, Koo J ei pvm An assessment of risks with the bow-tie method and designing plans for lowering risks
10. Anjuman Shahriar RSST (2012) Risk analysis for oil and gas pipelines: a sustainability assessment approach using fuzzy based bow-tie analysis. J Loss Prev Process Ind 25(3):505–523
11. Khakzad N, Khan F, Amyotte P (2012) Dynamic risk analysis using bow-tie approach. Reliab Eng Syst Safety, Osa/vuosikerta 104:36–44

Gefördert durch:

Bundesministerium
für Wirtschaft
und Energie

aufgrund eines Beschlusses
des Deutschen Bundestages

Climate Change and Coastal Transport Infrastructure—How Do We Keep Australia Moving?

Greg Fisk, Fahim Tonmoy and David Rissik

Abstract Transport infrastructure across the spectrum of airports, seaports, road and rail involves assets that are long-lived, and what is designed today must be done so in the context of expected increases in the intensity of extreme weather events. Much of Australia's transport infrastructure is located close to the coast and is vulnerable to sea level rise and its associated processes (e.g. erosion, inundation), and other climate change-related extremes storms, heatwaves, droughts and floods. Faced with the uncertainties of the timing and severity of climate change, decisions about what, where and how to build new coastal transport infrastructure as well as maintaining existing ones will become more and more challenging in the future. This paper summarises the risks to coastal infrastructure from climate change and the key drivers that owners and operators of transport infrastructure in Australia should consider to help them adapt to the effects of coastal climate change and extreme weather events. This includes both the siting and design of new infrastructure as well as strategies to build resilience of current infrastructure to future impact. Showcasing the National Climate Change Adaptation Research Facility's (NCCARF) on-line adaptation decision support tool, CoastAdapt, the paper outlines guidelines and information available to infrastructure owners and operators to build resilience and adapt to future climate risks including a recent case study undertaken with North Queensland Airports.

G. Fisk (✉)
Market Lead (Environment) BMT WBM Pty Ltd., Brisbane, Australia
e-mail: greg.fisk@bmtwbm.com.au

F. Tonmoy · D. Rissik
National Climate Change Adaptation Research Facility (NCCARF), Southport, Australia
e-mail: f.tonmoy@griffith.edu.au

D. Rissik
e-mail: d.rissik@griffith.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_17

1 Introduction

As the growing risk of climate change become more evident, nearly 200 Governments agreed in 2015 as part of the Paris Agreement to pursue efforts to further reduce emissions and halt global temperature increases [8]. However, even if current carbon mitigation targets are achieved, this will not avoid major global impacts and as such there is an urgent need for strong adaptation and mitigation strategies [6].

Coastal transport infrastructure such as airports, seaports, road and rail play a significant role in the national economy and community through passenger movement, imports and exports of goods. A significant proportion of Australia's transport infrastructure is located close to the coast and vulnerable to climate change, sea level rise and its associated processes (e.g. increased intensity of extreme weather events, erosion and inundation). The design life of transport infrastructure is long making it critical that it is designed with consideration of expected future increases in the intensity of extreme weather events. Faced with the uncertainties of the timing and severity of climate change, decisions about what, where and how to build new coastal transport infrastructure as well as maintaining existing assets will become more and more challenging in the future. A number of guidelines are available to assist the transport sector to understand climate risks and adaptation (e.g. Ng Ak et al. [5]). However, there are few guidelines available which apply across coastal infrastructure types and are specific to the Australian context. The Australian Government commissioned the National Climate Change Adaptation Research Facility (NCCARF) to develop a national coastal climate change risk management tool called CoastAdapt which includes a range of national climate change and sea level rise projections, guidelines on risk assessment and adaptation planning. Although not specifically designed for the transport sector, CoastAdapt's information can assist decision makers of this sector to manage their future climate change risks. This was demonstrated by a recent project with Northern Queensland Airports where CoastAdapt was used to investigate climate change risks of Mackay and Cairns airport (see Fisk [3]). This paper highlights how some of the CoastAdapt information can be used by the coastal transport sector in Australia in managing their climate change risks.

2 Climate Risks for Transport Infrastructure

The key coastal climate risk parameters that are considered most relevant to transport infrastructure assets and operations (including workforces) are outlined in Table 1.

Table 1 Key climate change parameters of concern for transport infrastructure

Parameter	Effects
An increase in the severity of cyclones, hurricanes or extreme storm events	That could result in: • Increased high wind events • Increased storm surge (tide) events • Increased major flooding events in the form of either: – overland flooding (rainfall leading to flooding from the catchment or watershed that flows into the port) or – flooding from the sea (storm tide inundation)
Changes in rainfall patterns	That could result in: • Heavier rainfall and fog events (causing impacts to visibility and safe maritime navigation) • More frequent flooding and/or flooding at different times of the year to current conditions • Changes to water supply and availability
Increasing temperatures	That could result in: • Increased incidents of very hot days and heatwaves • Increased evaporation and fire risk • Increases in water temperature
Sea level rise	That as a result of higher mean sea level could result in: • More frequent erosion events • More frequent and far-reaching tidal inundation associated with storm surge and storm tide events • Permanent inundation of coastal areas • Exacerbate the effect of cyclones and extreme storm events listed above.

3 What Are the Major Drivers for Taking Action?

High-level vulnerability mapping assessments of future climate change including sea level rise that have been undertaken at the Australia-wide level [1] and by the various State and Territory Governments have demonstrated that transport infrastructure owners and operators have powerful reasons to begin planning and adapting now to climate risks given their inherent vulnerability. Other drivers for consideration of climate change come from sources that are external to organisations. In particular, recent recommendations made by the Taskforce on Climate-Related Financial Disclosures in December 2016 reported on the need for organisations to include climate-related risks in their mainstream financial filings. The advice and recommendations apply to large asset owners and operators in particular, as they sit on the top of the investment chain. While still in its early stages, the recommendations of the Taskforce place an increasing importance on understanding climate risk and disclosing mitigation and adaptation actions to minimise this risk [8]. Similar guidance released from Australian Prudential

Regulation Authority (APRA) warned that climate change risks can threaten the entire financial system and advised investors to consider climate change risks in their decisions [7]. These reports make it clear that investors, owners and operators in the transport sector will need to start to better understand climate change risks and plan for action.

4 Climate Change Risk Management of Coastal Transport Infrastructure Using CoastAdapt

CoastAdapt has developed the Coastal Climate Adaptation Decision Support (C-CADS) framework which is based on ISO31000 risk management framework but tailored for climate change risks (Fig. 1). C-CADS provides general guidelines for climate change adaptation planning which can be used along with CoastAdapt datasets for developing resilience and adaptation pathways of coastal organizations (e.g. local councils, businesses). The following sections complement the C-CADS framework and set out key issues that need to be considered by the coastal transport sector which would allow them to better use this national resource CoastAdapt.

4.1 Assess Risks and Vulnerabilities

In a general sense and as outlined in DIICCSRTE [2], transport infrastructure can be exposed to significant climate risks through:

Fig. 1 Coastal Climate Adaptation Decision Support (C-CADS). *Source* www.coastadapt.com.au

- **risks to assets**. Changing climate may damage major assets including roads, rail, airport facilities and port facilities. Support structures and infrastructure (yards, workshops, etc.) and energy infrastructure and communications facilities will also be potentially affected. Damage to assets can force owners to retire assets early or make major upgrades following damage.
- **risks to operations**. More frequent or intense extreme weather events may disrupt business operations or else the usability or reliability of services. Examples include the effect of floods on rail and road transport and electricity supply, disruption of ports and airports by cyclones and major storms, and reduced productivity of outdoor workers due to high temperatures.
- **critical dependencies**. Extreme weather events may interrupt supply chains or services such as transport, cargo, electricity, gas or water supply on which businesses depend. A changing climate may also affect global trading patterns, for example, by changing the supply of agricultural products or mining products.
- **national economy**. Climate change impacts are a potential drag on the national economy, with a flow-on effect to individual businesses.
- **insurance and capital markets**. Climate change shocks may affect the availability of insurance and access to capital, either locally or worldwide.

From a business perspective, future climate change has the potential to affect the valuation of infrastructure assets by impacting cash flows, operational cost and capital expenditure. The risk posed by climate change may not necessarily be to the asset itself but to the goals and objectives that may be compromised if the asset is impacted or damaged. Thus the primary risks that need to be considered are those to the reliability and performance of the assets and whether relevant targets and performance metrics can continue to be met.

To begin to assess these risks, CoastAdapt advocates a 3-phase risk assessment methodology and provides guidance on three potential risk assessment processes that can be undertaken by an asset manager of a transport authority: (i) a first pass risk screening can be used with existing national datasets to identify most at risk assets and relevant stakeholders; (ii) a second pass risk assessment can be done by having a risk workshop with relevant stakeholders to investigate potential consequences of climate change risks and identify the most critical ones; and (iii) a third pass (detailed) assessment which can be conducted to gain better information for any critical at-risk sector through site-specific biophysical modelling. Key information and data inputs for a climate risk assessment process will generally need to include the following contextual information:

- Picking a timeframe for the assessment (for example—current day, 2050 and 2100), the climate change emissions scenario to be used (most likely, best case, worst case), the spatial study area of the assessment and scope of the assessment.
- Knowledge of climate risks and trends for the local and regional area including expected changes to sea level rise, rainfall patterns, temperature and sea state conditions.

- Knowledge of asset registers and the condition of assets at the study area, the relative value and significance of assets and where available, more detailed vulnerability or hazard mapping.

In order to assist risk assessments, CoastAdapt provides access to sea level rise projections, temperature and rainfall extreme projections for all coastal councils in Australia which can be used by the transport sector to start understanding their future risks (a first-pass or second-pass assessment). CoastAdapt also provides present-day geomorphic information (*Smartline* showing erodibility of the coastline and *Sediment compartment* providing high-level information on sediment movement within a secondary compartment) and historical flood information (*Water Observation from Space* showing the historical presence of water analysing 27 years of Landsat data). These data can assist transport authorities to understand the likelihood of impact from climate change in a risk assessment processes.

Consequence is the other half of the risk equation—that is assessing the consequence of climate hazards on asset condition, life and function or on the operations of the transport authority. Assessing consequence can be far more difficult and subjective to assess without a detailed understanding of the asset and/or operation being impacted and should generally involve content specialists such as asset managers and maintenance staff.

The concept of the adaptive capacity of assets and operations is also important for evaluating risk consequence. Some key considerations here include:

- The design life and resilience of the asset to impacts (particularly if it is not a permanent asset or structure);
- Existing emergency/evacuation management procedures and protocols that can mitigate the risk;
- The degree to which the asset can be reconfigured or redesigned to accommodate changes in climate or extreme weather events;
- Technological changes including the ability to work longer/function during periods of more challenging conditions.

4.2 Identify Options for Risk Treatment

In responding to risks identified, there will be a range of planning and management options available. Focussing on assets, adaptation for transport infrastructure will need to be split between existing assets and new (proposed) assets—with different adaptation tools available to manage current and future risks. Figure 2 shows an example decision tree for assessing options for resilience and adaptation for transport infrastructure, separated between existing assets and new assets.

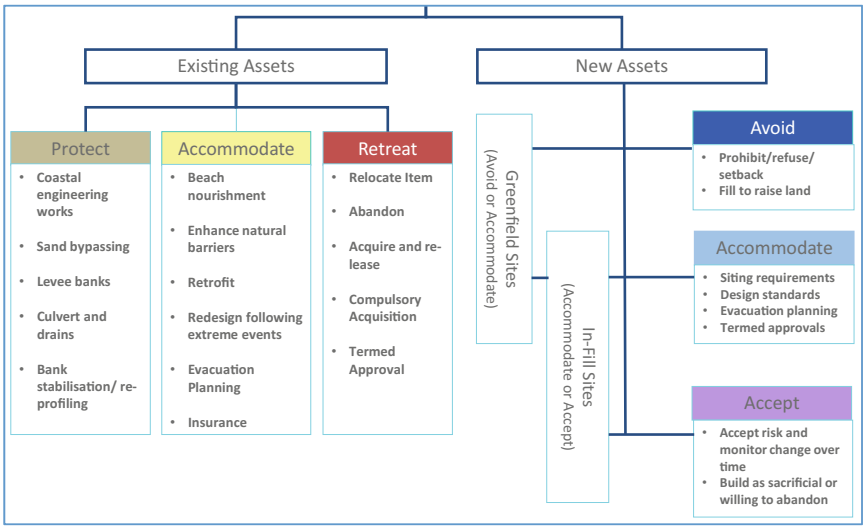


Fig. 2 Decision tree for application of resilience and adaptation options for transport infrastructure. *Source* Authors

4.3 Taking Action and Adaptive Pathways

Implementation of identified options can be done in stages as outlined in CoastAdapt as a pathways approach, originally proposed by Kwadijk et al. [4]. It can allow transport authorities to minimize their current expenditure on managing climate change risks and keep options open. Under the approach, rather than determining a final outcome or decision at an early stage, decision-makers are able to build a strategy that will follow changing circumstances over time. The timing of adaptation activities can also be very important. For example, it makes sense to delay expensive engineering retrofits until threats are more immediate. Alternatively, the implications of sea level rise need to start to be considered now as new port infrastructure will likely be expected to have a design life into the latter part of the 21st century and beyond.

4.4 Case Study on Climate Risk Assessment—NQA

This section outlines a short case study on how a transport infrastructure entity (North Queensland Airports) is starting to understand and consider climate change risks for its assets and operations. This case study was undertaken using the tools and guidance contained within CoastAdapt and further information about this and other case studies are available from https://coastadapt.com.au (Figs. 3 and 4).

Fig. 3 Aerial image of Cairns Airport. *Source* © Google Earth

Fig. 4 Aerial image of Mackay Airport. *Source* © Google Earth

Case Study—Climate Risk Assessment for North Queensland Airports

North Queensland Airports (NQA) operates the Cairns and Mackay Airports on leased state land. The airports are situated on the tropical North Queensland coast. Both have been built on low elevation coastal land (reclaimed mangrove ecosystems), are situated in cyclonic regions, experience high temperatures during summer and have climate-sensitive assets and operations. In particular, intense cyclones and associated storm surges and flooding can damage airport infrastructure as well as impact operations by causing temporary closures. Under future climate change and sea level rise, these impacts are likely to intensify.

To better understand climate risks to the airports both now and in the future, NQA undertook an internal risk screening and risk assessment process using the mapping tools and guidelines published in CoastAdapt. The risk screening and assessment process looked at present-day and future risks (at 2030 and 2070).

A broad range of climate change risks were selected for assessment, including:

- increasing average temperatures, including more extremely hot days (greater than 35 °C), and increasing evaporation/drought
- average wind speed increase (noting that higher winds affect aircraft operations and ground handling procedures)
- coastal erosion
- increasing frequency and severity of storm surges and storm tides
- more intense rainfall leading to flash flooding
- more intense storms leading to larger riverine and overland flooding
- higher groundwater table during rainfall peaks (affecting pavements)
- more frequent and more intense bushfires (smoke events)
- more intense lightning storms and events, and fog events
- cumulative impacts from adjacent development (that address or exacerbate climate change impact).

Risks were identified and assessed using the first pass and second pass guidance documentation provided in CoastAdapt (NCCARF 2016). Through a series of workshops (facilitated by the authors), NQA staff assigned risks on the basis of how they could potentially affect each airport's built assets (i.e. runways, taxiways, buildings, aviation precincts) and/or airport operations (i.e. efficiency, productivity and safety). Current climate risks at the airports were all assessed as 'Low' on the basis of the range of risk mitigation and treatment measures already being implemented. Given their existing exposure to coastal hazards, the assessment identified the specific engineering solutions already in place at the airports such as the storm tide levee and pumping system at Cairns Airport and the flood detention drainage system at Mackay Airport.

Future risks from storm surge and tidal inundation exacerbated by higher sea levels were identified as relevant at both airports in the longer term (by 2070). Other future risks include increased flash flooding (from more intense rainfall events) and hence localised erosion. The risk assessment and associated information will be used by NQA to inform their long-term planning.

5 Conclusions

With powerful reasons and drivers to build resilience and adapt because of their exposure and limited adaptive capacity, transport infrastructure on the coast will benefit from an adaptive, pathways approach to respond to future climate change. In addition to the existing suite of guidance available from CoastAdapt, specialist guidelines being funded by NCCARF will assist transport authorities and organisation to navigate through an appropriate risk assessment process and to develop appropriate treatment options and triggers.

References

1. Department of Climate Change and Energy Efficiency—DCCEE (2011) Climate Change Risks to Coastal Buildings and Infrastructure A Supplement to the First Pass National Assessment. Canberra
2. Department of Industry, Innovation, Climate Change, Science, Research and Tertiary Education [DIICCSRTE] (2013) Climate Adaptation Outlook: A Proposed National Adaptation Assessment Framework. Canberra
3. Fisk G (2017) Climate risk assessment for North Queensland Airports. Snapshot for CoastAdapt, National Climate Change Adaptation Research Facility, Gold Coast
4. Kwadijk JCJ et al (2010) Using adaptation tipping points to prepare for climate change and sea level rise: a case study in the Netherlands. Wiley Interdisc Rev: Clim Change 1(5):729–740
5. Ng AK, Becker A, Cahoon S, Shu-Ling Chen, Earl P, Yank Z (eds) (2016) Climate change and adaptation planning for ports. Routledge Publishing, New York
6. Parry M, Palutikof J, Hanson C, Lowe J (2008) Squaring up to reality. Nat Rep Clim Change 2:1–3
7. Summerhayes G (2017) Australia's new horizon: climate change challenges and prudential risk. Australian Prudential Regulation Authority (APRA)
8. Taskforce on Climate-Related Financial Disclosures (TCFD) (2016) Recommendations of taskforce on climate-related financial disclosures

Ultrasonic Phased Array on Time-of-Flight Diffraction for Non-destructive Testing via Numerical Modelling

Tat-Hean Gan, Channa Nageswaran and Mario Kostan

Abstract An inspection technique for steel pipes that utilises phased array (PA) probes in time-of-flight diffraction (TOFD) configuration to continuously monitor the defect growth over time is being developed, so that when the defect reaches a critical size the power plant can be shut down and maintenance can take place before failure. The numerical models for PA/TOFD in either symmetric or asymmetric pitch-catch configuration were developed using CIVA simulation platform developed specifically for non-destructive testing applications. The probe characteristics were selected and the beam profile was predicted for different points in the weld and heat affected zone (HAZ). The probes positions and interspacing in emission and reception were also selected in order to achieve max inspection coverage. The PA probes use piezoelectric elements for generation and reception of the ultrasound beam. Single crystal gallium orthophosphate ($GaPO_4$) was selected for impedance analysis as a candidate for application in the PA probe. Impedance characteristics of $GAPO_4$ elements were investigated up to 580 °C and then used to derive the material properties of these elements up to the temperature stated above. Finally, the calculated properties were used to evaluate the developed PA/TOFD technique at high temperatures (HT) using COMSOL simulation package.

T.-H. Gan (✉) · M. Kostan
Brunel University London, Kingston Lane, UB8 3PH Uxbridge, Middlesex, UK
e-mail: tat-hean.gan@brunel.ac.uk

M. Kostan
e-mail: mario.kostan@brunel.ac.uk

C. Nageswaran
TWI Ltd., CB21 6AL Granta Park, Great Abington, Cambridge, UK
e-mail: channa.nageswaran@twi.co.uk

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_18

1 Introduction

A number of power plants are being granted extensions, as they are reaching the end of their expected service life. Hence, the maintenance of these ageing plants is very important and the operators have to carry out more regular outages to ensure their safe operation.

Outages involve the shutdown, erection of scaffolding and removal of insulation to gain access which is a significant part of the inspection cost. Since the inspection is performed during an outage, the techniques used are effective at ambient temperature only. However, if a vital part needs to be monitored in service following an outage, such as pipes carrying superheated steam, it may present a problem, e.g. HT and pressures experienced in those pipes, can lead to creep, fatigue and corrosion type defects, which if undetected may have catastrophic consequences. Hence, structural health monitoring (SHM) techniques need to be developed to retain reliability and extend the lifetime of aging plants.

Combined ultrasonic techniques, such as PA/TOFD, can be used to detect and monitor defects at ambient temperature. However, currently, there are no such combined ultrasonic systems that can allow operation at up to the temperature level required in this work. Commercially offered HT ultrasonic systems, together with their temperature limitations, are listed in Table 1 (Sonotech, Olympus).

The wider context is to develop PA/TOFD ultrasonic probes for continuous monitoring of defect growth over time up to 580 °C [3]. Figure 1 depicts the new SHM system, which consists of the following components: (i) HT probes, (ii) PA pulser-receiver unit and (iii) signal processing and visualisation software. The key novelty presented here is the impedance analysis method of the gallium orthophosphate piezoelectric crystal that potentially can be used to build the probes for HT operation.

The PA/TOFD probes use piezoelectric elements for generation and reception of ultrasound beam needed for inspection. Lead Zirconate Titanate (PZT) is the most commonly used piezoelectric for PA probes today but has a max operation at ~180 °C which is not suitable for this HT application [4]. Hence, alternative piezo-materials need to be considered. A number of piezoelectrics that can withstand HT have been reported [1]. Single crystals such as gallium orthophosphate ($GaPO_4$), langasite (LGS) and aluminium nitride (AlN) stand out because they exhibit no Curie temperature and no domain-related aging behaviour while showing good sensitivity and the ability to function over a broad temperature range [8].

Table 1 Temp. limitations of the commercially offered HT ultrasonic systems

Technique	Application	Temp. limitation (°C)
TOFD	Defect detection	485
Phased array	Defect monitoring	400
Corrosion mapping	Corrosion assessment	350
Thickness gauging	Spot checking	500

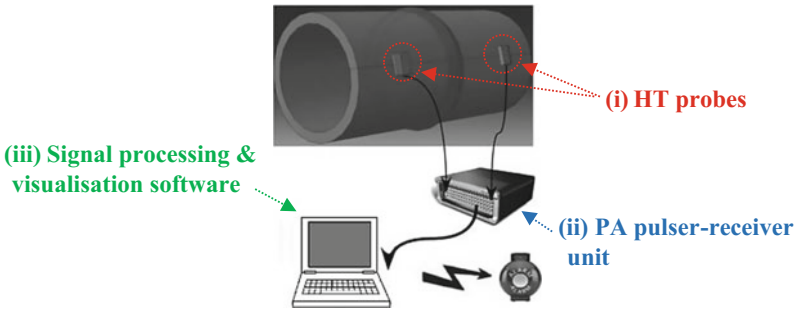


Fig. 1 New HT SHM system under development

Due to the thermal stability of most of its key properties up to 970 °C (≫580 °C) as well as commercial availability, $GaPO_4$ has been selected for further study as a candidate for application in the HT PA/TOFD probes under development [5].

The PA/TOFD probes should function continuously in between two outages. However, HT properties of $GaPO_4$ were not known. The impedance method was used to determine the properties such as thickness coupling factor or charge constant. These properties can be derived from the measured values of the resonant and anti-resonant frequencies, capacitance, density and dimensions of appropriately cut and excited $GaPO_4$ elements, as outlined in the Standard on Piezoelectricity [2]. From the derived properties, it is possible to evaluate the efficiency and sensitivity of PA/TOFD probes that would use these elements for inspection at HT.

The typical defect is a crack, especially in the weld region. The SHM system under study aims to detect the creep damage at stage 4 when the cavities are joined into microcracks. The new system will monitor the growth of the crack before failure. The cracks at this stage require the use of an ultrasonic technique that allows detection of diffraction signals from the tips of the crack. The detection of tip diffraction is also used for height sizing. This is the reason why PA/TOFD was selected to be developed.

2 General Principle for PA/TOFD

PA/TOFD is the combination of the pitch-catch technique with the generation of a range of angle ultrasonic beams. The transmission and the reception of the sound are separated by two probes positioned on either side of the weld. Figure 2a shows how the two probes are set-up for TOFD examination of a weld.

It has been reported that focussed PA probes were capable of detecting diffraction signals from crack-like defects, enabling detection and through wall sizing [6]. Therefore the PA technology combined with TOFD principle allows coverage of a large weld volume and the HAZ with a single probe set and reduces

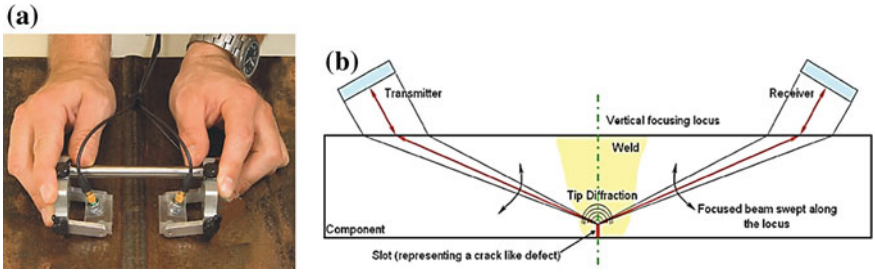


Fig. 2 **a** TOFD examination of a weld and **b** PA/TOFD concept where one linear PA probe operates as a transmitter focusing the sound and the other as a focused receiver

the need to use several probe sets and mechanical scanning. This presents the advantage of simplifying inspection implementation. Figure 2b shows the PA/TOFD concept.

A number of points in the region of interest (RoI) were selected to measure. Due to the positions of the points regarding the PA transmitter and receiver, two configurations for PA/TOFD were investigated: (i) symmetric PA/TOFD where the delay law at emission is the same as reception and (ii) asymmetric PA/TOFD where the delay laws for the emission and reception are different.

The beam profile was modelled in order to verify whether the characteristics of the beam were suitable for inspection of the weld volume and HAZ. The modelling was carried out considering the case where the beam was steered from 35° to 75° and the case where the beam was focussed at different points in the weld volume and HAZ, Fig. 3a. For each model, the beam profile and the amplitude of the beam were recorded at the max point and the point of interest of the weld.

Figure 3a shows the points where the beam was focussed. At each point the beam in emission and in reception meet. Figure 3b shows the amplitude of the beam strength taken at each focus point. On the chart the amplitudes were normalised to the max value recorded for all focus points. The amplitudes are displayed in dB. The max amplitude for all focus points was recorded at X = 5 mm and Z = 20.25 mm with max amplitude of 0.82 in CIVA absolute value. The ultrasonic beam is 5.4 dB stronger than the non-focussed beam steered at 67°. Moreover, it can be noted that the difference between the max beam amplitude at X = 5 mm and Z = 20.25 mm and the point of min amplitude X = −10 mm and Z = 9 mm is 4.6 dB. This difference can be easily compensated for by suitable calibration applied during the data collection. These data allow concluding that the PA probe can generate an efficient beam when focussed over the full volume of the weld and HAZ.

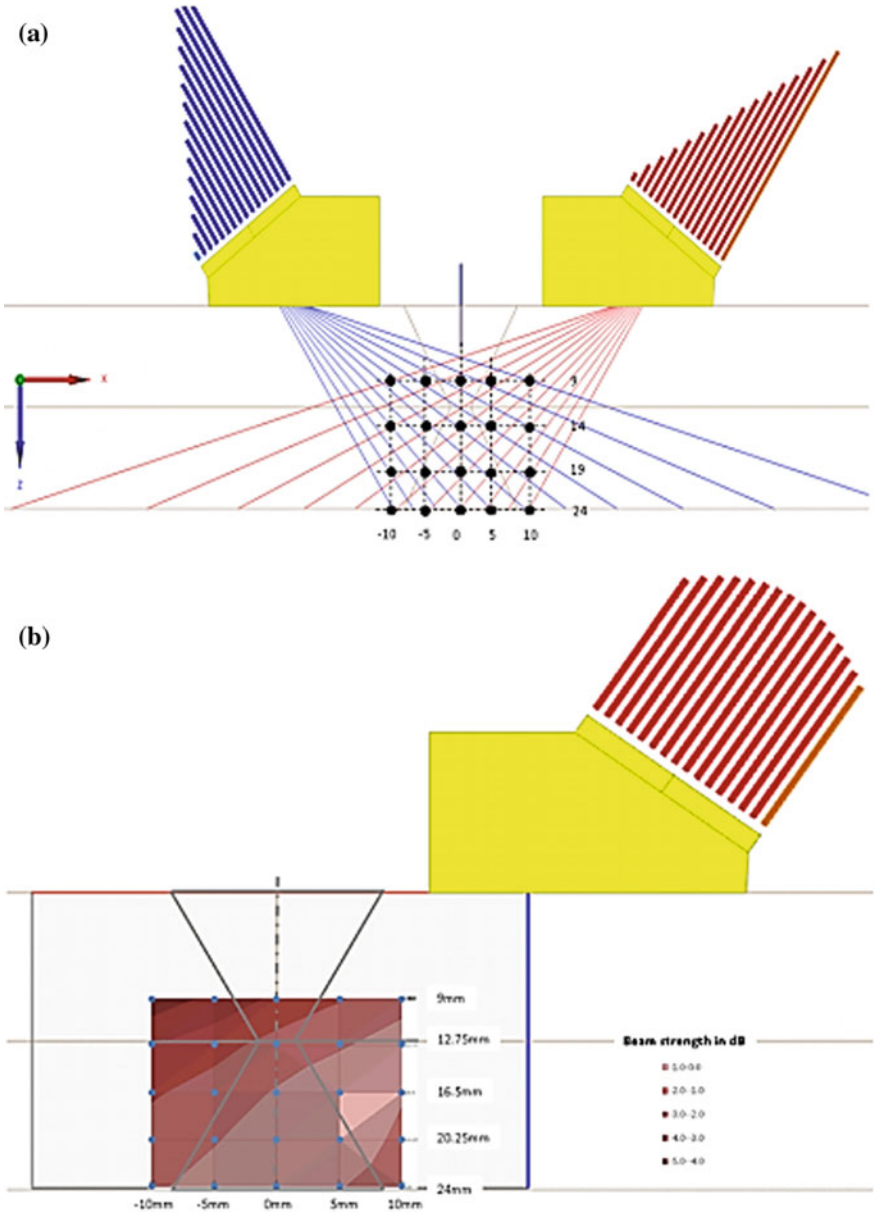


Fig. 3 **a** Focussing points in the weld and HAZ and **b** ultrasonic beam strength in emission at the point of focussing identified to cover the full volume of the weld and HAZ

3 Development of HT PA Probes in CIVA

There are a variety of available PA probes but mainly limited to operation at ambient temperature. There are some PA probes for operation up to 200 °C (max contact time of 10 s), which use an HT wedge [7]. PA probes that can work up to 580 °C currently do not exist. Hence, such PA probes need to be developed by selecting suitable piezo-material and other components to operate at 580 °C.

The key issues for designing such PA probes are: (1) HT piezo-material/elements, (2) HT coupling, (3) HT wedge, (4) HT backing, (5) HT matrix, (6) HT housing and (7) HT wiring. Among these, the performance of the HT piezoelectric material/elements is the most important and the rest of issues can be solved through selecting commercial products from the market. In this paper, the HT performance of the selected $GaPO_4$ material is investigated experimentally using impedance method. Once developed, the PA probes will be subjected to prolonged HT tests to validate their thermal stability. This will be done in the lab, prior to field trials.

3.1 Probe Settings

In order to demonstrate the HT PA/TOFD system concept, it has been decided to develop a 5 MHz linear PA probe with an array count of 16 elements. To determine if such a probe will be able to produce a satisfactory ultrasonic result, a representative probe was designed and modelled in CIVA. The key indications of good ultrasonic performance that were considered were fine beam spot with weak side lobes. In the model, the optimum positioning in relation to the weld and HAZ was also considered, as well as the size of the probe components such as wedge.

The gap between the elements and the element width are two of the key parameters to consider in terms of PA probe design. These were studied in CIVA to identify the gap and element width that would achieve the best ultrasonic characteristics. Considering the ultrasonic characteristic generated in the weld volume and HAZ, the PA probe selected was as shown in Table 2:

Table 2 Selected PA probe's parameters

Type	Count	Width	Gap	Pitch	Frequency
Linear PA	16	0.9 mm	0.1 mm	1 mm	5 MHz

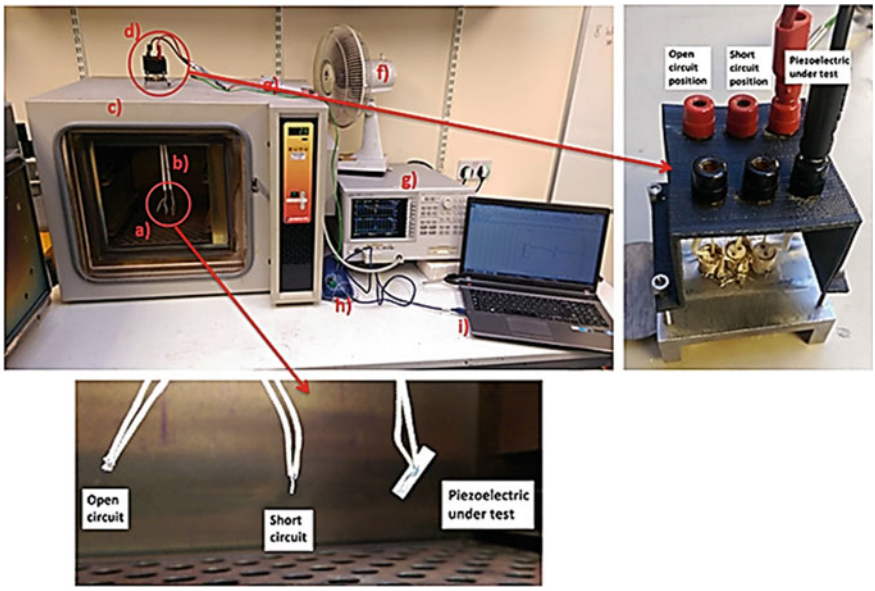


Fig. 4 Experimental setup used for impedance analysis of $GaPO_4$ elements at HT

4 Experimental Setup and Procedure

Five $GaPO_4$ elements were used for impedance analysis. The elements had a velocity of 4356 m/s and a resonant frequency of 2.17 MHz. The dimensions and density at ambient temperature were determined using a calliper, and for HT they were calculated using coefficients of thermal expansion (CTEs) provided in (Piezocryst).

In Fig. 4, one can see the experimental setup used for impedance analysis at HT. The setup consisted of: (a) $GaPO_4$ element under test, (b) insulation tubing, (c) furnace, (d) panel with plug sockets for open and short circuit compensation, (e) BNC leads, (f) fen used for cooling of the element holder, (g) impedance analyser, (h) temperature logger with K type thermocouple and (i) PC to collect the data.

5 Results

5.1 Impedance Characteristics

The impedance characteristics of five $GaPO_4$ elements were taken from ambient temperature up to 580 °C, with an increment of 50 °C. The hold time of 10 min at each temperature was applied to ascertain an isothermal measurement. The typical

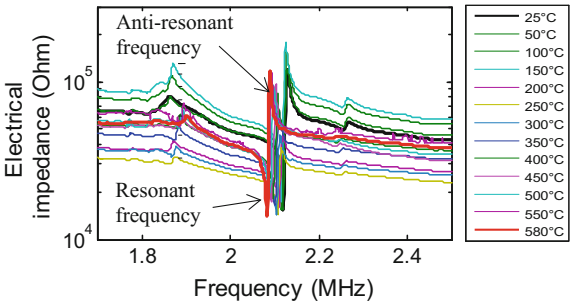


Fig. 5 Impedance characteristics of a $GaPO_4$ element from 25 °C up to 580 °C

characteristics of a $GaPO_4$ element, plotted to show the effect of increasing temperature on the resonant and anti-resonant peaks, can be seen in Fig. 5.

5.2 *Derived Material Properties*

The values of the resonant and anti-resonant frequencies were used to derive a number of material properties, and two of these, piezoelectric charge constant d_{11} and thickness coupling factor k_t, can be seen plotted against temperature in Fig. 6a, b, respectively. The derived material properties, averaged for five $GaPO_4$ elements, stayed stable all the way up to 580 °C.

6 PA Probe Simulation in COMSOL

The model for PA/TOFD transmission and reception up to 580 °C was studied in COMSOL. As shown in Fig. 7a, the model of PA transmitter T_X and receiver R_X on a steel sample was in the symmetric configuration. The probe settings were the same as in 3.1. To have a good signal, the focus was at the back wall and the time delay for each element of T_X at the focus is in Table 3. Due to symmetric position of T_X and R_X, the receiving time delay was the same as the transmitting one. An example of the beam focused at the back wall at 580 °C is in Fig. 7b.

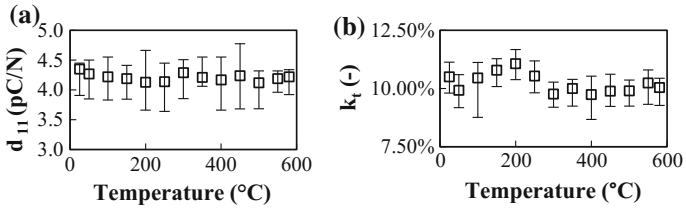


Fig. 6 **a** Piezoelectric charge constant d_{11} and **b** thickness coupling factor k_t

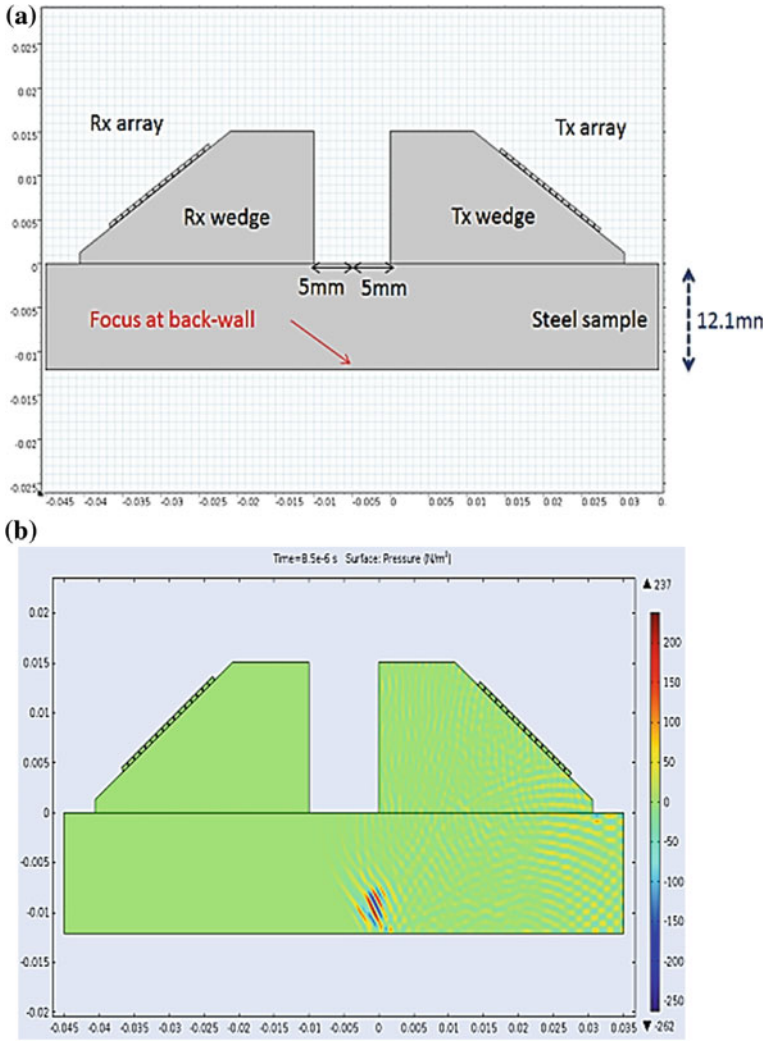


Fig. 7 **a** Model for a pair of PA probes for transmission and reception and **b** pressure distribution when the beam propagates towards the focusing point at 580 °C

After implementing the time delay on the receiver PA, an average of 16 signals received at R_X is calculated as the overall received signal. Four different temperatures 20, 200, 400 and 580 °C, are simulated by using the derived properties from 5.2. The averaged received signals at *Rx* at four temperatures are shown in Fig. 8. A reflection with large amplitude is detected at ~13.4 μs for each temperature, which is the reflection from the back wall of the steel sample. This simulation showed that the acoustic performance of the PA from $GaPO_4$ was temperature-independent, which matched the experimental results discussed in 5.1.

Table 3 Time delays for each of the 16 elements of PA, when the focus is at back-wall

Element	1	2	3	4
Delay (μs)	0	0.073	0.142	0.208
Element	5	6	7	8
Delay (μs)	0.270	0.328	0.383	0.432
Element	9	10	11	12
Delay (μs)	0.478	0.519	0.555	0.586
Element	13	14	15	16
Delay (μs)	0.613	0.634	0.650	0.660

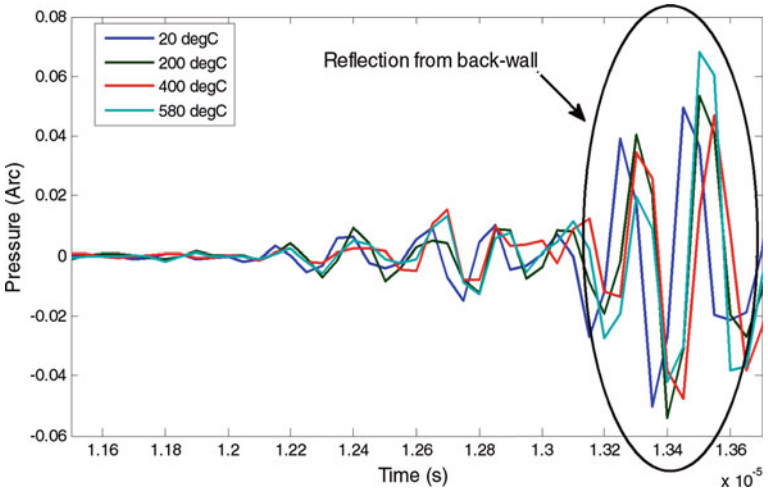


Fig. 8 Averaged signals received at R_X PA at 20, 200, 400 and 580 °C

7 Conclusion

The combined PA/TOFD inspection technique was developed using CIVA. The beam profile and amplitude for different points in weld and HAZ were simulated. This resulted in optimal probe settings of a linear PA with 16 elements, pitch 1 mm and frequency 5 MHz. To allow inspection up to 580 °C, the PA will be designed from a HT piezoelectric $GaPO_4$. The material properties of $GaPO_4$ were measured using impedance method which confirmed it to be convenient for the probe design. Finally, the derived properties were used to simulate the ultrasonic response of the 16-element PA and this was done using COMSOL. Stable amplitudes of the simulated echoes confirmed good acoustic performance of the PA up to 580 °C, and the next step will be to build the PA probe using the same piezoelectric material and to test it in the lab.

References

1. Damjanovic D (1998) Materials for high-temperature piezoelectric transducers. Curr Opin Solid State Mater. Sci 3:469–473
2. European Standard (2002) Piezoelectric properties of ceramic materials and components: Methods of measurement-Low power. EN 50324-2:2002
3. FP7 Project (2013) HotPhasedArray. [Online] Available at: http://www.hotphasedarray.eu/project
4. Hooker MW (1998) Properties of PZT-based piezoelectric ceramics between −150 and 250 °C. NASA/CR
5. Krempl P (1994) Quartz homeotypic gallium-orthophosphate-a new high tech piezoelectric material. UT Symp. IEEE Proc, pp 949–954
6. Nageswaran C, Bird C (2008) Evaluation of the phased array transmit-receive longitudinal and time-of-flight diffraction techniques for inspection of a dissimilar weld. Insight 50(12):1–7
7. SIUI Ltd. (2015) Phased array probes and wedges. [Online] Available at: file:///C:/Users/PoolUser/Downloads/Phased_Array_Probe_5.pdf
8. Zhang S, Yu F (2011) Piezoelectric materials for high temperature sensors. J Am Ceram Soc 94:3153–3170

An Approach to Quantify Value Provided by an Engineered Asset According to the ISO 5500x Series of Standards

Vicente González-Prida, Antonio Guillén, Juan Gómez, Adolfo Crespo and Antonio de la Fuente

Abstract The purpose of any asset is to provide value to the organization and its stakeholders. In Asset Management, the concept of value encompasses quantitative and qualitative, as well as tangible and intangible benefits that assets may provide to an organization. The definitions of asset and value are not only closely linked but also complementary. An "asset" provides the means for the realisation of "value" thus the management of an asset is strategic and has to be linked to an organization's value norms. This paper extrapolates from the definitions in ISO 5500x series of standards to describe a generic approach for quantifying the value provided by engineered assets deployed by a business organisation.

1 Introduction

The purpose of any asset is to provide value to the organization and its stakeholders [1]. In Asset Management (AM), the concept of value considers all those aspects (obtained or expected to be obtained from an asset) that provide any kind of benefit (expressed in specific terms) to the organization [2]. The definition of both, asset and

V. González-Prida (✉) · A. Guillén · J. Gómez · A. Crespo · A. de la Fuente
Department of Industrial Management, University of Seville, Seville, Spain
e-mail: vicente.gonzalezprida@gdels.com

A. Guillén
e-mail: ajguillen@us.es

J. Gómez
e-mail: juan.gomez@iies.es

A. Crespo
e-mail: adolfo@etsi.us.es

A. de la Fuente
e-mail: antoniodela84@gmail.com

V. González-Prida
UNED, Madrid, Spain

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_19

value, are complementary to each other. Indeed, asset and value are two concepts closely involved where there is no sense one without the other. In few words, the "asset" is the tool or the instrument for materializing the "value" expected by the company, and this fact encompasses the company strategic objectives. Each asset must be defined and managed based on its link to the concept of the organization's own value [2]. The definition of value for each company/organization, must serve to determine and describe, accurately, the asset portfolio of each organization, answering the questions of what those assets are and how they contribute to precisely achieve or conserve this value [3].

Asset management should be founded on the value control. But many times the value concept is not actively integrated in the management tools, and only financial/economic variables—only limited to cost most cases—are used for the control and decision-making. The idea of value is present in these processes, of course, but not in a formal way, neither including objective evaluation methods. The challenge here is to design a methodology that allows the quantification of such a value and its effective use in asset management. Even ISO 55000, where the capital role of value has definitively been put on the table, does not address the value measurement. The effective control of value, which needs the introduction of value assessment methods, can promote great benefits for the overall organization management [3] throughout this document, a generic approach to quantify assets value is presented, according and aligned to the concept defined in the standards ISO 55000 [1, 4, 5].

In this context, a deep revision about the current use of value as key concept in AM is needed including, or starting from, a philosophical discussion about the definition of value concept. But this goes far beyond the scope and the extension intended by this paper. Nevertheless, in order to promote the discussion some references have been included here that can be consulted as a starting point: Cronin et al. [6] includes a review of value, quality and satisfaction concepts in services marketing literature; Harrison et al. [7] offer the stakeholder perspective on value, draw attention to those factors that are most closely associated with building more value for stakeholders; Heitz et al. [8] present a formal model for decision making in asset management, based on the general concepts of ISO 55000, in particular the notion of value realization; Hofmann and Locker [10] present a use case of value-based performance measurement concept in supply chains, Campbell et al. [9] present a general framework for Asset Management and maintenance where interesting value-risk references can be found; Kaplan and Norton [10] with the Balance Score Card they developed, the fact, a practical method to aboard the inclusion of approaches different from economical aspect in management of any kind of organization, and in this sense, this is an evident precursor of value approach; Allee [11] gives a good approach to value analysis for intangible assets management. Among these references, other fields appear too in addition to engineering and AM as economics market, customer studies, marketing and intellectual capital management etc. All these areas present more mature visions about value modelling and its utility for management. These fields treat this issue from different perspectives that, somehow, can help the engineer and asset manager to build their own vision that contributes to the evolution of AM.

The paper is organized as follow: Sect. 2 introduces the approach to the value concept in an organization that summarizes the vision of this paper to the value definition. This section includes two subsections: in Sect. 2.1 the difference between a negative assessment of value, or "loss of value", and positive assessment of value, or "gain of value" is discussed; Sect. 2.2 gives examples to understand the value assessments approaches. Finally, conclusions are summarized in Sect. 3.

2 The Value Concept in an Organization

Since the publication of ISO 5500x, the term Value has become the buzz word in Asset Management [12]. However, in the literature review cannot be found yet many contributions dealing specifically with the value concept for an organization. In general terms, value is intuitively easy to understand. However, its objective definition is not easy to delimit, moreover, when the goal is to establish methods of measurement that allow the decision-making to be made in terms of value [13]. Basically, each organization has to define its own value concept. In other words, the organization value concept (OCV) as a specific element must include and describe all the relevant components of the overall value view of the organization. A property of value is that it is created during the operation of the asset portfolio. This means: Value is generated over time. Thus, value is defined more precisely as a value stream or value creation rate v(t). For example, a production system that produces items with a specific production rate can be seen as generating value during its operation, and this value might be expressed in units "$ per time" [8]. Another property of value is that one that can typically be defined just at a portfolio level, not on the single asset level. In most cases, a single asset as such has no value at all.

Consider, for example, an electrical transformer of an electricity supply network. Taken alone, this transformer has no value at all. It only creates value as a part of a network, i.e. within an asset portfolio. Thus, it is quite natural (and often required) to define value creation on the portfolio level [8]. The value is implicitly treated in the definition of the Organizational Objectives, then in the Asset Management Policy (AMP) and, finally, in the Asset Management Objectives (AMO) within the SAMP [1]. We say implicitly because, although relevant information about the value is presented in these descriptions, there is no more mention to value once the objectives are introduced. So the link with value assessment or control is indirect through Objectives (OO and AMO). With the generalization of AM (not only limited only to ISO 55000) and supporting new information technologies, the effective assessment of value is now possible. Value Components (VCs): The first step is the determination of value components. VCs are those general aspects that can be included in the value definition. In other words, the value can be expressed through its components. These components are usually employed in the Objectives description. VCs can be: social impact, other stakeholders' impact, safety, environmental, operation/capability, profits, etc. Value components can include aspects which are very difficult to evaluate in monetary terms. Most of the cases, almost any

value aspect can be translated to monetary terms. But, in many cases the models to do this are very complicated and they are not accurate at all, presenting a great level of uncertainty. Two aspects are fundamental in the value of industrial systems.

For example, in the case of a railroad system [14]:

(i) The social or public service related to the system in which it is integrated. Aspects, such as utilization, assured service times, etc. have a direct impact on the life of cities. This impact is not only on its inhabitants but on the attraction of visits and investments. Public transport is one of the main image references in any city;
(ii) The user safety. This is a key aspect in the exploitation of any means of transport. In the case of railroad systems, safety is one of their main advantages in front of any other urban transportation.

Value Factors (VFs): in this second step, we try to translate the VC to a specific term that can be used by management approach. Therefore, VFs are defined from VCs, but:

- A VF has a very accurate definition.
- A VF can be measured using specific indicators.
- The measure can be quantitative or semi-quantitative.
- A VF should be additive and scalable.
- One VF can include references to one or more VCs (i.e. we can define a unique VF for Security and Safety).
- Each VF has a relative weight for the Value composition.
- VF is related to global optimization point.

Finally, the management objects are the assets. We are going to measure and act on the assets to get value from them. Asset portfolio definition, once value has been described through VC and VF is part of OCV. It is also necessary to define how each asset contributes to gain or to loss of value.

Summarizing, the OCV include:

1. Defining the VCs (concept of value within the whole system operation). Determining the key aspects of the organization/business strategy and how to express them in terms of value.
2. Designing tools to measure and control the asset value. For this purpose, Value Factors (VF) will be used. The VF allows expressing in measurable terms all thoe aspects that constitute the value for a given business/organization.
3. Defining assets of the whole system. The definition of value determines “for what” the organization needs or will use the assets. Only on that basis, it is possible to precisely define the assets of an organization. It translates (at a practical level), the structure and tasks that include the own definition and description of the established value criteria.

2.1 Positive and Negative Asset Value Assessment (AVA)

Establishing the OCV, the own concept of value according to the company or organization point of view, it is possible to define methodologies to measure the value. The measures are always related to the asset. The asset is the element that provides value. These methodologies should be preferably objective and be included in the decision-making processes of the company. In general terms, processes can be classified as a strategic, tactical and operational level. This vision is complemented by the time horizon considered in the decision making. Consequently, methodologies for the evaluation of value according to the asset impact over the whole system process and goals will be fundamental in any asset management model. In our approach, value is composed as commented by a series of factors. These VFs can be used in two types of methods:

- Positive Control: Assessing the contribution or gain during the asset useful live. It can be the contributed value until a specific moment or the expected gain until the end of the useful life.
- Negative Control: Assessing the value reduction or loss due to events that have caused a reduction of the value or it is foreseen they will cause an expected loss of value. Failures fall into this category. Each potential asset failure has, as a risk consequence, a potential loss of value. This can range from low impact to catastrophic.

The suitable application of these methods will depend on how the VF is defined. The positive contribution, for instance, may be calculated considering the contribution of asset availability in reference to the whole system availability, in order to obtain an expected production or gain. The negative contribution may be calculated from criticality assessment of the specific asset. Both, positive and negative terms, will depend on the nature of the asset itself, but also, on the asset location throughout the whole system configuration.

2.2 Asset Value Assessment Method Example

As commented, there is no consensus for the definition of the value provided by a physical asset. One option proposed here is to define factors similarly to those used in the criticality analysis. Criticality analysis is basically a methodology that allows hierarchizing systems, facilities and equipment (assets). Figure 1 represents the results of real Criticality Analysis of industrial plant in the Oil and Gas sector.

Usually, it is measured similarly to risk, taking into account the frequency of a failure event, and the consequences that can be generated when that failure occurs. Failure frequency is usually a measurable aspect (or there are databases or other reference sources). On the other hand, consequences refer to the effects in terms of costs (operating, environmental, maintenance, etc. costs). It often requires a

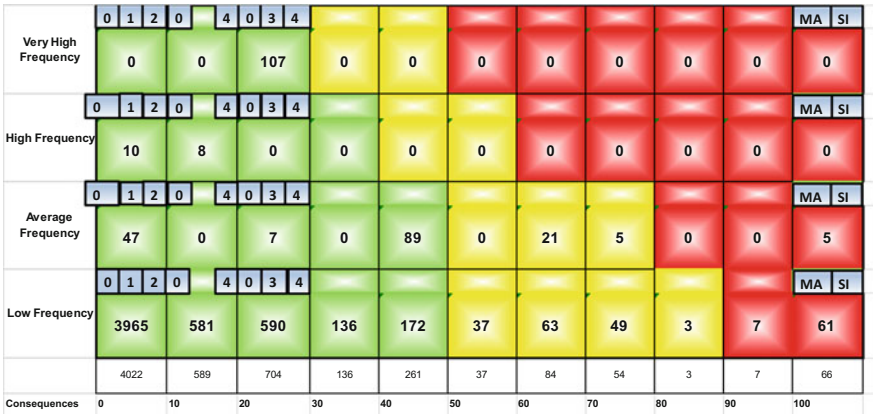


Fig. 1 Risk and value based asset criticality analysis matrix

consensus among experts. The criteria to be agreed would be, for example: Safety (are dangerous working conditions for people? are there safety measures?); Environment (is the failure dangerous for the environment? are there contingency measures?); Production (does the failure totally paralyze the plant? if not, how soon is it operational again?); O&M Costs (how much does it cost to re-commission the equipment?). Mathematically, we could formulate the concept of criticality as:

$$Risk = Likelihood \times Consequence$$

Therefore, criticality can be assessed through the estimation of loss of value when a failure event occurs. Consequences can be measures through a Value Factors accurate definition, independently from costs aspects. Moreover, it is important to consider the "value" interpretation given by stakeholders or interested parties and, in some specific cases, the value observed by the user as an intangible social benefit.

As an illustrative example (and not intended to be exhaustive), the following figure (Fig. 2) shows graphically a value quantification related to an industrial case from the point of view of the safety dimension. This *a priori* intangible dimension is forced to be quantified as an accumulation of days without accidents. The theoretical cumulative value, in this case, represents the total of days operating the system throughout the whole life cycle (35 years in this example). The Total Life Cycle Value (LCV) would represent the realistic behaviour of the physical asset throughout its useful life. No accidents would mean the achievement of the total maximal value for the asset, observed from the point of view of the safety dimension.

Future research works may be focused on the development of similar examples, although adapted to the particular casuistry of the diverse scenarios under study and different dimensions for the concept of value. Particularly, extensions to this

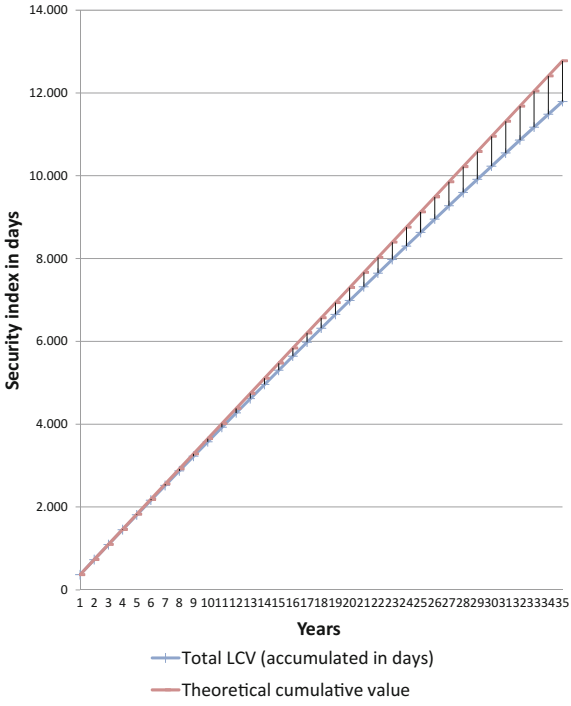


Fig. 2 Example of Life cycle value assessment

contribution may deal with the development of specific mathematical methods that allow calculating the value perception of an organization in reference to a specific asset, being such value translated into numerical quantities.

3 Conclusions

As a conclusion, it should be emphasized that physical assets have usually been managed by the departments or areas dedicated to maintenance, being in charge of the control of the installations as well as the repair and revision work, in order to assure the regular operation of the service and/or production, as well as preserving the appropriate state of the infrastructures. However, this activity has evolved in order to reach the current concept of AM by which it is understood as: "those coordinated activities in an organization, intended to obtain value from its assets" [12]. To this aim, what has been proposed in this document is an efficient approach to face the changing challenges in the industrial sector. Under these assumptions, the benefits of implementing an AM model, with an integrated approach in order to achieve value throughout the asset life cycle, are solidly proven in industry, improving service quality through its contribution to safety, human health and environmental protection.

Acknowledgements This research work was performed within the context of Sustain Owner ('Sustainable Design and Management of Industrial Assets through Total Value and Cost of Ownership'), a project sponsored by the EU Framework Program Horizon 2020, MSCA-RISE-2014: Marie Skłodowska-Curie Research and Innovation Staff Exchange (RISE) (grant agreement number 645733—Sustain-Owner—H2020-MSCA-RISE-2014); and the project "DESARROLLO DE PROCESOS AVANZADOS DE OPERACION Y MANTENIMIENTO SOBRE SISTEMAS CIBERO FISICOS (CPS) EN EL AMBITO DE LA INDUSTRIA 4.0", Ministerio de Economía y Competitividad del Gobierno de España, Programa Estatal de I+D +i Orientado a los Retos de la Sociedad.DPI2015-70842-R, financed bycon FEDER (Fondo Europeo de Desarrollo Regional).

References

1. ISO (2015a) ISO 55000:2015. Asset management—overview, principles and terminology
2. Sola A, Crespo A, Guillen A (2015) Bases para la mejora de la gestión de activos en las organizaciones. Industria Química
3. López M, Crespo A (2010) Modelling a maintenance management framework based on PAS 55 standard. Qual Reliab Eng Int. https://doi.org/10.1002/qre.1168
4. ISO (2015b) ISO 55001:2015. Asset management—management systems—requirements
5. ISO (2015c) ISO 55002:2015. Asset management—management systems—guidelines for the application of UNE-ISO 55001:2015
6. Cronin JJ, Brady MK, Hult GTM (2000) Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. J Retail 76(2):193–218
7. Harrison JS, Wicks AC (2013) Stakeholder theory, value, and firm performance. Bus Ethics Q 23(01):97–124
8. Heitz C, Goren L, Sigrist J (2016) Decision making in asset management: optimal allocation of resources for maximizing value realization. In Proceedings of the 10th world congress on engineering asset management. https://doi.org/10.1007/978-3-319-27064-7_25
9. Campbell JD, Jardine AKS, McGlynn J (2016) Asset management excellence: optimizing equipment life-cycle decisions. CRC Press, Boca Raton
10. Kaplan RS, Norton DP (1992) The balanced scorecard: measures that drive performance. Harvard Bus Rev 70(1):71–79
11. Allee V (2008) Value network analysis and value conversion of tangible and intangible assets. J Intellect Capital 9(1):5–24
12. Sola A, Crespo A (2016) Asset management principles and frameworks. Aenor, Madrid
13. Amadi-Echendu JE, Willett R, Brown K, Hope T, Lee J, Mathew J, Yang BS (2010) What is engineering asset management? Definitions, concepts and scope of engineering asset management. Springer, London, pp 3–16
14. UIC (2016) UIC railway application guide—practical implementation of asset management through ISO 55001. UIC
15. Hofmann E, Locker A (2009) Value-based performance measurement in supply chains: a case study from the packaging industry. Prod Plann Control 20(1):68–81
16. IAM (2011) Asset management—an anatomy. Institute of Asset Management, UK

An Optimised Energy Saving Model for Pump Scheduling in Wastewater Networks

Neda Gorjian Jolfaei, Bo Jin, Christopher Chow, Flavio Bressan and Nima Gorjian

Abstract The cost of electricity used for pumping in water and wastewater networks typically represents the largest part of the total operational costs. These days, South Australia has the highest electricity rate in Australia. Hence, energy management is becoming increasingly more critical in water and wastewater sectors in this state. Most sewer networks and pump stations operate based on the common high/low sewage levels and not taking into account energy costs associated with pumping. Scheduling the pumps in these systems is a smart choice for saving more electricity cost. The intelligent and smart control of the utility's assets could present operational cost saving opportunities by considering the price of purchased electricity from the spot market; however, this must be balanced with the environmental constraints to manage system odours, spills and overflow. The purpose of this study is to improve and optimise the pump control switching with the aim of reducing electricity consumption costs. To this end, the hydraulic modelling approach using Infoworks ICM has been used to simulate the performance of the pump controller. This model considers both the electricity spot price and sump elevation as inputs into smart control logic programming to operate pumps more efficiently. Results show this smart controller improves conventional pump switching models in terms of energy optimising and cost savings.

N. Gorjian Jolfaei (✉) · B. Jin
School of Chemical Engineering, University of Adelaide, Adelaide, Australia
e-mail: neda.gorjianjolfaei@adelaide.edu.au

B. Jin
e-mail: Bo.Jin@adelaide.edu.au

C. Chow · N. Gorjian
School of Natural and Built Environments, University of South Australia, Adelaide, Australia
e-mail: Christopher.Chow@unisa.edu.au

N. Gorjian
e-mail: Nima.GorjianJolfaei@unisa.edu.au

F. Bressan
South Australia Water Corporation, Adelaide, Australia
e-mail: Flavio.Bressan@sawater.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_20

1 Introduction

Energy is a key business and Australia is the world's ninth largest energy producer, accounting for about 2.4% of world energy production [9]. Energy costs comprise part of the largest expenditure for nearly all water and wastewater utilities worldwide [10]. Literature shows that utilities use approximately 3% of total electricity production in developed countries such as United Kingdom, United States and Australia [11]. A large amount of electricity is consumed by wastewater operations due to an increased rate of sewage every year. This is largely due to the spreading of cities and their resulting population growth. For this reason, there is a pressure to reduce the energy consumption in public and private sewage operations [6]. One of the greatest potential areas for energy cost savings is the scheduling of daily sewage pump operations [10].

Between 90 and 95% of the electricity purchased is used for sewage pumping in utilities [1]. The cost of energy is often related to the time of day at which the energy is used. In order to promote the use of off-peak energy and hence provide smoother loading of energy production facilities, different energy rates have been introduced by many energy providers. Therefore, avoiding peak hour pumping and having effective and optimised pump scheduling is one of the ways to reduce energy costs and thus decrease operating and maintenance costs for wastewater network operators [12].

Generally, sewer networks are divided into two types, gravity and forced by pumps networks [2]. Ideally, efficient sewer systems are designed to drain sewage by gravity related to the topology where the sewer flows from the high point to the low point [3]. However, pump stations are often required subject to topology, ground conditions, location of wastewater treatment plants and other factors [5]. Pump stations are typically controlled by conventional on/off switching based on sewage elevation in the inlet wet well without considering energy costs. This type of control would lead to poor performance across a variety of performance indicators, including energy costs, hydraulic performance and efficiency. It is a major challenge to improve the conventional switching for energy optimisation and cost savings. In this case study, the smart controller intakes two main inputs including electricity spot price and sump elevation in both dry and wet weather conditions. This smart controller has a list of logic and rules in the form of an IF-THEN statement that combines these inputs to generate pump control commands [8].

According to the large part of electricity used by sewage pump stations and the highest electricity rate of South Australia in Australia, South Australian Water Corporation is strongly seeking methods to reduce the electricity consumption in sewage pump stations. Hence, this case study attempts to decrease the electrical energy usage by optimising the pump switching regime in wastewater networks. The remainder of this paper is organised as follows. Section 2 illustrates the conventional on/off pump switching in wastewater networks. Section 3 describes the proposed smart controller model. Then, a case study is conducted to simulate and validate the proposed model. Section 5 shows the case study results. Section 6 presents the conclusions.

2 Conventional Pump Switching

A sewage pump station consists of a wet well that holds wastewater and a number of pumps in order to empty the wet well in accordance with the control programme executed by a Programmable Logic Controller (PLC). In general, there are two pumps (i.e. 'duty' and 'standby' pumps) which are utilised in alternation under the normal operating condition. When the duty pump requires maintenance, the standby pump will be turned on. However, during a wet weather event it is expected that both pumps would operate. Figure 1 is an example of a sewer pumping station operation envelope.

As can be seen in Fig. 2, sewage pumping control is a simple on/off control system based on the fluid level in the wet well.

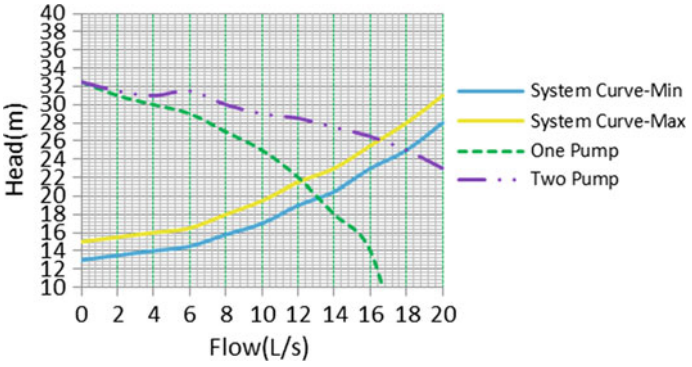


Fig. 1 Example of a sewer pumping station operation envelope [7]

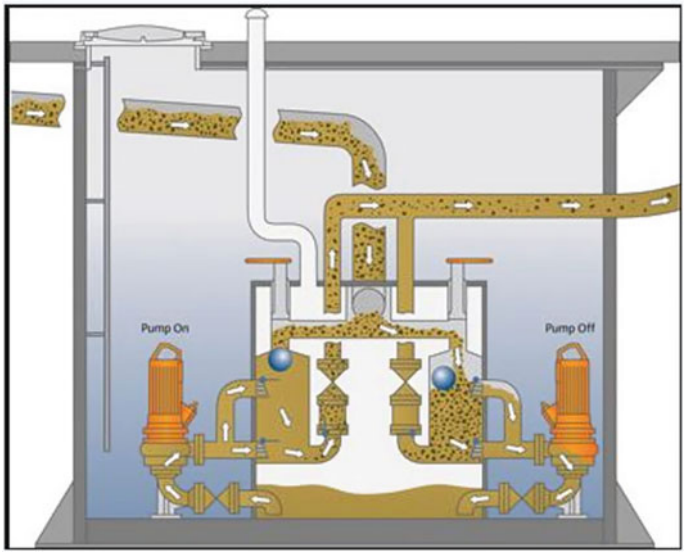


Fig. 2 A typical conventional sewer wet well

Local PLCs start and/or stop the pumps based on sensors or float switches detecting the level within the wet well using predetermined set points. The PLC generally uses the control logic algorithm. This algorithm includes accumulating wastewater in the wet well until the liquid reaches the duty pump switch on level and then this pump will be started. It runs until the switch off level in the wet well is reached. The pump duty assignment is cycled between the pumps ensuring both pumps are operated approximately equally in order to retain equipment reliability, availability and maintainability.

3 Smart Controller Modelling

The smart controller uses two key inputs including the sewage elevation and electricity price in order to generate pump status (on/off) as an output. The starting point to model the smart controller is to create logic control programming. To this end, a series of logic ranges would be allocated to inputs. As can be seen in Fig. 3, four ranges are considered for electricity price from one to four (i.e. Low Tariff, Low-Medium Tariff, High Medium Tariff and High Tariff). In addition, three ranges from Low to High are assigned to the sump elevation. The smart controller has a list of logic rules in the format of 'IF-THEN' statements that can calculate the pump status as an outcome. Infoworks ICM and its built-in-type Real Time Control (RTC) editor are applied to assess and simulate 'IF-THEN' logic rules.

Modelling the smart controller within Infoworks ICM requires extensive use of the RTC function. The RTC function contains six commands that can be used to control flows via pumps. The six commands are described as follows.

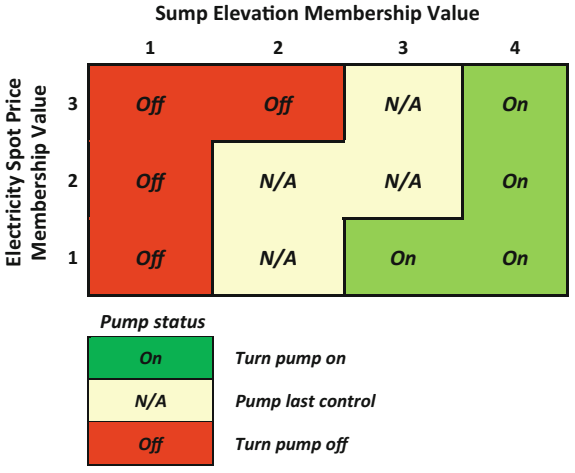


Fig. 3 Smart controller rules and logics

Range A range can be set for a variable, either created or from within the simulation.

Table This allows for data to be entered that can be used in a number of ways. In this simulation it has been used to input data in order to create a variable.

Variable This enables the user to create a new variable that is not currently within the simulation. The new variable may be entirely from inputted table data or related to current variables within the simulation.

Logic This command/function essentially sets up the IF side of the control. It is required to select different ranges and an operator (e.g. AND, OR) that will be used in the IF statement.

Rule This command/function acts as the THEN side of the control. Rule allows the pump to be set at either a status (on/off) or flow rate (i.e. 10 L/s) if a 'logic' or 'range' is satisfied.

The output of these functions is dependent on the order in which they are listed in the RTC editor. Therefore, if any command/function is referencing another then it should be listed below. Within the RTC editor, the commands/functions can be entered either under a global section or an individual structure's (pump) control. Any commands/functions entered under global will affect all structures within the network and be available to use for any range, variable, logic or rule command/functions in structure controls.

4 Case Study

The Murray Bridge wastewater network was selected for this case study. The Murray Bridge wastewater network is located on the bank of the Murray River in the south-east of Adelaide. The wastewater network contains 31 pump stations in total which covers both residential and industrial users (refer to Fig. 4). Out of 31 pump stations, 26 of them have been modelled using the smart controller. Infoworks ICM simulation has been applied in this case study for the energy optimisation and cost savings. Infoworks ICM is hydraulic extended period of wastewater networks.

Pump station 1 has variable speed control pumps; therefore, it would be more complicated to apply the smart controller modelling. Additionally, this pump station operates within a small wet well elevation with pumps in operation close to all the time; hence no optimisation in the order of pumping would be possible for this pump station. The other four pump stations were not modelled as they either were not included within the hydraulic model created nor have all required data readily available for the control system to be implemented.

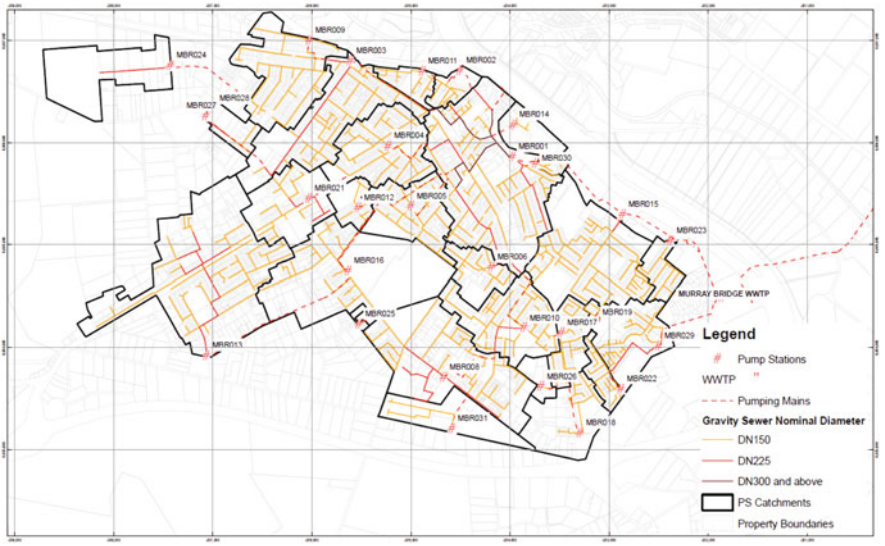


Fig. 4 Murray bridge wastewater network [4]

Scenarios run on both dry and wet weather conditions. Inputs have been provided from different sources. Electricity spot price and sump elevation are considered as inputs in the smart controller and 'IF-THEN' statements determine the pump status including on, off or remaining on the last control which is shown in Fig. 3. Electricity spot prices have been collected from Australian Energy Market Operator (AEMO) at 5-min increments in July 2016. Initial sump elevation and initial pump status collected from the hydraulic model using Infoworks ICM.

Modelling the smart controller using Infoworks ICM needs to use the RTC editor which developed as a version of the smart controller. The RTC commands and functions are; Range, Table, Variable, Logic, Controller and Rule. Figure 5 is an example of how the RTC editor has been used to model a smart controller via Infoworks ICM. Below the figure is a detailed version of the concept used in each step of the real-time control.

During creating the RTC controls, all global commands and functions are completed first. The first step is to create a range for time set as that time span of the electricity data used. A table is then formed to enter electricity pricing data that is a function of the time range set above. A variable is then created from the table of electricity data such to allow ranges to be created from different times in which electricity prices are classed as high, medium and low. Ranges of electricity prices can then be entered through using the range command. In the example three ranges used with range 1 signifying low prices and range 3 signifying extreme prices.

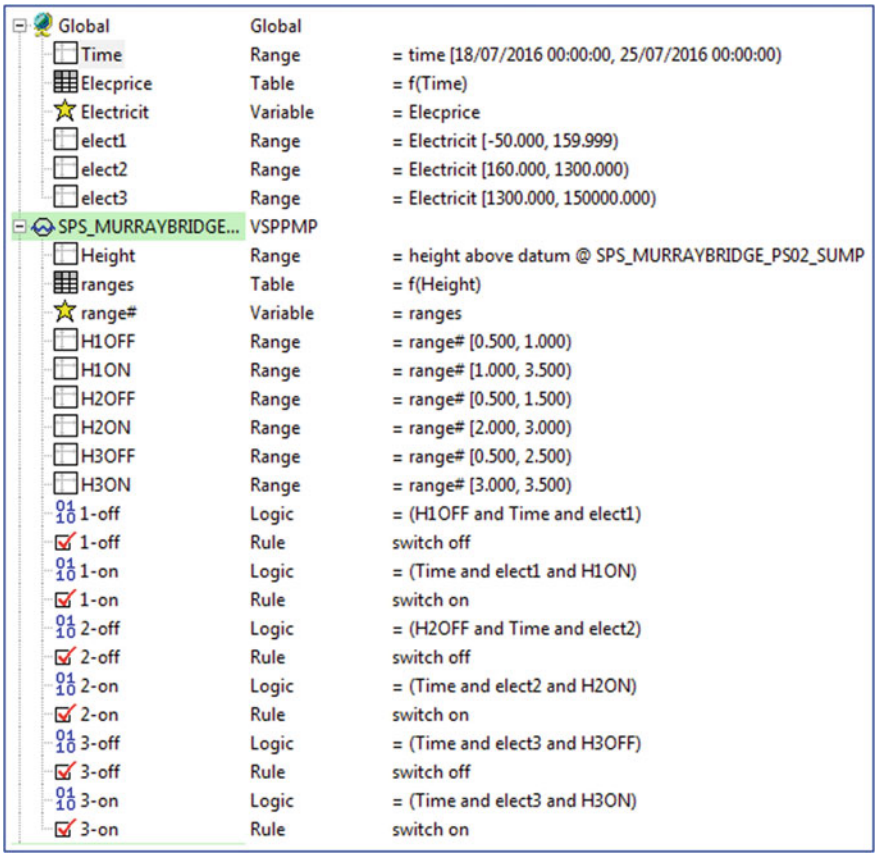


Fig. 5 RTC editor controls for pump stations (extracted from Innovyze Infoworks ICM application)

In addition, a table is then created as a function of the sumps elevation and pump specific elevation ranges are entered with their associated membership values. Membership value boundaries for elevation can be adjusted by the user. For example, all pumping stations may have the same boundaries at 25, 50 and 90% or the other option is for boundaries to be individually customised. A variable is then created from the table for additional ranges based on these membership values. Ranges are used to control the heights at which pumps will be turned on or off for each electricity range. Ranges are set up based on the variable created with membership values ranging from 1 to 3. Within each OFF range the minimum has to be less than 1 and each ON range has a maximum greater than 3. It should be

noted that there is a gap between the maximum of the OFF range and the minimum of the ON range. The gap left therefore acts as a last control function. The logic function is used to combine each electricity range with the associated ON and OFF elevation ranges. Through combining these three ranges an IF style statement: IF (Time is within suitable range AND Electricity Price is Low AND Elevation is Low). According to results of this study, the implemented model could enhance the energy cost saving successfully. Therefore, the logic used to optimise the pump switching based on electricity spot market price can apply in any controller for same cases in other industries (e.g. agricultural industries and recycle water industries).

5 Results

The smart controller for optimising energy usage and cost savings in the sewer networks has been modelled within 24-hour electricity spot prices in both dry and wet weather conditions using Infoworks ICM. According to the evaluation performance of the smart controller, in terms of reducing energy costs, results for the base pump switching control (control base on sewerage elevation) has been compared with the using smart controller for pumping.

Tables 1 and 2 demonstrate the electricity cost of various pump stations in Murray Bridge wastewater network. In addition, they compare this cost for conventional pump switching and using the smart controller in one day with dry weather condition and one day with wet weather condition respectively.

Results in Table 1 show 79% energy savings during the dry weather condition across these pump stations. Table 2 shows energy saving results at the wet weather condition. These results demonstrate 84% energy cost savings across these pump stations under specific dry and wet weather patterns which tends to suggest using such as operating regime on the Murray Bridge sewerage networks could be advantages. Impact on sewage delivered into these conditions, sewage treatment plant and its optimal operation may become another constraint that needs future consideration as well.

Figures 6 and 7 present the sump level and pump switching pattern under conventional pump switching and using smart controller during a week. It is obvious that conventional pump switching works based on sump level without considering the tariff rate. However, by using the smart controller, a low level of sump in the wet well is maintained during the low-cost tariff and the high level of sump is kept up through the high-cost tariff. Thus, electricity cost has been saved.

Table 1 One day results using Infoworks ICM—dry weather condition

Pump number	02	03	04	05	06	08	09	10	11	12	13	14	15	16
Base cost	$0.76	$15.09	$3.81	$3.65	$0.51	$2.00	$2.24	$9.78	$0.20	$0.13	$6.76	$0.05	$3.45	$1.08
Smart controlled cost	$0.35	$2.50	$0.88	$0.72	$0.16	$0.56	$1.06	$1.20	$0.10	$0.02	1.5	0.013	1.3	0.24
Saving ($)	$0.41	$12.59	$2.93	$2.93	$0.35	$1.44	$1.18	$8.58	$0.11	$0.11	$5.26	$0.04	$2.15	$0.84
Saving (%)	54	83	77	80	69	72	53	88	52	85	78	74	62	78
Pump number	**17**	**18**	**19**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**29**	**30**		**Total**
Base cost	$0.63	$0.73	$0.08	$0.42	$0.40	$2.85	$0.93	$0.00	$0.02	$0.84	$1.44	$0.04		$**57.89**
Smart controlled cost	$0.14	$0.19	$0.00	$0.14	$0.00	$0.15	$0.35	$0.00	$0.00	$0.07	$0.30	$0.00		$**11.94**
Saving ($)	$0.49	$0.54	$0.08	$0.28	$0.40	$2.70	$0.58	$0.00	$0.02	$0.77	$1.14	$0.04		$**45.95**
Saving (%)	78	74	99	67	99	95	62	99	80	90	79	95		**79**

Table 2 One day results using Infoworks ICM—wet weather condition

Pump number	02	03	04	05	06	08	09	10	11	12	13	14	15	16
Base cost	$1.23	$21.33	$5.82	$5.49	$0.63	$3.16	$4.58	$10.73	$0.44	$0.18	$6.86	$0.07	$5.06	$1.22
Smart controlled cost	$0.45	$2.50	$1.05	$0.91	$0.18	$0.66	$1.30	$0.01	$0.12	$0.03	$0.66	$0.01	$1.60	$0.30
Saving ($)	$0.78	$18.83	$4.77	$4.58	$0.45	$2.50	$3.28	$10.72	$0.32	$0.15	$6.20	$0.06	$3.46	$0.92
Saving (%)	63	88	82	83	71	79	72	100	73	83	90	86	68	75
Pump number	**17**	**18**	**19**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**29**	**30**		**Total**
Base cost	$0.81	$1.14	$0.13	$0.65	$0.50	$0.69	$2.05	$0.00001	$0.07	$1.26	$2.15	$0.06		**$76.31**
Smart controlled cost	$0.22	$0.22	$0.00	$0.17	$0.05	$0.22	$0.73	$0.0001	$0.01	$0.12	$0.40	$0.00		**$12.00**
Saving ($)	$0.59	$0.92	$0.13	$0.48	$0.45	$0.47	$1.32	−$0.0001	$0.07	$1.14	$1.75	$0.06		**$64.31**
Saving (%)	73	81	100	74	90	68	64	−25	93	90	81	97		**84**

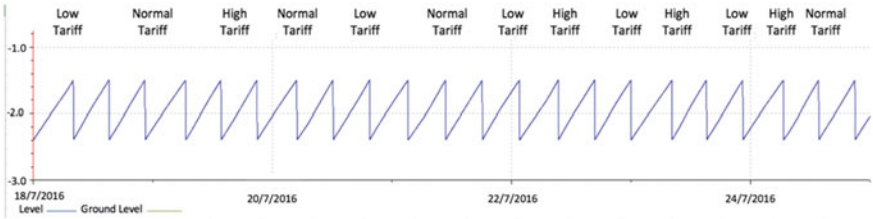


Fig. 6 Sump elevation pattern under conventional pump control during a week

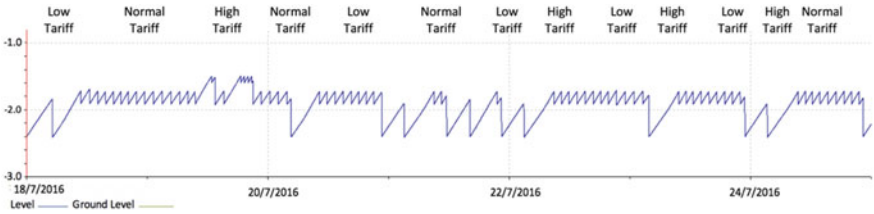


Fig. 7 Sump elevation pattern using smart controller for pump control during a week

6 Conclusions

The large amounts of energy are consumed daily in modern wastewater networks to operate pumps. Consequently, energy optimisation in pump station consumption and electricity cost savings are critical to manage and operate these systems in cost-effective manner. The majority of pump stations in sewer networks utilise conventional control for pump switching based on only sewerage levels. In order to have optimised operating expenditure, it is essential to consider not only sump elevation but also the electricity spot price to generate an optimal smart controller model in sewer networks. This study presents a new intelligent and smart controller to optimise pump switching at the Murray Bridge wastewater pump station. To this end, two major inputs, electricity price and wet well level and a list of rules including IF-THEN statements were considered to produce pump switching commands by the smart controller. In addition, Infoworks ICM and its built-in-type RTC editor were used as an assessment and simulation method for modelling the smart controller. The potential energy consumption optimising and cost savings demonstrated within the case study. According to the results, the smart controller could help to reduce energy usage and then enhance electricity cost savings. Findings show 79% cost savings of electricity after using the smart controller compared to the conventional controller in a day with the dry weather simulated condition. Moreover, in a day with the wet weather condition, 84% electricity cost savings has been exposed by applying this smart controller. Further study could be carried out to assess and analyse influences of pump operations on pump maintenance in terms of the number of pumps switching and try to decrease this factor to

improve asset reliability and maintainability. Moreover, environmental constraints such as sewage defect time and spill need to investigate more in future.

References

1. Bunn S, Reynolds L (2009) The energy-efficiency benefits of pump scheduling optimisation for potable water supplies. IBM J Res Dev 53(3):5–13
2. Ermolin Y (1999) Mathematical modelling for optimised control of Moscow's sewer network. J Appl Math Model 23:543–556
3. Fiter M, Guell D, Comas J, Poch M (2005) Energy saving in wastewater treatment process. J Environ Technol 26:1263–1270
4. GHD Report for SA Water (2015) SA water Murray bridge wastewater catchment model build report, 33/17514. s.l.: South Australian Water Corporation Internal Report—Not Published
5. Hao R, Liu F, Ren H, Cheng S (2013) Study on a comprehensive evaluation method for the assessment of the operational efficiency of wastewater treatment plants. J Stoch Environ Res Risk Assess 27:747–756
6. Hass D, Dancey M (2015) Wastewater treatment energy efficiency. J Aust Water Assoc 42 (7):53
7. Hayde P, W X (2012) Technical guideline—hydraulic design of sewer network s.l.: South Australian Water Corporation Internal Report—Not Published
8. Konetschka M et al (2017) Developing intelligent, cost saving pump controls for wastewater networks through integration with the electricity spot market. s.l., OZWater-In Press
9. Liu Y, Ganigue R, Sharma K, Yuan Z (2016) Event-driven model predictive control of sewage pumping stations for sulfide mitigation in sewer networks. Water Res 1:376–383
10. Ostojin S, Mounce S, Boxall J (2011) An artificial intelligence approach for optimising pumping in sewer systems. J Hydroinform 13:295–305
11. Walski T, Andrews T (2015) Energy savings in water and wastewater systems. A Bentley White Paper 1–16
12. Zhang Z, Kusiak A (2011) Models for optimization of energy consumption of pumps in a wastewater processing plant. J Energy Eng 137:159–168

A Novel Approach to Sensor-Less Daylight Harvesting in Commercial Office Buildings

B. Harris, J. Montes and N. Forbes

Abstract The rise of energy costs is negatively impacting operating budgets for buildings. Energy efficient programs are being implemented during the design phase of buildings to maximise occupancy comfort and reduce energy consumption. The problem for investors is that the cost to install these energy efficient products have questionable investment returns. Through theoretical and subsequent empirical analysis, this paper introduces a mathematical equation that replaces existing sensor controlled daylight harvesting control systems. This is achieved using an exponential distribution that takes a weather station lux reading to identify the natural light at any given point in a building. This equation makes it possible to calculate the dimming level (the level by which the light fitting reduces in power to create artificial light) required to achieve a designed lighting level that uses both natural and artificial light. This sensor-less daylight harvesting method will provide building owners with an alternative, less expensive, and more efficient daylight harvesting control system.

1 Introduction

Daylight Harvesting (DH) in a commercial building considers the illumination effect of natural light to offset artificial lighting illuminance at a working plane or building gathering place [10]. The DH benefit to the owner is that the demand for power is reduced, thereby reducing light fitting power consumption costs and maintenance factors as the life of the light fittings is extended. Academic research

B. Harris (✉)
Queensland University of Technology, Brisbane, Australia
e-mail: bharris@fredon.com.au

J. Montes · N. Forbes
Fredon QLD Pty Ltd., Underwood, Brisbane, Australia
e-mail: fredoncad@fredon.com.au

N. Forbes
e-mail: nforbes@fredon.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_21

suggests that the power demand of artificial lighting equates to "approximately one-third of electricity used in commercial buildings" [7]. Understanding the actual power consumption of a light is needed to identify the return on investment (ROI) and to evaluate the effectiveness of a proposed lighting cost minimisation strategy. Observations of industry practice revealed that DH in both new construction projects and existing buildings is not regularly adopted. This is supported by research that suggests "lighting control is not popular", and that the unenthusiastic adoption of DH lighting control is due to the absence of local daylight data and lighting cost analysis [4].

The energy consumption of a lamp is "roughly proportional" to the output illuminance [8]. Due to this proportionality, [4] identified predictive energy cost savings is possible by using a daylight factor (DF). This supports the evidence that dimming light fittings with sensors will decrease energy costs [2, 3]. However, the perceived problem with implementing this theoretical finding lays in the functionality of lighting control systems that is limited to manual test and pre-set scenes, or installing lux sensors in the ceiling located at determined intervals around the perimeter of an office space. Whilst the benefit of dimming control in relation to energy efficiency has been thoroughly researched, linking this literature to investor ROI appears limited. A google search of DH identified numerous companies that promote the benefits of their products and the energy savings that is possible from DH, but don't provide any ROI data, highlighting the gap in industry and literature on DH processes and lighting control interfaces.

This novel approach differs from other sensor-less DH (SDH) studies that are limited to using photovoltaic (PV) generation [9]. The PV method is deemed an unacceptable approach to harvesting power from dimming lights in commercial buildings. This is because the study does not address regulatory lux level requirements, variables such as shadowing from neighbouring buildings, cost of implementation, and the effect that random clouding has on internal building light levels. This paper demonstrates that SDH is possible, and with future research it can be automated to provide owners with a cost-effective DH solution.

2 Methodology

An SDH system requires an addressable lighting control system, mathematical software, lux level meters, data logger, and a roof top weather station. Light modelling software such as AGI or Dilux can be used to provide an estimate of natural light levels, but is not necessary. The option to use computer aided drawings (CAD) to simulate the light conditions is a costly exercise, however for this simulation, the AGI data was used to simulate the office light levels.

2.1 Lux Modelling Software

Current lighting design practices generally do not include external lighting sources other than artificial light sources modelled in the AGI32 software program. To estimate the natural light level, a transition glass setting is used. Verifying the calibration process to achieve a "real life" internal natural lux plot calculation was not carried out, as this process was proven possible by Kacprzak and Tuleasca [2].

The AGI32 software program was used to identify the relationship between external light and internal light at three internal plot (x) locations (Fig. 1). This provided a platform to verify existing literature, and subsequently led to the discovery of a scalar factor (SF) that is defined in Sect. 2.3. In addition to this, recording the lux levels at varying distances from the window (Fig. 2) identified an exponential lighting distribution curve that has been recognised in literature as a fact [6].

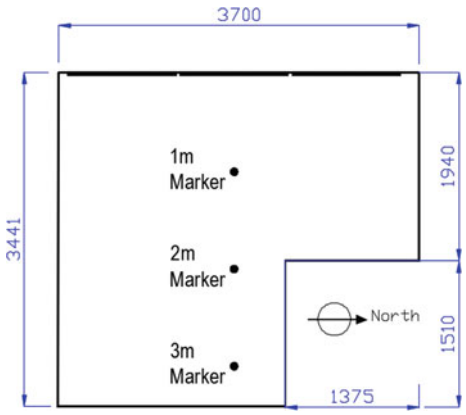


Fig. 1 Room layout

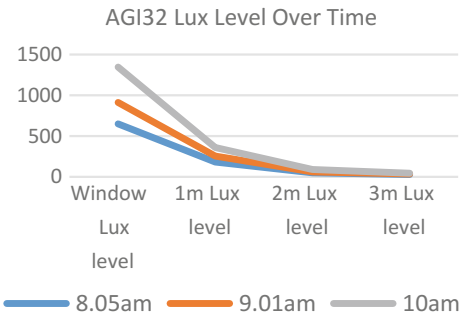


Fig. 2 AGI32 lux level in room exponential shape test

2.2 Modelling the Effects of Lighting Variables

To effectively maintain lighting levels to Australian/International Standards, SDH considered the following variables and the ease at how these variables can be calculated and incorporated in the design and commissioning phase of a SDH system.

External	Internal
1. Atmospheric variables that include; (a) clear sky with direct sunlight, (b) random cloud events, and (c) overcast skies 2. Position of sun during the four seasons, winter, autumn, summer, and spring 3. Shadowing impact to the lux levels from block structures such as neighbouring buildings, trees, or bridges etc	1. Window tinting and internal partitioning 2. Interior partitions, interior design/art work, and alterations over time 3. External light entering a room from different sides 4. Blinds (blinds in a building that do not have auto control will be set in the SDH program to the translucency factor of the blind)

The **external variable** is managed by a weather station and onsite commissioning that adjusts the SF to the floor areas that are affected by shadowing structures. The **internal variables** are deemed to be constant, and therefore the SF accounts for these variables during the SDH commissioning phase. Lighting control systems have the capability of managing blinds, but that is a lighting control design parameter that will be addressed in future research.

2.3 Exponential Function—One Side Light Entry Only

Experimental tests using both the AGI32 software and samples taken using the light meter demonstrated a correlation between the rise in the indoor to outdoor lux levels. This ratio of internal illuminance to external illuminance is defined as 'Daylight factor' (DF) by the International Commission on Illumination (CIE). The mathematical definition is as follows:

$$DF = \frac{internal\ illuminance}{external\ illuminance} * 100\% \tag{1}$$

The difficulty with using this DF equation to accurately predict the dimming threshold to control artificial light, is that the variables impacting external light penetrating the building façade remain unaccounted for. Furthermore, reflectance of internal room surfaces, the orientation and location of the room, the size of the room, and amount of direct and or indirect light are all variables that need to be calculated to accurately determine internal illuminance. Likewise, external illuminance depends on the time of day, the date, atmospheric conditions and weather patterns. CIE adopt the approach of minimising some of these variables by

standardising the DF definition to overcast sky conditions [1]. This research used the AGI32 software to provide internal lux levels for natural light using CIE overcast sky conditions [5]. The simulation of a full day revealed that once the initial variable in the building structure is calculated (which was the size of the window and the layout of the room in the sample), the rise in internal and external lighting levels was linear. Therefore, the hypothesis is that irrespective of the weather condition, as long as the variables remain constant, a linear relationship between the amount of external illuminance, in proportion to the internal illuminance at a particular point will occur. SF is dependent on the exponential function of the decaying luminance intensity, and is different (exponentially) at any given distance from the window.

$$SF = \frac{external\ Lux\ value}{internal\ Lux\ value} \tag{2}$$

Using Eq. 2, the external lux value can be determined by using an exponential function that considers the horizontal grid sample of internal lux values at incremental distances (x) from the window, and multiplying by SF. Similarly, the external lux value divided by SF will result in the internal value at each grid point. This linear relationship of external light to internal light, results in SF with the proportionality being dependant on 3 variables as described in International Commission on Illumination (CIE); sky component (SC), externally reflected component (ERC), and the internally reflected component (IRC) at each sample grid point.

Through the calculation of SF from actual grid point samples, light level variables can be accounted for, and SF can be adjusted during the "building tuning" phase to provide greater accuracy to light levels at the working plane [2]. Table 1 shows the lux values recorded from the AGI32 room simulation and the distance plots from the window.

The function of the dataset from Table 1 can be represented using the following base intercept form:

$$\begin{aligned} &P(0)a^x = P(0)e^{rx} \\ &which\ can\ be\ solved\ for\ a\ to\ achieve \\ &a^x = e^{rx} = e^{(r)x} \quad therefore, \\ &a = e^r\ and\ r = ln(a) \end{aligned} \tag{3}$$

Table 1 Lux levels from office space experiment

Distance (x) from window (m)	Internal lux level (y)	External lux level
1	1345	15,113
2	913	15,113
3	650	15,113

where a is calculated using the sample values of (x1, y1) and (x2, y2).

$$a = \left(\frac{y2}{y1}\right)^{\frac{1}{x2-x1}} \tag{4}$$

To calculate y using P(0) (initial value of x from 1 to ∞). The sample data in Table 1 of 1 m = 1345 lx, and 2 m = 913 lx is calculated.

$$\begin{aligned} r &= \left(\frac{\ln(913) - \ln(1345)}{2-1}\right) r = -0.387413411 \\ y &= P(0)e^{-0.387413411x} \\ &\text{and } y = 1345 \\ 1345 &= P(0)e^{-0.387413411*1} \\ P(0) &= 1345/(e^{-0.387413411}) \\ P(0) &= 1934.76 \\ y &= 1934.76e^{-0.387413411x} \end{aligned} \tag{5}$$

Equation 5 is derived from two coordinate points which can produce a higher error rate. A more accurate approximation of Ilx can be derived by calculating P(0) and r at each sampled point of x, where n is the number of sampled values:

$$\frac{1}{n}\sum_{i=1}^{n} P(0)i \text{ and } \frac{1}{n}\sum_{i=1}^{n} ri \tag{6}$$

Therefore, to improve the accuracy of P(0) and r, at least 3 coordinate point samples is recommended, as shown in Eq. 6. Increasing the number of samples will increase the accuracy of the rate of change, however considering the commercial application of SDH, it is determined that 3 lux plot tests will satisfy compliance to lux level standards.

$$y = 1919.57e^{-0.363588465x} \tag{7}$$

To verify the accuracy of the algorithm, Table 2 shows y (internal lux level) for each value of x (distance from window) in comparison to the original lux values.

Table 2 Internal lux sample compared to exponential lux value

Distance (x) from window (m)	Internal lux level (y)	Internal exponential lux level (y)
1	1345	1334.44
2	913	927.68
3	650	644.9

2.4 Calculating Dimming Values

When Calculating the dimming value (DimX) at any sample location, variables such as reflection and shadowing can have an impact on measurements. These effects can be minimised by controlling the sampling environment. Controlling the effects of these variables becomes more problematic when a room records light entering from multiple directions. To model the dimming value in this scenario, lighting software can be calibrated to reflect the artificial light to daylight relationship [2] which will provide the values needed to calculate SF at each lighting point. However, site sampling under each light is considered the best calibration option. This is more expensive to implement, but the accuracy that this testing ensures is that lighting standards are met. Further research into the effects of shadowing and reflection, particularly in multiple side light entry scenarios is recommended. This paper however, deals with the single side lit scenario and a combination of sampling lux levels to identify Elx where the internal lux level = Ilx is;

$$Elx = Ilx \times SF \text{ or } Ilx = \frac{Elx}{SF} \text{ or } SF = \frac{Elx}{Ilx} \tag{8}$$

Using 2 of the 3 variables, Elx, Ilx or DimX, the light level standard that determines the amount of light in a room can be maintained by introducing the design value (DV) into the following equations;

$$DimX = 1 - \left(\frac{Ilx}{DV}\right) \text{ or } Ilx = (1 - Dimx) \times DV \tag{9}$$

Dimming calculation (DimX) from external lux level:

$$Elx = ((1 - DimX) \times DV) \times SF \tag{10}$$

Therefore

$$Ilx = (1 - DimX) \times DV \text{ or } \frac{Elx}{SF} = (1 - DimX) \times DV \tag{11}$$

3 Exponential Comparison Between Single Side Light and Multiple Light Entries

The exponential calculation in Sect. 2.3 is used to determine internal light levels and corresponding dimming levels at a specific point. When there is a single side light source penetrating a room, a minimum of 3 sample plots are taken at a distance from the window to the focal point (darkest part of the room). In Fig. 3 the focal

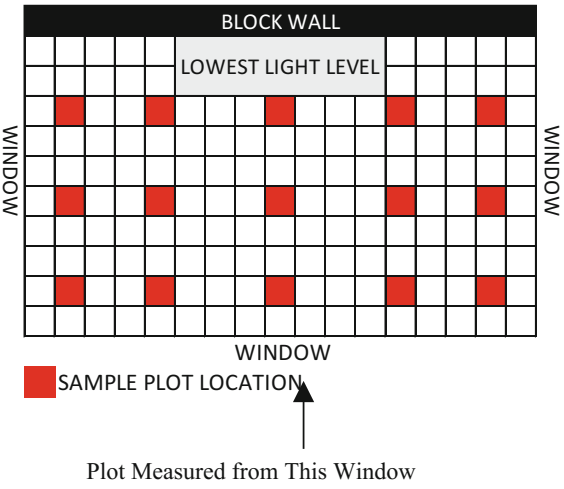


Fig. 3 Lux plot sample examples

point is the back of the room (block wall) and measurements are taken at equal intervals in a straight line forming a grid to capture the light from the 3 windows.

Using the algorithms in this paper and the AGI sample plots (Table 3) in MATLAB, the exponential light distribution is demonstrated graphically in Fig. 4 using an external lux value of 500 lux. The plots in Fig. 4 identify the natural light level that determines the dimming levels that is needed to maintain a design of 320 lux in the sample space.

Table 3 Sample plot (in bold) exponential distribution

Distance from window (m)	Grid spacing: position	Outside lux			15,356	
		Grid no. 1	Grid no. 2	Grid no. 3	Grid no. 4	Grid no. 5
1	Pos no. 1	7196	5827	5503	5830	7195
3	Pos no. 2	5745	3997	3572	4001	5740
5	Pos no. 3	5196	3273	2784	3270	5189
7	Pos no. 4	4915	2873	2369	2869	4912
9	Pos no. 5	4406	2448	2028	2443	4396

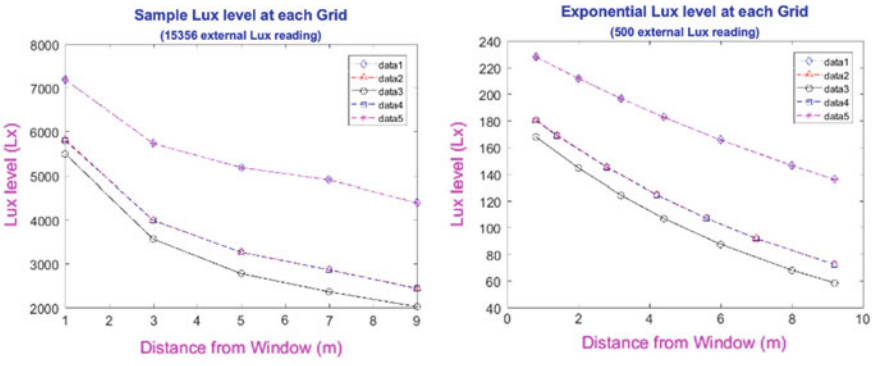


Fig. 4 Sample plot and exponential lux plot at 500 lux

4 Future Research

The practical use of this algorithm to automate SDH using existing lighting control technology has been implemented on a commercial project by Fredon Industries. Research regarding the efficiency of SDH and its financial benefits needs to be expanded to include; (a) lifecycle benefits using AS/NZS 4536 (Life cycle costing—an application guide) to estimate ROI, (b) technical integration requirements into lighting control designs, (c) the possibility of removing the building weather station and use a Really Simple Syndication (RSS) feed from the Bureau of Meteorology to determine SF, (d) modelling lux dimming levels using matrices, (e) the monitoring of causal condition benefits that will support a business case for adopting a SDH system, (f) the impact that varying artificial light has on occupants wellbeing, and (g) the effects of shadowing and reflection when sampling lux levels, particularly for multiple side lit spaces when light enters a space from four directions.

5 Conclusion

The above simulation proves that a sensor-less approach to daylight harvesting is possible. This research measures external light levels and uses an exponential equation to determine the natural lighting level at any location on a building floor. Once a floor is modelled, SF is used to identify the dimming value and 3-way relationship between internal, external, and artificial light to maintain a designed lux level. This novel approach to daylight harvesting will provide owners with increased energy savings, extend the life of a light fitting, and provide metrics to improve energy efficiencies, which will benefit Building standards reporting.

References

1. Bian Y, Ma Y (2017) Analysis of daylight metrics of side-lit room in Canton, south China: a comparison between daylight autonomy and daylight factor. Energy Build 138:347–354
2. Kacprzak D, Tuleasca I (2013) Daylight savings calculation technique for smart buildings. Intell Control Autom 4:102–107
3. Li DHW, Lam TNT, Wong SL (2006) Lighting and energy performance for an office using high frequency dimming controls. Energy Convers Manag 47:1133–1145
4. Li DHW, Mak AHL, Chan WWH, Cheng CCK (2009) Predicting energy saving and life-cycle cost analysis for lighting and daylighting schemes. Int J Green Energy 6:359–370
5. Lighting A (2016) AGi32 version 17 [Online]. Available: http://docs.agi32.com/AGi32/Content/daylighting/Daylighting_Overview.htm. Accessed 10-12-2016
6. Singhvi V, Krause A, Guestrin C, Garrett JH, Matthews HS (2005) Intelligent light control using sensor networks. In Proceedings of the 3rd international conference on Embedded networked sensor systems. ACM, pp 218–229
7. Soori PK, Vishwas M (2013) Lighting control strategy for energy efficient office lighting system design. Energy Build 66:329–337

8. Ticleanu CM (2015) Aiming to achieve net zero energy lighting in buildings. Revista Romana de Inginerie Civila 6:63–78
9. Yoo S, Kim J, Jang CY, Jeong H (2014) A sensor-less LED dimming system based on daylight harvesting with BIPV systems. Opt Express 22:A132–A143
10. Yu X, Su Y (2015) Daylight availability assessment and its potential energy saving estimation—a literature review. Renew Sustain Energy Rev 52:494–503

Development of Autonomous Hammering Test Method for Deteriorated Concrete Structures Based on Artificial Intelligence and 3D Positioning System

Katsufumi Hashimoto, Tomoki Shiotani, Takahiro Nishida, Hideo Kumagai and Katsuhiko Kokubo

Abstract It is very important nowadays to establish economic and efficient management systems for existing concrete infrastructures, in order to fulfill their service designed lives and even to extend them. Hammering test, which is a manual inspection technique detecting defects and damages, is widely used to judge the soundness of tunnels, bridges, roads and railways. However, the reliability of the diagnosis results using hammering test data is not described quantitatively but handled only with inspectors' professional experience. This study aims to propose an autonomous hammering test and deterioration diagnosis system, developing artificial intelligence (AI) systems which recognize the internal defects of reinforced concrete members. As for the three-dimensional (3D) positioning system installed in the autonomous solenoid hammering device with a microphone, it is also equipped with a gyro sensor and an accelerometer so that it is ongoingly collecting the hammering sound and detecting the 3D location of hammering location in real time of inspection on site. Consequently, in the present study, the frequency spectrum analysis based on the wavelet transform of acoustic hammering sound data gives the recognition for deterioration probability

K. Hashimoto (✉) · T. Shiotani · T. Nishida
Kyoto University, Kyoto, Japan
e-mail: hashimoto.katsufumi.8a@kyoto-u.ac.jp

T. Shiotani
e-mail: shiotani.tomoki.2v@kyoto-u.ac.jp

T. Nishida
e-mail: nishida.takahiro.6e@kyoto-u.ac.jp

H. Kumagai
Tamagawa Seiki Co., Ltd., Iida, Japan
e-mail: hideo-kumagai@tamagawa-seiki.co.jp

K. Kokubo
I.T. Engineering Co., Ltd., Kawata, Japan
e-mail: kk-kokubo@it-e.co.jp

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_22

followed by the artificial neural networks with the diagnosis algorithm consisting of image binary processing of the wavelet tomogram and output system.

1 Introduction

In inspection and health monitoring of concrete structure, using a technique for identifying the deteriorated and internal defects or damage, appropriate maintenance strategy has become greatly important in these days. For this reason, many researches on non-destructive testing and evaluation methods to objectively and quantitatively evaluate the physical properties or degradation degree of concrete have been promoted widely. Applying concrete hammering test is a method for determining an inner defect from the concrete surface with the sound generated when hitting the concrete with a hammer [1]. In general, the reliability of hammering test for deteriorated reinforced concrete member is high just based on experience and knowledge of the inspector with skills and careers.

However, some previous reports regarding degradation assessment by hammering test are not yet up to indicate quantitatively the reliability of the evaluation results. In this study, frequency spectrum analysis of acoustic data was carried out after obtaining the hammering sound from the simulated deteriorated concrete specimen with deferent types of defect properties. Replacing the ability of skilled inspection technicians with quantitative values visually, this study aimed at finding new information for practice of the heuristics in degradation evaluation method with hammering test applied to RC specimens explained below.

This study aims to propose an autonomous hammering test and deterioration diagnosis system, developing artificial intelligence (AI) systems which recognize the internal defects of reinforced concrete members. As for the three-dimensional (3D) positioning system installed in the autonomous solenoidal hammering device with a microphone, it is also equipped with a gyro sensor and an accelerometer so that it is ongoingly collecting the hammering sound and detecting the 3D location of hammering location in real time of inspection on site. Consequently, in the present study, the frequency spectrum analysis based on the wavelet transform of acoustic hammering sound data gives the recognition for deterioration probability followed by the artificial neural networks with the diagnosis algorism consisting of image binary processing of the wavelet tomogram and output system.

2 Development of AI for Recognizing the Depth of Defects with Hammering Sound

2.1 Specimen

Reinforced concrete specimen (length 1000 mm × width 1000 mm × thickness 200 mm) simulating the defect was prepared with reinforcing bars. Figure 1 and Table 1 show the conditions of the overview and embedded objects. A cross (×) mark in the figure indicates a targeted point to record the hammering sound.

The objects used to simulate the defect are styrofoam with a thickness of 30 mm. After setting the reinforcing bars in the form, concrete of a formulation shown in Table 2 was cast. The experimental works subjected to the hammering test were followed by exposure to Asahikawa, Hokkaido over 28 months after concrete casting.

2.2 Hammering Test and Acoustic Data for Frequency Analysis

Detectable depth for the defects is up to 5 cm in general, although it is depending on the size of the test hammer size or weight. According to the diagnosis by two skilled technicians for hammering test on targeted No. 1, No. 2 and No. 3 (defect depth: 100, 75 and 50 mm) in Fig. 1, it was difficult to detect the defect properly.

FFT (Fast Fourier Transform) analysis was carried out to evaluate the frequency spectrum. It should be noted that all the peak patterns were normalized to the maximum value in this study. The maximum value of frequency spectrum intensity was defined as 1.0 to display the frequency spectrum on the peak intensity pattern

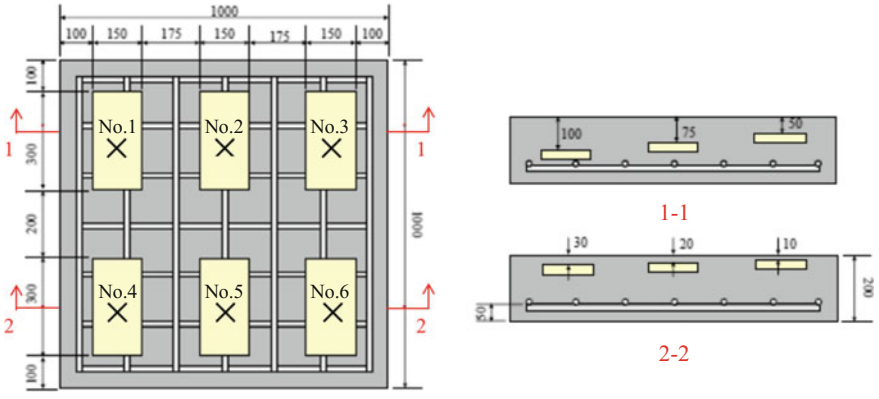


Fig. 1 Design of simulated deteriorated specimen

Table 1 Properties of embedded objects

Mark	Depth of objects (mm)	Object	
		Type	Size
①	100	Stylofoam	x:300 mm y:150 mm z:30 mm
②	75		
③	50		
④	30		
⑤	20		
⑥	10		

Table 2 Mix proportion and property of concrete

W/C (%)	s/a (%)	Unit weight (kg/m^3)					Slump (cm)	Air (%)	Temp (° C)
		W	C	S	G	ad			
49.6	41.4	157	317	746	1063	3.17	13	5.9	19

in the range of 0–1.0, so that the difference at each measurement point could be extracted in the spectral shape. Figure 2 illustrates the flow of hammering test and frequency analysis.

As shown in Fig. 3, peaks of frequency spectrum patterns shift clearly from 2000 to 200 Hz when the depth is less than 50 mm. On the other hand, the defect depth is larger than 50 mm, it is not able to find the difference in the spectrum patterns. Therefore, the present research work aimed to obtain new findings to propose quantitative information in order to visualize inspection engineers' skills and experience. In addition to the conventional frequency spectrum analysis by the Fourier transform of hammering sound waveform from the defect simulated specimen, binarization process of the time-frequency spectrum based on the wavelet transformation technique is performed.

In this paper, wavelet transformation is applied to the waveform analysis of hammering sound from using the binarized image of time-frequency spectrum is built, where the defect depth determination (diagnosis) result is automatically

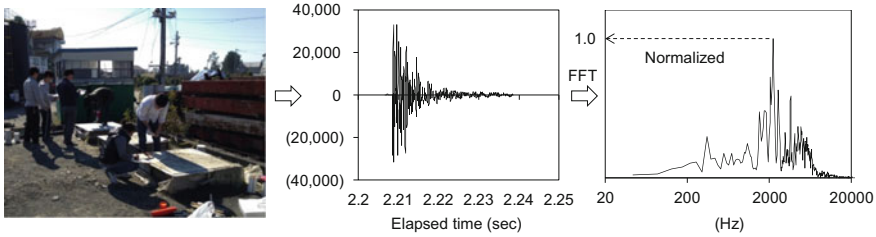


Fig. 2 Outline of hammering test and frequency analysis

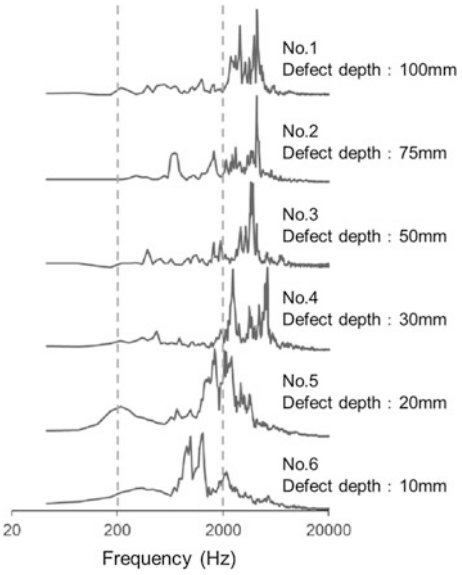


Fig. 3 Normalized frequency spectrum patterns

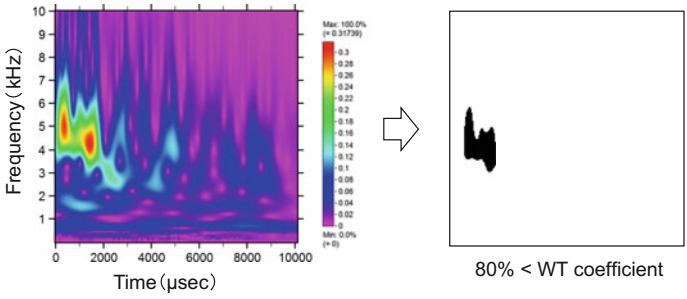


Fig. 4 Example of wavelet transform to binary image

output [2, 3]. Figure 4 shows how to pick out black and white image (320 × 320 pixels text data) by extracting an area where the wavelet coefficient is the top 20% in black.

2.3 Pattern Recognition of Binary Image

As for recognition of planar curves such as images (figure shapes), complex autoregressive model and PARCOR coefficient have been proposed as a plane curve natural description (invariant feature extraction method) [4]. Currently, image

pattern recognition (identification) techniques are used in various fields and widely used in medical care, industrial and so on.

Figure 5 shows the flow of obtaining an image contour extraction. After wavelet transformation is applied to the hammering sound at concrete surface, the following processes are conducted to the binary image of time-frequency spectrum;

(1) Focus on a black (=1) part in the black or white (=0) text image
(2) Count the number of black (=1) squares around the focused black part
(3) Extract the contour by leaving the squares numbered from 4 to 6
(4) Start creating x-y coordinates (in the case of Fig. 5, N0 to point N15)

The x-y coordinates that form the contour line of the targeted image obtained by the above-mentioned process are described by a complex expression as shown in Fig. 6.

The m-th order complex PARCOR coefficient p_m is equivalent to a_m of the m-th order complex autoregressive coefficient. Therefore, as shown in Fig. 7, as an example where m = 5, complex PARCOR coefficients p_1, p_2, p_3, p_4 and p_5 are calculated when [1, 1], [2, 2], [3, 3], [4, 4] and [5, 5] are determined. As for the m-order complex autoregressive coefficient, in general, it is possible and noted to make the judgment accurate as the number of m increases.

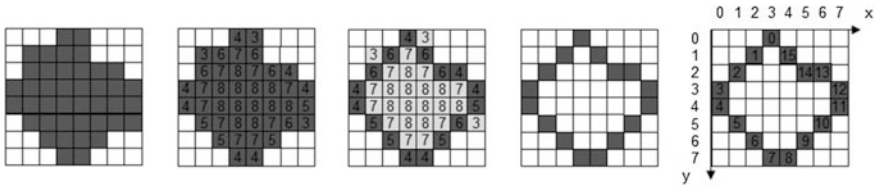


Fig. 5 Contour extraction of binary image

Fig. 6 Complex representations of x-y coordinates for contour line

	x	y			x	y
N0	3	0	→	N0 =	3 +	0 i
N1	2	1	→	N1 =	2 +	1 i
N2	1	2	→	N2 =	1 +	2 i
N3	0	3	→	N3 =	0 +	3 i
N4	0	4	→	N4 =	0 +	4 i
N5	1	5	→	N5 =	1 +	5 i
N6	2	6	→	N6 =	2 +	6 i
N7	3	7	→	N7 =	3 +	7 i
N8	4	7	→	N8 =	4 +	7 i
N9	5	6	→	N9 =	5 +	6 i
N10	6	5	→	N10 =	6 +	5 i
N11	7	4	→	N11 =	7 +	4 i
N12	7	3	→	N12 =	7 +	3 i
N13	6	2	→	N13 =	6 +	2 i
N14	5	2	→	N14 =	5 +	2 i
N15	4	1	→	N15 =	4 +	1 i

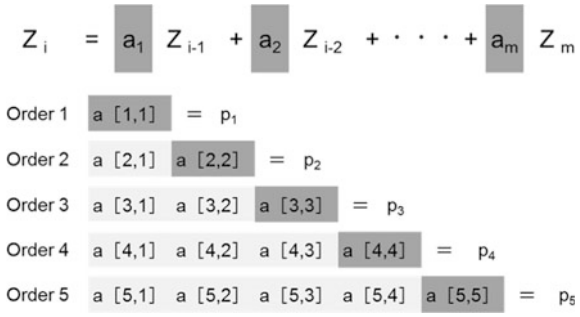


Fig. 7 Complex autoregressive coefficient and complex PARCOR coefficient

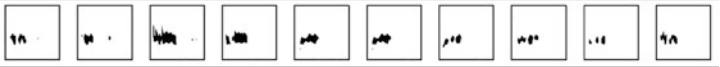
The complex PARCOR coefficient p_m afore-obtained is compared with the complex PARCOR coefficients of other images, and it is recognized as the nearest (the least error-induced) image. In this study, the average of the PARCOR coefficients of the m-th real part and imaginary part are used for the learning data, which are calculated as teacher images of 10 binarized images obtained for 10 hammering waveforms at concrete surface with interior defect. The complex PARCOR coefficients of the unlearned image obtained from three different hammering waveforms are tested, comparing with the complex PARCOR coefficient of the teacher image (learning data) to output the diagnosis result (defect depth) with minimum error sum of squares.

2.4 *Results on Pattern Analysis and Recognition of Hammering Sound*

Table 3 shows the binarized image (n = 10) with different defect depths as teacher images (learning data). They are respectively featured with only black color and originally the time-frequency spectrum by wavelet transformation as explained earlier. In the region of more than 10 kHz and 10,000 μs, the top 20% wavelet coefficient was not extracted. 10 years and more experienced technicians also manually tried to qualify the hammering sound and judge how deep the defects are embedded. However, they could not detect or make out the defects at depths of 50 mm or more. As also shown in Table 3, it is found that even the binarized images are similar when the depth of defect is more than 50 mm and it is difficult for skilled and experienced technicians to distinguish the difference of hammering sounds.

Figure 8 shows the diagnosis results for intact or defect depth using the unlearned data (n = 3, Test 1, 2 and 3) based on the learning data (n = 10, teacher images) with showing the error sum of errors on the complex PARCOR coefficients. The error of the complex PARCOR coefficients is minimum for any depth. It is indicated that the proposed pattern recognition algorithm gives accurate diagnosis results for defect depth under the concrete cover.

Table 3 Binary images of wavelet transform of hammering sound at different defect depth

Depth (mm)	Engineers' judge	Binary images
Intact	○ Easy	
100	× Difficult	
75	× Difficult	
50	△ Not clear	
30	○ Easy	
20	○ Easy	
10	○ Easy	

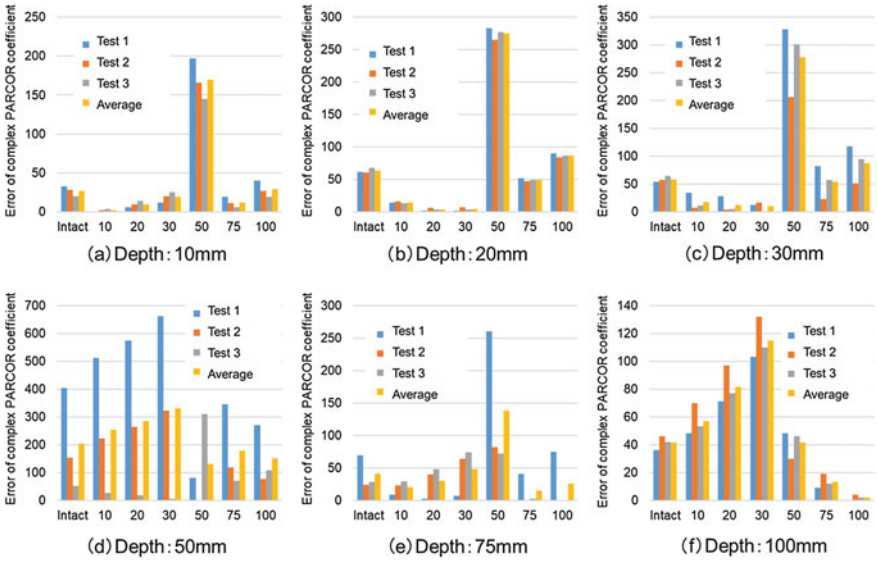


Fig. 8 Diagnosis results on complex PARCOR coefficients for defect depth

3 Autonomous Hammering Tools and System

In this study, an autonomous hammering test method is proposed. Conventionally some inspection technicians, who have much experience, access the targeted position of structures to manually hit the surface of the concrete. In the recent social situation where fewer workers are present at the inspection site, there is a need to have an autonomous hammering technique which does not require any skills and experience, so that anyone can technically complete the hammering inspection. Artificial intelligence installing the pattern recognition algorithm helps the concrete degradation diagnosis and 3D positioning system for hammering device with a microphone to collect a hammering sound enables the inspection work as shown in Fig. 9. Figure 10 shows the solenoid hammer, which uses a gyro sensor and accelerometer for 3D location, to keep the hit strength with a steel sphere, and an example of a hammering waveform.

As shown in Fig. 10, the 3D location path result for hammering device with carrying gyro sensor and accelerometer gives some errors approximately less than 10 cm for x-y. However, it is found that there is error accumulation with distance from original start point (time elapsed and longer measurement duration) as for x-z.

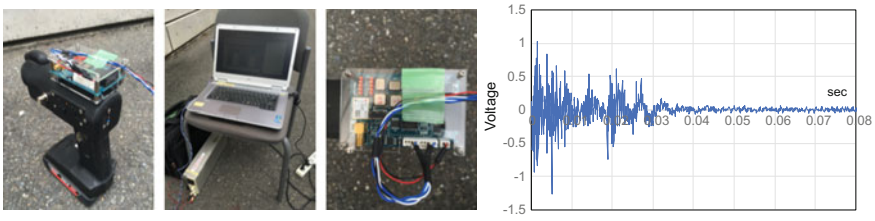


Fig. 9 Overview of autonomous hammering device with 3D location system

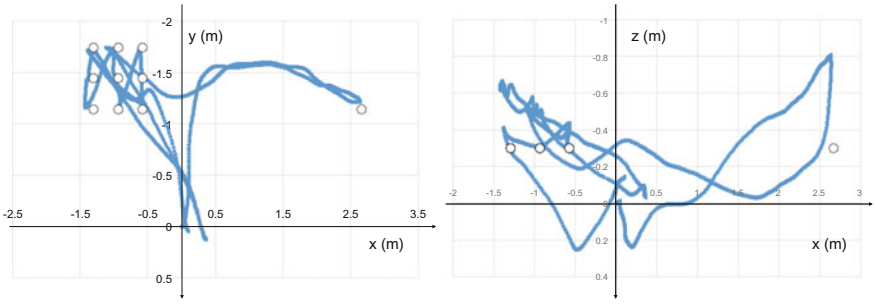


Fig. 10 3D location of hammering device with carrying gyro sensor and accelerometer

4 Conclusions

Defect depth diagnosis algorithm was built with binarization process on time-frequency spectrum by wavelet transformation of hammering sound waveform. It was shown that accurate judgment was clearly provided from the recognition technique for defect depth in concrete up to 100 mm. Also, autonomous hammering tools and system were introduced to establish the authors' further investigation for any existing structures. It is still needed to accumulate this data to improve the system. When the 3D positioning system and autonomous hammering device are integrated with the pattern recognition AI proposed in this study, inspection works for concrete structures will be drastically improved socially and economically.

Acknowledgements This research work has been supported by Association for Disaster Presentation Research, Japan and Kinki Kensetsu Kyokai (Construction Services in Kinki Region, Japan).

References

1. Asano M, Kamada T, Kunieda M, Rokugo K (2003) Frequency characteristic of impact acoustic waves in concrete slab with artificial defects. In 3rd Kumamoto international workshop on fracture, acoustic emission and NDE in concrete
2. Aggelis DG, Shiotani T, Kasai K (2008) Evaluation of grouting in tunnel lining using impact-echo. Tunnel Undergr Space Technol 23:629–639
3. Suzuki H, Kinjo T, Hayashi Y, Takemoto M, Ono K, Hayashi Y (1996) Wavelet transform of acoustic emission signals. J Acoust Emission 14(2):69–84
4. Sekita I, Kurita T, Otsu N (1992) Complex autoregressive model for shape recognition. IEEE Trans Pattern Anal Mach Intell 14(4):489–496

Predictive Maintenance as an Integral Part of Asset Life Cycle Maintenance Model

Md Mahdi Hassan, Carla Boehl and Mahinda Kuruppu

Abstract In a lean economy and stringent maintenance budget, it is necessary to implement a flexible and dynamic Asset Life Cycle Maintenance Model (ALCMM). Also, it is important to achieve the availability and reliability goals. Maintenance best practice entails a sequence of activities in a maintenance strategy-proactive maintenance, predictive maintenance, preventive maintenance and reactive maintenance. However, in the current mining economic downturn, this trend was interrupted due to maintenance budget cuts. In reality, this trend costs more money than following the best maintenance practice sequence. The author aims to depict a simple approach which integrates the predictive maintenance (using vibration analysis) with preventive maintenance targeting the critical equipment to maximise the equipment reliability and availability. The proposed approach will be implemented in three steps-working out the reliability of equipment using First Order Reliability Method (FORM), comparing with the preventive maintenance timeline and finally using the right logical decision to undertake the maintenance of equipment to avoid failures.

1 Introduction

Asset management has been defined as: "a strategic, integrated set of comprehensive processes (financial, management, engineering, operating and maintenance) to gain greatest lifetime effectiveness, utilisation and return from physical assets (production and operating equipment and structures)" [12].

M. M. Hassan (✉)
Western Australian School of Mines, Kalgoolie, Australia
e-mail: mdmahdi.hassan@postgrad.curtin.edu.au

M. M. Hassan · C. Boehl · M. Kuruppu
Curtin University, Perth, Australia
e-mail: Carla.boehl@gmail.com

M. Kuruppu
e-mail: M.Kuruppu@curtin.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_23

In the total time horizon of asset lifecycle, operation and maintenance phase claims the lengthiest period. Maintenance accounts for 40–60% of the total operating cost in mining industry, therefore, this area needs to get prime attention. One of the key areas to improve is identifying and prioritising work orders with a simple and effective model. This can be done by integrating predictive maintenance with preventive maintenance and by evaluating the effectiveness of the proposed maintenance model.

Evaluating maintenance effectiveness requires looking at the distribution of maintenance work hours within the facility [9]. Best practice benchmarks should be utilised to determine whether we are performing effective maintenance, but the following can be used as general guidelines:

- 60% preventive and predictive work (predominantly predictive is preferable)
- 30% corrective work as follow-up work from preventive and predictive maintenance activities
- 5% Proactive work
- <5% emergency maintenance work

Reliability analysis has been carried out using a statistical approach for more than 70 years to determine the uncertainty. Application of Monte Carlo Simulation (MCS) started in 1947. Because of its complexity and time-consuming nature, in some cases, MCS loses its advantage. Therefore, after 30 years, in 1974, Hasofer and Lind [7] introduced the First Order Reliability Method and they applied this method for structural problems. First Order Reliability Methods are well explained in Ditlevsen [4], Shinozuka [15], Chen and Lind [3], Ang and Tang [1], Der Kiureghian et al. [8], Wu and Wirsching [17], Melchers [11], Haldar and Mahadevan [6], Zhao and Ono [19], Seo and Kwak [14], Rahman and Xu [18], Beacher and Christian [2] and Low and Tang [10]. However, very limited work has been done to implement this method for mechanical equipment performance. First Order Reliability method needs very few iterations for convergence compared to other established methods namely Second Order Reliability Method and Monte Carlo Simulations, therefore FORM is regarded as a highly useful method to analyse the reliability of equipment. To implement the FORM in reliability analysis of equipment, a small modification is needed to current Low and Tang's FORM approach. Since, reliability analysis considering the state of machine vibration will be used in this investigation to prioritise the equipment for maintenance, a factor of safety must be considered while defining the limit state. Using vibration velocity as a random variable X1, comparing the random variable with its ultimate limit variable X2 which can be obtained from ISO 2372, and incorporating a factor of safety M, the objective function can be defined as:

$$g(X) = X2 - (X1 * M) \tag{1}$$

The new X1 variable will accommodate M value. Therefore, $g(X1, X2) > 0$ will be regarded as a safe zone and $g(X1, X2) < 0$ will be regarded as a failure zone.

The purpose of incorporating factor of safety is to avoid failure and getting enough time for maintenance for that equipment. The work order prioritisation process will be done by evaluating the reliability of the asset by comparing with the unacceptable level as well as with the unsatisfactory level of vibration. Whenever the data from ISO 2372 for unacceptable level is considered (i.e. higher allowable vibration velocity) the corresponding reliability of the asset is higher. When unsatisfactory level is considered (i.e. lower vibration velocity) the reliability of the asset is lower. The work order prioritisation should be given to those assets which have lower reliability. It should be noted that the best practice would be to prepare the maintenance planning for the assets which have poor reliability at unsatisfactory level of vibration.

2 First Order Reliability Method (FORM)

Reliability is defined as the probability of a performance function g(X) greater than zero, i.e. P{g(X) > 0} [5]. In other words, reliability is the probability that the random variables are in a safe region which is defined by g(X) > 0. The probability of failure is defined as that the random variables are in the failure region which is defined by g(X) < 0.

The name of the First Order Reliability Method (FORM) comes from function g (X) is approximated by the first order Taylor expansion (linearization).

The probability of failure is evaluated with the integral

$$\mathrm{pf} = \mathrm{P}\{\mathrm{g}(\mathrm{X}) < 0 = \int_{\mathrm{g}(\mathrm{x})<0} \int_{\mathrm{x}} (\mathrm{x})\mathrm{dx} \tag{2}$$

The reliability is computed by

$$\mathrm{R} = 1 - \mathrm{pf} \ = \mathrm{P}\{\mathrm{g}\ (\mathrm{X}) > 0\} = \int_{\mathrm{g}(\mathrm{x})<0} \int_{\mathrm{x}} (\mathrm{x})\mathrm{dx} \tag{3}$$

The probability integration in Eqs. 2 and 3 is the volume underneath the surface of the joint probability density function $\int_{\mathrm{x}} (\mathrm{x})$ of the random variable X, which forms a hill which is cut by a knife that has a blade shaped with the curve g(X) = 0.

The direct evaluation of the probability integration in Eqs. 2 and 3 is difficult because the variables are random and the probability integration is multidimensional. Therefore, to make the probability integration easier to compute, two steps will be involved in these approximation methods. The first step is to simplify the integrand $\int_{\mathrm{x}} (\mathrm{x})$ so that its contours become more regular and symmetric, and the second step is to approximate the integration boundary g(X) = 0. Following these steps, an analytical solution for the probability integration can be easily found.

2.1 Simplifying the Integrand

The simplification is achieved through transforming the random variables from their original random space into a standard normal space. The space that contains the original random variables X = (X1, X2, …, Xn) is called X-space. To make the shape of the integrand $\int_x (x)$ regular, all the random variables are transferred from X-space to a standard normal space, where the transformed random variables U = (U1, U2, …, Un) follow the standard normal distribution. The transformation from X to U is based on the condition that the cumulative density function (cdf) of the random variables remains the same before and after the transformation. This type of transformation is called Rosenblatt transformation [13], which is expressed by

$$F_{xi}(X_i) = \Phi(U_i) \tag{4}$$

In which Φ() is the cdf of the standard normal distribution.
The transformed standard normal variable is then given by

$$U_i = \Phi^{-1}[F_{xi}(X_i)] \tag{5}$$

For example, for a normally distributed random variable $X \approx N\ (\mu,\ \sigma)$. Equation 5 yields

$$\mathbf{U} = \mathbf{\Phi}^{-1}[\mathbf{F_{xi}}(\mathbf{X_i})] = \mathbf{\Phi}^{-1}\left[\mathbf{\Phi}\left(\frac{\mathbf{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}\right)\right] = \frac{\mathbf{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}} \tag{6}$$

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\sigma}\mathbf{U} \tag{7}$$

2.2 Approximation of the Integration Boundary

In order to further make the integration of probability easier to be evaluated, in addition to simplifying the shape of the integrand, the integration boundary g(U) will also be approximated using the first order Taylor expansion.

To minimise the accuracy loss, it is natural to expand the performance function g (U) at a point that has the highest contribution to the probability integration. In other words, it is preferable to expand the function at the point that has the highest value of the integrand, namely, the highest probability density. With the integration going away from the expansion point, the integrand function values will quickly diminish. The point that has the highest probability density on the performance g (U) = 0 is termed as the Most Probable Point (MPP). Therefore, the performance function will be approximated at the MPP. The MPP is the shortest distance point from the origin to the limit state. This minimum distance $\beta = ||u^*||$ is called reliability index also known as the Hasofer-Lind index which is equivalent to

$$\beta = min_{x \in F} \sqrt{\left(\left[\frac{x_i - \mu_i^N}{\sigma_i^N} \right]^T [\mathbf{R}]^{-1} \left[\frac{x_i - \mu_i^N}{\sigma_i^N} \right] \right)} \tag{8}$$

where, μ_i^N = equivalent normal mean, σ_i^N = equivalent normal standard distribution, R = correlation matrix, F = failure domain, x_i = random variable (vibration velocity).

Therefore, the procedure of the FORM is briefly summarised below:

i. Transforming the original random variables from X-space to U-space by Rosenblatt transformation.
ii. Searching the MPP in U-space and calculating the reliability index β
iii. Calculating the reliability R = Φ(β) which can be obtained using R = NORMDIST(-β)

2.3 Vibration Analysis

Most of the machine failures caused by vibration problems are due to fatigue [16]. The time taken to cause fatigue failure is determined by both how far an object experiences fatigue loading (stress) and the rate of deflection (strain). Using the principle of energy conservation it can be shown that both the fatigue stress and the strain are proportional to the vibration velocity.

In summary, the vibration velocity is a direct indicator of the likelihood of fatigue damage. The severity of vibration of industrial machinery and their acceptable limits depending on the size of machines are shown in ISO 2372. The advantages of measuring velocity are explained as given below:

i. Vibration velocity is a direct indicator of the severity of fatigue since it takes into account both stress and strain
ii. It is not necessary to know the frequency of vibration in order to evaluate the severity of vibration since frequency and velocity are related.
iii. A measurement of overall vibration velocity is a valid indicator of the overall condition of a machine whether the vibration is simple (single frequency) or complex (having more than 1 frequency).

For the reasons listed above, vibration velocity has become the industry standard for evaluating machinery condition based on vibration.

3 Results and Analysis

By using Hasofer and Lind reliability index and Low and Tang's algorithm, an analysis of vibration severity has been carried out in a ball mill with the specification of power = 4300 kW, RPM = 960 and capacity = 220 tons/h. This rotating machine is rigidly mounted on foundations; therefore it comes under Class III according to ISO 2372.

The vibration velocity data was collected over a period of 1 month and converted to mean and standard deviation as lognormal distribution Y. For the specification of the ball mill, the standard data range was obtained from ISO 2372 and processed as lognormal distribution Z. Again, a factor of safety has been considered and the data was processed and represented as the distribution Extreme value 1 of M.

Finally, a performance function is defined as:

$$g(X) = (Y * M - Z) = 0 \text{ which is the limit state where}$$

Y Real-time vibration velocity
M factor of safety
Z ultimate limit for the vibration velocity

Two scenarios were examined. Scenario 1 was based on considering vibration severity range as unsatisfactory (range from 6.3 to 15.67 mm/s which is 4.5 mm/s (RMS) to 11.2 mm/s (RMS)) as given in ISO 2372 and using the FORM the result was obtained. Similarly, scenario 2 was based on considering vibration severity range as unacceptable (range from 15.67 to 62.99 mm/s which is 11.2 mm/s (RMS) to 45 mm/s (RMS)) as given in ISO 2372 and again using the FORM the result was obtained.

3.1 Scenario 1

Comparing the vibration velocity with the ISO 2372 unsatisfactory range, it was found that the reliability is 69.91%. A summary of the calculations is given below:

Distribution		Mean	StDev
Lognormal	Y	1.3628	0.1052
Lognormal	Z	2.2168	0.9919
ExtValue1	M	1.2	0.2

x*	μ^N	σ^N	Correlation matrix [R]			n_x	g(x)	β
1.346	1.3587	0.1038	1	0	0	-0.122	-1E-09	0.5132
1.6234	1.981	0.6935	0.97	1	0	-0.516		
1.206	1.1657	0.1934	0	0	1	0.2087		

=YM - Z (x* values)

$(x^* - \mu^N)/\sigma^N$

Probability of failure 0.304
Reliability= 69.61 %

3.2 Scenario 2

Comparing the vibration velocity with the ISO 2372 unacceptable range, it was found that the reliability is 100%.

Distribution		Mean	StDev
Lognormal	Y	1.3628	0.1052
Lognormal	Z	6.8567	1.6837
ExtValue1	M	1.2	0.2

x*	μ^N	σ^N	Correlation matrix [R]			n_x	g(x)	β
1.282	1.3565	0.0988	1	0	0	-0.754	-2E-07	4.495
2.8344	5.2553	0.6858	0.997	1	0	-3.53		
2.211	0.5434	0.5319	0	0	1	3.1354		

=YM - Z (x* values)

$(x^* - \mu^N)/\sigma^N$

Probability of failure 3E-06
Reliability= 100 %

An algorithm was formulated to compute the reliability index, β as follows:

Step 1: Define the appropriate performance function g(X).
Step 2: Assume the initial value for the design point X^* as the mean value of the n − 1 random variable. For the last random variable, its value is obtained from the performance function to ensure that the design point is located on the limit surface g (X) = 0
Step 3: Obtain the design point in transformed reduced space, $U = \frac{X-\mu}{\sigma}$ or $X = \mu + \sigma U$
Step 4: Find the correlation matrix of X^*, μ^N, and σ^N
Step 5: Compute the reliability index as $\beta = \sqrt{\left[\frac{x_i-\mu_i^N}{\sigma_i^N}\right]^T [R]^{-1} \left[\frac{x_i-\mu_i^N}{\sigma_i^N}\right]}$
Step 6: Obtain the new design point $X^{*'}$ or (U′) for the n − 1 random variables

$$X^* = \mu + \sigma U'$$

Step 7: Determine the value of the last random variable in the original space such that the estimated point belongs to the limit state surface g(X) = 0
Step 8: Repeat steps 3–7 until convergence of β is reached.

The proposed model prioritises the work order based on the reliability achieved from the above two scenarios. When the result tends to scenario 1 but not reaching the scenario 1 result (69.91%) then the asset needs to undergo maintenance. Therefore, in this case, the benchmark for the ball mill maintenance work order can be established when the reliability decreases to 70% (but not less than 70%). The purpose of generating scenario 2 is to confirm that the asset is in a safe state with respect to the coarser standard level of vibration and using FORM.

4 Conclusion

FORM is a simple and efficient method to determine the reliability of any asset. Though Monte Carlo Simulation and Second Order Reliability Method give competitive results, considering sample size and number of iterations, FORM is superior to the other two. In the asset maintenance, very often a company needs to take a decision within a short period of time. FORM can deliver a quicker result in this occasion. Once the reliability of the asset is known using FORM and the preventive maintenance timeline is analysed (based on the information supplied by the Original Equipment Manufacturer (OEM)) then by comparing the two states as shown in the example the work order prioritisation can be determined.

References

1. Ang AHS, Tang WH (1984) Probability concepts in engineering planning and design. Wiley, New York
2. Baecher GB, Christian JT (2005) Reliability and statistics in geotechnical engineering. Wiley, New York
3. Chen X, Lind NC (1982) Fast probability integration by three-parameter normal tail approximation. Struct Saf 1(4):269–276
4. Ditlevsen O (1979) Generalized second moment reliability index. J Struct Mech 7(4):435–451
5. Du X, Chen W (2005) Collaborative reliability analysis under the framework of multidisciplinary systems design. Optim Eng 6(1):63–84
6. Haldar A, Mahadevan S (2000) Probability, reliability, and statistical methods in engineering design. Wiley, New York
7. Hasofer AM, Lind NC (1974) Exact and invariant second-moment code format. J Eng Mech Div 100(1):111–121
8. Der Kiureghian A, Lin HZ, Hwang SJ (1987) Second-order reliability approximations. J Eng Mech 113(8):1208–1225
9. Life Cycle Engineering (2013) [online] Available at: www.LCE.com
10. Low BK, Tang WH (2007) Efficient spreadsheet algorithm for first-order reliability method. J Eng Mech 133(12):1378–1387

11. Melchers RE (1987) Structural reliability: analysis and prediction. Wiley, New York
12. Mitchell JS, Carlson J (2001) Equipment asset management–what are the real requirements. Reliab Mag 4:14
13. Rosenblatt M (1952) Remarks on a multivariate transformation. Ann Math Stat 23(3):470–472
14. Seo HS, Kwak BM (2002) Efficient statistical tolerance analysis for general distributions using three-point information. Int J Prod Res 40(4):931–944
15. Shinozuka M (1983) Basic analysis of structural safety. J Struct Eng 109(3):721–740
16. Shreve DH (1994) Introduction to vibration technology. In Proceedings, predictive maintenance technology conference, annual meeting
17. Wu YT, Wirsching PH (1987) New algorithm for structural reliability estimation. J Eng Mech 113(9):1319–1336
18. Xu H, Rahman S (2004) A generalized dimension-reduction method for multidimensional integration in stochastic mechanics. Int J Numer Meth Eng 61(12):1992–2019
19. Zhao YG, Ono T (2001) Moment methods for structural reliability. Struct Saf 23(1):47–75

Adapting Infrastructure to Climate Change: Who Bears the Risk and Responsibility?

Samantha Hayes

Abstract The projected effects of climate change on infrastructure assets create a driver for climate adaptation measures in infrastructure design, construction and operation. To support effective adaptation approaches, it is important that roles and responsibilities of stakeholders are understood and clearly communicated. Synthesising academic, industry and government literature, this review identifies climate risks for a selection of infrastructure stakeholders, including designers, constructors, planners and regulators, operators, investors and insurers. It provides examples of global best practice adaptation responses and proposes several adaptation considerations for each of these stakeholders moving forward. The paper serves as a quick reference guide for practitioners and decision-makers.

1 Introduction

Existing infrastructure assets have been designed for a tolerance to extreme weather events which may not suffice under future climate pressure, where altered rainfall patterns, extreme weather and rising sea levels are expected to lead to increased operation and maintenance costs, reduced lifespan and high cost retrofit and repairs [12, 8]. In Australia and internationally, climate vulnerability assessments and strategic adaptation planning will be pivotal to managing vulnerability of infrastructure networks. To facilitate this action and move beyond ad hoc, champion-based initiatives, it is necessary to clearly identify major risk areas and adaptation strategies available to stakeholders throughout the infrastructure asset lifecycle. This investigation and synthesis of risks, current best practice and priority

This document is a heavily abridged version of a research project undertaken at University of Sydney as a component of the Master of Environmental Law. Please contact the author for further details.

S. Hayes (✉)
Griffith University, Brisbane, Australia
e-mail: samantha.hayes3@griffithuni.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_24

areas for future action seeks to contribute to ongoing efforts to clarify roles and responsibilities for stakeholders.

2 Climate Risks in the Infrastructure Asset Lifecycle

Table 1 and the ensuing analysis outline climate risks for key stakeholders in the infrastructure asset lifecycle. The risks listed here reflect those commonly cited as significant in the industry, government and academic literature.

Table 1 Climate change risks in the infrastructure asset lifecycle

Stakeholder	Climate change risks
Planning and regulators	Cost of rebuilding Legal liability for failing to consider climate impacts in planning Mal-adaptation Operations, maintenance and repair costs
Investors	Potential for asset failure or project delays Litigation of key parties (i.e. investees)
Investees	Litigation Access to capital due to increasing insurance premiums and risk
Designers	Legal liability for failing to consider climate risk Inconsistent application of adaptation approaches across projects Mal-adaptation
Constructors	Program delays Supply chain exposure Productivity impacts Impacts of altered weather patterns or extreme events on asset longevity and resilience (impact on guarantees and defects liability) Mal-adaptation
Operators and maintenance	Forced retirement of assets Repair and maintenance requirements Disruption of operations Accelerated degradation of materials, structures and foundations Increased insurance scope and premiums Increased investment risk Mal-adaptation Litigation
Insurers	Inability to insure in high-risk locations Inability to deliver on premiums (where risk has not been adequately considered/mitigated) Failure to provide pricing signals Lack of coverage Potential to dis-incentivise adaptation

Planning and regulatory risk: Government agencies face numerous challenges when seeking to establish regulatory mechanisms that support adaptation without fostering inefficiency or maladaptation. Competing stakeholder priorities and resource constraints can impede progress in introducing new regulation, and where implemented without bipartisan support, the threat of future repeal remains. While Local Government actors are often best placed to implement adaptation measures and stand to be directly impacted by inoperable infrastructure [21], limited funding, inaccessibility of detailed local datasets and a lack of clarity around roles and responsibilities each constitute tangible barriers to climate adaptation efforts at the local government level [21, 24].

Investment risk (investors and investees): Infrastructure investors face several risks from anthropogenic climate change, particularly where extreme weather events and more gradual shifts in climatic patterns impact the commercial viability of the project during both construction and operation, including cash flow continuity over time [3]. Construction delays, as well as the inability to generate revenue during the operations phase are obvious examples. Further to this, the risk of legal claims against any of the key parties involved in project poses a threat to cash flow and financial viability—a risk heightened by the emerging body of case law in regards to climate resilience in infrastructure.

Legal action to assert design or construction negligence for failure to consider climate impacts has targeted designers, constructors, planners and operators, citing their failure to adequately recognise, mitigate or adapt to such impacts in infrastructure assets [e.g. 13, 15]. Such cases have had limited success to date, however they indicate a growing focus on the allocation of liability for failure to incorporate climate risk into infrastructure projects—an issue which will become more pertinent as the impacts of climate on infrastructure continue to be realised over time. Investees, on the other hand, are likely to encounter increased scrutiny of their climate risk management, resulting in increased insurance premiums and potential restrictions on access to capital where climate risk management does not satisfy investor expectations.

Design risk: Design risks can also be significant. Selection of materials, site location and alignment, height and durability of the asset are all elements that may expose designers to litigation risk should the design prove insufficient under altered climate conditions. The inconsistent application of adaptation approaches across a range of infrastructure assets, the potential for mal-adaptation or utilisation of inaccurate or outdated climate data; or a failure to alert clients to known climate risks may all provide potential triggers for liability and reputational damage. Depending on contractual arrangements, including detailed specifications, defect, and warranty obligations, these effects may extend well beyond the design and construction phases and into the commissioning, operations and maintenance phases of the asset lifecycle [2, 16, 14].

Construction risk: The potential effects of climatic change on construction are becoming increasing clear [3]. Climate risks for contractors (including an increase in wet weather days and extreme heat days), may affect productivity, scheduling and construction costs. In addition to natural disaster impacts, increased rainfall,

flooding and sea level rise will alter erosion patterns during both the construction and operations phases, and the intrusion of salt water into freshwater reserves as a result of sea level rise may reduce the availability of water for infrastructure construction projects [1, 2]. Extreme weather and changes to rainfall and temperature patterns are also likely to impact on the supply chain through disruptions to production and transport of supplies domestically and internationally. The global supply chain repercussions of natural disasters have become evident in recent years, affecting a broad range of sectors, products and services. Such global fluctuations have the potential to significantly influence supply and demand of construction materials, making supply chain risk mapping increasingly important [32].

Operations and maintenance risk: In operation, risks include unforeseen increases in operating costs, asset degradation and importantly, deficiencies in other interconnected infrastructure assets, which may limit the ability of the asset to operate effectively [3]. Many of the physical risks faced by infrastructure assets will be borne during the operations and maintenance phase, ranging from flooding, extreme winds and storm events, to accelerated degradation of materials and structures [2, 14].

Insurance risk: Insurance plays an important role in managing compensation for damages resulting from extreme weather events. Moving forward, however, such risks could become uninsurable as limits to insurance uptake and magnitude of claims lead to an inability to deliver on pay-outs in the aftermath of disaster [19].

Historically, natural disaster events have largely been considered *force majeure*, however increasing scientific certainty around anthropogenic climate changes and the availability of climate modelling and forecast impacts mean that failure to adequately account for extreme weather events during construction and operation will increasingly be viewed as an avoidable failure by key players in the infrastructure asset lifecycle. The absence of legislated requirements for local governments to insure council infrastructure, and limitations on the availability of infrastructure insurance products within the market may also contribute to vulnerability and under-insurance [21].

3 Current Status and Recommendations

Within Australia and internationally, stakeholders are taking action to address these risks throughout the infrastructure asset lifecycle. Table 2 provides examples of action taken by both the private and public sectors to respond to and manage climate risks. Further investigation of the effectiveness and drivers of these actions will support greater stakeholder understanding of available strategies at each stage of the asset lifecycle.

These adaptation examples include both legislated and voluntary efforts and establish a catalogue of adaptation strategies for infrastructure stakeholders moving forward. Table 3 synthesises the findings on climate risks and current practice to propose recommended priority areas for key stakeholders moving forward.

Table 2 Examples of adaptation actions throughout the asset lifecycle

Stakeholder	Adaptation actions
Planning and regulators [12, 21, 24]	Regulatory frameworks mandating adaptation measures Consideration of climate risk in planning instruments Development of climate adaptation plans—local and state government Development of state government portfolios and committees to address climate risk and disaster resilience Information and guidance on project risk and adaptation options Local government associations—collaboration for consistency Development of toolkits and models for adaptation projects Requirements for private sector risk and adaptation planning and reporting.
Investors [5, 9]	Enhanced disclosure of triple bottom line risk Divestment of assets and investments with high inherent climate risk Incorporation of climate risk into investment decisions and analysis
Designers and Constructors [4, 6, 7, 11, 16, 17, 18, 25, 26, 30]	National and international standards Industry standards and rating schemes to provide consistency in determining best practice Assessment of climate risk across property portfolios Assessment of climate risk and opportunities in construction of infrastructure assets Development of 'adaptive' adaptation in design, which allows for progressive adaptation over time based on actual extent and rate of climate impacts Early incorporation of adaptation into international development programs Multi-purpose infrastructure assets which offer dual transport and disaster mitigations functions
Operators and Maintenance [4, 7, 10, 16, 17, 29, 31]	Australian and international standards Industry standards and rating schemes to facilitate best practice Assessment of climate risk across portfolios
Insurers [20, 22, 23, 28]	Pricing of insurance products to encourage climate risk management Provision of climate risk data and modelling Development of risk assessment tools, frameworks and guidance Legislating for public liability and professional indemnity insurance for local governments Mutual liability schemes available to local councils

Source Compiled by author with references as noted

Table 3 Moving forward—adaptation considerations at each phase of the asset lifecycle

Stakeholder	Adaptation considerations
Planning and regulators	State and local government required to have documented response to climate risks and approach to adaptation Government agencies adopt a network approach to vulnerability assessments as opposed to individual assets Major projects required to assess vulnerability and identify adaptation options Key major private sector participants required to periodically report on measures taken to assess and reduce climate risk in their operations
Investors	Investors to expand current due diligence processes to include assessment of climate risk and adaptation measures Investor indices to strengthen focus on adaptation If introduced, private sector reporting on adaptation measure would provide greater transparency into investment risk associated with climate change impacts
Designers	Engineering and design guidelines amended to more explicitly address climate risk and adaptive design Government specifications mandate use of these standards/guidelines
Constructors	Broader uptake of AAS5334 and rating schemes Amendment of construction standards to incorporate and allow for adaptive construction methods and materials selection to enhance resilience
Operators and maintenance	Broader uptake of AAS5334 and rating schemes Assessment of climate vulnerability across portfolios Cost-benefit analysis of soft and hard adaptation measures Adjustment of budgeted maintenance regimes to accommodate impacts on operations, structure integrity and material longevity
Insurers	Parties seeking insurance conduct vulnerability assessments to inform insurance choices and determine where adaptation measures should be introduced in lieu of/in conjunction with insurance mechanisms Utilisation of best available forecast data to inform policy instruments Insurers to assess climate risk mitigation in the determination of policy payout Potential for mandatory insurance requirements

Planning and regulation: Planning and regulation mechanisms will be integral to ensuring the consistent application of adaptation measures across the infrastructure asset lifecycle. Regulatory bodies should consider climate vulnerability of infrastructure networks in zoning, master planning and the location of essential services and emergency responses, including the incorporation of climate risk assessment into Federal and State Government approval and impact assessment processes. Infrastructure projects should be located in minimal risk zones where possible, and infrastructure standards and specifications amended in order to increase the integration of vulnerability assessment and adaptation. At the Local Government level, climate change considerations can be integrated into operation, maintenance and repair strategies, with climate adaptation incorporated into the

Local Government Acts for each State or Territory. Local Government can endeavour to manage adaptation in line with the AS5334-2013 standard to ensure that infrastructure resilience to climate risk is consistently incorporated into planning, development, approvals and ongoing management of infrastructure assets at the local level. To ensure consistency, requirements should be established and agreed to in collaboration with relevant national bodies. Roles and responsibilities for adaptation should be clearly articulated and communicated across all levels of government.

The national adaptation strategy can provide a strong and consistent guiding framework for adaptation efforts nationally, including consideration of appropriate regulatory and funding mechanisms. Regulatory approaches could include a streamlined mechanism requiring organisations to report on their approach to climate risk assessment and adaptation. Coverage could be restricted to major infrastructure projects and other key assets, essentially providing a more consistent and transparent mechanism than the existing asset-level contract specifications which vary in scope, transparency and robustness. This would also allow government to identify barriers to adaptation, recognise best practice and evaluate the status of adaptation measures nationally. Should this not occur, however, government agencies acting as project proponent, client and operator, are well placed to incorporate requirements for vulnerability assessment and adaptation evaluation into the contractual arrangements and project specifications, as well as adjusting material, design and engineering specifications. This provides government with a practical avenue for catalysing private sector adaptation in the absence of a broader regulatory driver.

Investment: The inclusion of climate risk and adaptation considerations within investor due diligence processes will drive a greater focus on reducing infrastructure vulnerability. Such considerations should assess whether climate adaptation is sufficiently integrated into business processes to provide confidence that the business and its assets are resilient and sustainable in the face of climate change. Integration of such analysis into investment decisions supports consistent, uniform evaluation of adaption approaches across investment portfolios. Just as investor rating schemes have requested details on company approaches to climate change mitigation, a growing focus will now be placed on the assessment of climate change adaptation, risk assessment and preparedness. Climate adaptation and resilience to climate risk should become an integrated component of the due diligence process for acquisition of existing infrastructure assets moving forward. The Investor Group on Climate Change (IGCC) and similar associations will likely play an important role in facilitating this change.

Design: Design standards should evolve to reflect and accommodate changing weather and climate patterns, as opposed to relying on historical datasets and traditional approaches. Given the long lifespan of bridge and causeway structures, particular attention should be paid to adaptation measures for these structures. Arguably, these changes will occur naturally as a result of more widespread climate risk analysis throughout the sector. As such, it is the consistent implementation of

vulnerability assessment in the design phase that will be key to climate-proofing infrastructure design.

Construction: Similarly, to minimise risk of programme and non-recoverable cost overruns during the construction phase, it will be necessary for construction contractors to utilise forward projections, as opposed to historical data only, in their estimation of wet weather days, flood impacts and extreme heat days. The adjustment of construction standards, particularly the consideration of climate risk in materials selection, will also support ongoing adaptation. The AS 5334 Standard provides a useful guiding framework for climate adaptation approaches in infrastructure, and can be adopted by all parties in the infrastructure asset lifecycle. Industry rating and performance frameworks, such as the Infrastructure Sustainability (IS) rating tool, also play an important role in facilitating adaptation in infrastructure, and could be more broadly adopted across the sector, whether in a formal or self-assessment approach.

Operations and maintenance: Climate vulnerability assessments and adaptation plans for the operations and maintenance phase can be facilitated by uptake of the IS Rating Tool in conjunction with AS 5334 to allow comprehensive climate risk assessment and evaluation of adaptation actions. Armed with this information, operators may identify 'soft' adaptation measures and 'hard' retrofit opportunities, as well as developing action plans to minimise the disruptions caused by extreme weather events. Maintenance regimes may be adjusted to account for accelerated weathering, structural degradation and reduced longevity. Preventative maintenance approaches will be integral to these responses.

Insurance: Throughout the asset lifecycle, climate risk and vulnerability assessments can be used to inform insurance strategies for infrastructure, improving understanding of risk profiles and enabling stakeholders to adequately reflect risk and insurance requirements within project pricing and contract conditions [3]. This also allows for identification of risks which may be considered uninsurable or excessive, as well as catalysing strategies to reduce exposure. Insurers may introduce requirements that insured parties have taken reasonable steps to avoid an insurable loss, including climate risk mitigation and emergency response planning [14]. An increasing focus on the availability and use of comprehensive climate risk data will offer numerous benefits, informing insurers in their development of insurance instruments and risk modelling, as well as providing a valuable resource for other stakeholders seeking to better understand and manage their climate risk profile.

4 Conclusion

Given the scale of projected climate change risks to infrastructure, it is important that stakeholders develop efficient and effective strategies for reducing vulnerability. Government, industry and academic literature offers insight into key risks faced by stakeholders throughout the asset lifecycle, as well as examples of their

approaches to dealing with those risks. Synthesising current approaches, it is evident that a broad spectrum of adaptation options exist, ranging from minimum regulatory requirements to world leading voluntary strategies. Here, these efforts have been consolidated into a series of key considerations for major infrastructure stakeholders, which move beyond current practice yet seek to remain feasible and practical for widespread adoption. If adopted consistently, these approaches would signify a major shift in understanding and management of infrastructure vulnerability to climate change.

References

1. Australian Green Infrastructure Council (Now ISCA) (2011) AGIC guidelines for climate change adaptation, AGIC
2. Austroads (2004) Impact of climate change on road infrastructure, Austroads
3. Baartz J, Longley N (2003) Construction and infrastructure projects—risk management through insurance, Allens Arthur Robinson
4. CEEQUAL (2014) CEEQUAL: improving sustainability, CEEQUAL
5. Climate Disclosure Standards Board (CDSB) (2012) Climate change reporting framework edition 1.1: promoting and advancing climate change related disclosure, CDSB
6. Colonial First State (2012) Climate change adaptation at Brisbane Airport. Case Study: New Parallel Runway Project, CBA
7. Council of Standards Australia (2013) 'AS5334-2013 climate change adaptation for settlements and infrastructure—a risk based approach. Council of Standard Australia
8. Cunite (2011) Climate change: science and solutions for Australia. CSIRO, Collingwood
9. Dow Jones Sustainability Index (DJSI) (2014) DJSI 2014 review results, Dow Jones Sustainability Index
10. Economic Commission for Europe (2013) Climate change impacts and adaptation for international transport networks (expert group report ECE/Trans/238, Geneva: UNECE, pp 59, 72, 87
11. Engineers Australia (2013) Guidelines for responding to the effects of climate change in coastal and ocean engineering, Engineers Australia, p 11
12. Garnaut R (2011) The garnaut climate change review 2011. Cambridge University Press, Port Melbourne
13. Gerrard MB (2012) Hurricane katrina decision highlights liability for decaying infrastructure. Law Journal 1(90):247. New York
14. Holper P, Lucy S, Nolan M, Senese C, Hennessy K (2007) Infrastructure and climate change risk assessment for Victoria: consultancy report to the Victorian Government prepared by CSIRO. CSIRO, Maunsell AECOM and Phillips Fox
15. Illinois Farmers Insurance Company et al v. The Metropolitan Water Reclamation District of Greater Chicago District, et al., (Ill, 14 CV 03251, June 3, 2014)
16. Infrastructure Sustainability Council of Australia (2016) Infrastructure sustainability rating tool: technical manual version 1.2, ISCA
17. Institute for Sustainable Infrastructure and the Zofnass Program (2012) The envision rating system, ISI and Zofnass
18. Lal PN, Thurairajah V (2011) Making informed adaptation choices: a case study on climate proofing road infrastructure in the Solomon Islands, IUCN
19. Linnerooth-Bayer J, Hochrainer-Stigler S (2014) Financial instruments for disaster risk management and climate change adaptation. Clim Change 133(1):85–100

20. Local Government Association of South Australia (LGASA) (2012) Local government South Australia climate change adaptation programme: final report, LGASA, Adelaide, 17 pp
21. Mukheibir P, Kuruppu N, Gero A, Herriman J (2013) Cross-scale barriers to climate change adaptation in local government. Australia, NCCARF
22. Munich RE (2014) Natural disasters 2013, NatCatSERVICE
23. Munich RE (2014) Loss events worldwide 2013: geographical overview, Munich RE
24. Productivity Commission (2012) Barriers to effective climate change adaptation (Inquiry Report No. 59), Productivity Commission, Melbourne
25. Ranger N, Reeder T, Lowe J (2013) Addressing 'deep' uncertainty over long-term climate in major infrastructure projects: four innovations of the Thames Estuary 2100 Project. EURO J Decis Process 1:233–262
26. Royal Academy of Engineering (2011) Infrastructure, engineering and climate change adaptation—ensuring services in an uncertain future. RAE, London
27. Stockland (2013) FY13 sustainability report. Stockland, Sydney
28. Swiss Re (2012) Flood—an underestimated risk: inspect, inform, insure. Mthenquai, Swiss Re, pp 11–16, 18, 21
29. Transurban (2014) Sustainability report 2014: climate change. Transurban
30. UK Climate Impacts Program (2011) Lend lease—approach to climate change adaptation, UKCIP
31. Water Sydney (2010) Climate change adaptation: program summary. Sydney Water, Sydney
32. Ye L, Abe M (2012) The impacts of natural disasters on global supply chains (Working Paper No.115, Asia-Pacific Research and Training Network on Trade, June 2012) 16

Calculation and Analysis of Anti-shock of a Marine Diesel Engine Turbocharger

Hu Lei, Yang Jianguo, Zheng Mingchao and Yu Yonghua

Abstract Anti-shock characteristics is one of the main technical indices of a turbocharger for a marine diesel engine. In order to study the anti-shock capability of the turbocharger, a nonlinear finite element model of the turbocharger considering the oil film is established in this paper. The shock loads of the turbocharger in the form of the shock spectrum are calculated and analysed as the anti-shock conditions. Furthermore, the shock spectrum is converted into the time-domain acceleration loads and applied as the external load to the nonlinear finite element model. Finally non-dimensional indices of the structural stress and the deformation of the turbocharger is obtained by simulating based on the time-domain shock loads, and then the anti-shock characteristics of the turbocharger are evaluated.

1 Introduction

A marine diesel engine is a power heart of a ship, the turbocharger with precise structure and high reliability is one of the key components of the marine diesel engine, so it is important to study anti-shock characteristics of the turbocharger. Methods used to study of the anti-shock characteristics mainly include the shock test method, the calculation and analysis method. The former can be simulated or tested in the actual shock environment, and the analysis and the assessment of the anti-shock more realistic and effective, but it requires a lot of the cost and the time. The latter the finite element model can be established to calculate the shock

H. Lei (✉) · Y. Jianguo · Z. Mingchao · Y. Yonghua
School of Energy and Power Engineering, Wuhan University of Technology, Wuhan 430063, China
e-mail: lgtc0616@aliyun.com

H. Lei · Y. Jianguo (✉)
Key Laboratory of Marine Power Engineering and Technology, Ministry of Communication PRC, Wuhan 430063, China
e-mail: jgyang@whut.edu.cn

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_25

response and to assess the anti-shock ability according to the relevant shock norms and standards [4, 11, 12].

There are three main methods to calculate the anti-shock, which are Static Equivalent Method (SEM), Dynamic Design Analysis Method (DDAM) and Real-time Simulation Method (RSM) [5]. SEM is that the dynamic load is equivalent to a certain static load, and then the static load is used to check the strength. Early relevant reports abroad put forward "the equal static acceleration of the shock movement" to the calculation method of the anti-shock for mechanical and electrical equipment, However, the shock environment of this method is weak, only has a certain reference on designing simple structure [6, 8]. SEM can obtain the first-order low frequency response, but can't analyse the high frequency breaking response of the equipment. The DDAM converts the n-degree-of-freedom system into the n single-degree-of-freedom systems based on the modal analysis theory. The equipment movement is the three-fold line spectrum of the displacement, the velocity and the acceleration changing with the frequency ratio. DDAM can be used to analyse the high frequency failure mode of the equipment [3], but its shock spectrum is mainly applied to the linear dynamic response. Real-time Simulation Method is a complement of DDMA, which converts the input shock into the time-domain load to calculate its transient response. O. Vinogradov and others proposed the non-linear shock response calculation of the shock spectrum considering viscous damping model. FENG Linhan also took the supercharged boiler as the object, and calculated and analysed the nonlinear shock response by taking the gap and damping into account [5, 10].

2 Calculation and Analysis of Anti-shock of Turbocharger

2.1 Analysis of Three-Fold Line Shock Spectrum

The shock spectrum is the relation curve between the maximum response amplitude and its natural frequency of a series of massless linear oscillators with the different natural frequencies, which are supposed to be subjected to the same transient excitation based on installing on the same shocked equipment [9]. At present, the shock environment of marine equipment refers to German Navy Construction Code BV043-85. The three-fold line shock spectrum can be represented as the shock load [1], the low frequency band is the constant displacement spectrum d_0, the middle frequency band is the constant velocity spectrum v_0, the high frequency band is the constant acceleration spectrum a_0, where: $f_1 = v_0/(2\pi d_0)$, $f_2 = a_0/(2\pi v_0)$, as shown in Fig. 1.

The acceleration and the velocity can be limited by DDAM shock spectrum curve and the relationship among the velocity, the acceleration and the displacement is also certain. The designed shock spectrum of DDAM can be extended to the three-fold line shock spectrum by constructing the constant—displacement

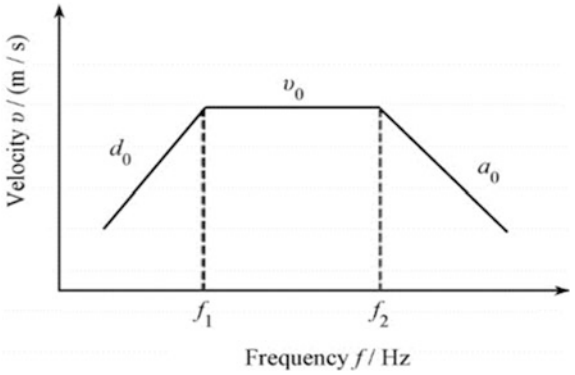


Fig. 1 Shock spectrum of three-fold line (logarithmic coordinates)

spectrum, the velocity spectrum and the acceleration spectrum. The spectral value is related to the ship type, the equipment installation location and the modal quality of each mode of the equipment. Therefore, it is necessary to establish the finite element model of the turbocharger at first.

2.2 Establishment of Finite Element Model of Turbocharger

The lubricating oil film of the turbocharger bearing has great influence on the dynamic characteristics of the rotor shaft. The oil film with damping and stiffness characteristics controls the vibration response of the rotor shaft. So it is necessary to consider the oil film effect of the bearing when calculating the anti-shock characteristic of the turbocharger. The oil film mainly affects the radial stiffness and damping of the rotor shaft, but has little influence on the circumferential stiffness and damping, generally not be considered [2]. While the perturbation is small, the Taylor series expansion of the perturbation parameters can be obtained to simplify the analysis of the oil film force of the oil film transferred to the journal, retain first-order available [13]:

$$\begin{cases} F_x = F_{x0} + \frac{\partial F_x}{\partial x}\Big|_0 \Delta x + \frac{\partial F_x}{\partial y}\Big|_0 \Delta y + \frac{\partial F_x}{\partial \dot{x}}\Big|_0 \Delta \dot{x} + \frac{\partial F_x}{\partial \dot{y}}\Big|_0 \Delta \dot{y} \\ F_y = F_{y0} + \frac{\partial F_y}{\partial x}\Big|_0 \Delta x + \frac{\partial F_y}{\partial y}\Big|_0 \Delta y + \frac{\partial F_y}{\partial \dot{x}}\Big|_0 \Delta \dot{x} + \frac{\partial F_y}{\partial \dot{y}}\Big|_0 \Delta \dot{y} \end{cases} \quad (1)$$

The oil film force is approximated as a function of the minor displacement and the velocity of the journal. In the formula, F_x and F_y are the components of the oil film force in X and Y direction, F_{x0} and F_{y0} are the components of the oil film force at the static equilibrium position in X and Y directions. Definite the oil film force increment induced by the unit displacement as the oil film stiffness coefficient, that is $k_{xx} = \frac{\partial F_x}{\partial x}\Big|_0$, $k_{xy} = \frac{\partial F_x}{\partial y}\Big|_0$, $k_{yx} = \frac{\partial F_y}{\partial x}\Big|_0$, $k_{yy} = \frac{\partial F_y}{\partial y}\Big|_0$; Definite the oil film force

increment induced by the unit velocity as the oil film damping coefficient, that is $c_{xx} = \frac{\partial F_x}{\partial \dot{x}}$, $c_{xy} = \frac{\partial F_x}{\partial \dot{y}}$, $c_{yx} = \frac{\partial F_y}{\partial \dot{x}}$, $c_{yy} = \frac{\partial F_y}{\partial \dot{y}}$, so the mechanical model of the bearing oil film stiffness damping is shown in Fig. 2.

The dynamic characteristic coefficients of the bearing oil film are related to the lubricating oil inlet temperature, the viscosity, the structure size and the speed and so on. Suppose the lubricating oil absorbs 80% friction heat, and the operating temperature of the bearing can be obtained by using the thermal equilibrium relationship [7].

$$T_{\text{oilinworking}} = T_{\text{oilsupplied}} + 0.8\frac{P}{C_V Q} = T_{\text{oilsupplied}} + 0.8\frac{\eta\Omega}{C_V}\left(\frac{R}{C}\right)^2 4\pi\frac{\bar{P}}{\bar{Q}} \tag{2}$$

The maximum temperature, T_{max}, of the oil film takes place in the carrying liner, which is

$$T_{\max} = T_{\text{oilinworking}} + \Delta T = T_{\text{oilinworking}} + \frac{\eta\,\Omega}{C_V}\left(\frac{R}{C}\right)\bar{T} \tag{3}$$

Dimensionless edge flow $\bar{Q}$, dimensionless friction power dissipation $\bar{P}$ and dimensionless temperature rise $\bar{T}$ of liner.

Expression of S, $\bar{Q}$, $\bar{P}$, $\bar{T}$ is as follows:

$$\begin{aligned} S &= \frac{\eta NDL}{W}\left(\frac{R}{C}\right)^2;\ \bar{Q} = Q\Big/\left(\frac{1}{2}\pi NDLC\right) \\ \bar{P} &= PC/(\pi^3\eta N^2 LD^3);\ \bar{T} = \Delta T\Big/\frac{\eta\Omega}{C_V}\left(\frac{R}{C}\right)^2 \end{aligned} \tag{4}$$

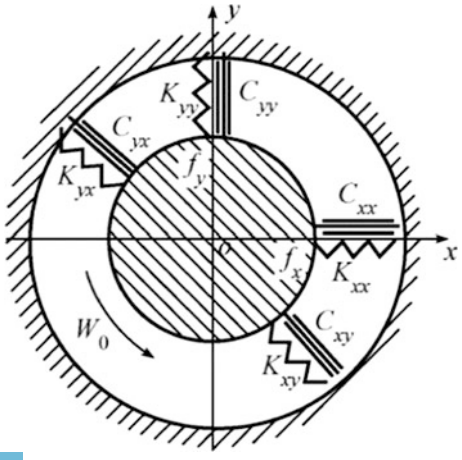


Fig. 2 Mechanical model of bearing oil film stiffness damping

where: D = 2R for journal diameter, N = Ω/2π for the journal rotation speed; L/D for the ratio of length to diameter of the journal; C_v for the specific heat of lubricating oil of unit volume.

The viscosity value can be found out from the oil temperature-viscosity curve, and the Moffield number S calculated. Because the carrying liners play a leading role on bearing characteristics, T_{max} was used as the temperature of choosing the lubricating oil viscosity. If T_{max} is different from the initially estimated temperature, a new value needs to be re-estimated and then calculated until the two temperatures are substantially in agreement. The dynamic coefficients are obtained by the interpolation method after obtaining S value [13].

According to the actual operating conditions of the turbocharger from manufacturer, speed n = 52,000 r/m, the lubricating oil type is CD40, the allowing oil inlet temperature is 40–0 °C, the operating temperature is about 78 °C, the setting lubricating oil operating temperature is about 80 °C. By calculating and verifying above method, the dynamic characteristics of the oil film are shown in Table 1. So the finite element model of the turbocharger is established as shown in Fig. 3.

Table 1 Oil film dynamic characteristic coefficient

Characteristic coefficient	S	K_{xx}	K_{xy}	K_{yx}	K_{yy}	C_{xx}	$C_{xy} = C_{yx}$	C_{yy}
Sliding bearings	0.628	240.0	114.3	−508.2	354.2	0.062	−0.049	0.176
Floating sleeves	2.835	381.3	639.0	−1426.4	320.1	0.244	−0.073	0.520

Note The unit of stiffness coefficient is N/mm, the unit of damping coefficient is N*s/mm

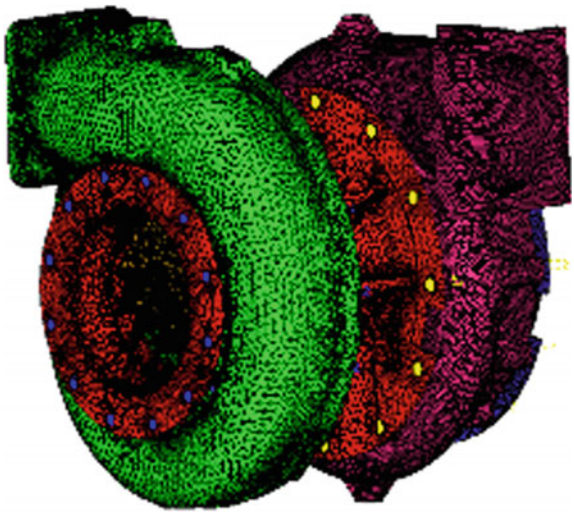


Fig. 3 Finite element model of the turbocharger

2.3 Calculation of Three-Fold Line Shock Spectrum

The shock acceleration value (A_a) and the shock speed value (V_a) of the naval vessels equipment are designed by the installation location. Blades and other parts of turbocharger are not allowed to produce permanent deformation after being shocked, and the elastic design criterion is selected [7].

The total mass of the model used in the modal analysis is 0.157t. The *X*, *Y* and *Z* directions of the finite element modelling of the established turbocharger are respectively corresponding to the transverse (port and starboard), the vertical and the longitudinal (fore and aft) direction of the hull. Minimum value of both $V_a\omega_a$ and A_a is taken as designed shock acceleration of the dynamic analysis system of the equipment in the given shock direction. The mechanical model with n degrees of the freedom for a given direction of the impact needs to analyse a sufficient number of vibration modes N to ensure that the total modal mass is not less than 80% of the total mass of the analysis system. The mode of analysis should include all modes with a modal mass greater than 10% of the total mass of the analysis system, and a lower frequency mode should be a priority [7] accumulated values of parameters for the constrained modal are shown in Figs. 4, 5 and 6.

The total modal mass exceeds 80% of the total mass of the turbocharger when *X* direction is 33 steps, *Y* direction is 44 steps, and *Z* direction is 52 steps by the result constrained modal calculation of the turbocharger. The three-fold line shock spectrum in the three directions can be obtained and shown in Table 2.

2.4 Determination of Impact Load in Time-Domain

The three-fold line shock spectrum is the frequency-domain load. Due to the influence of non-linear factor such as the oil film in the turbocharger, the frequency-domain load should be converted into the time-domain load, and

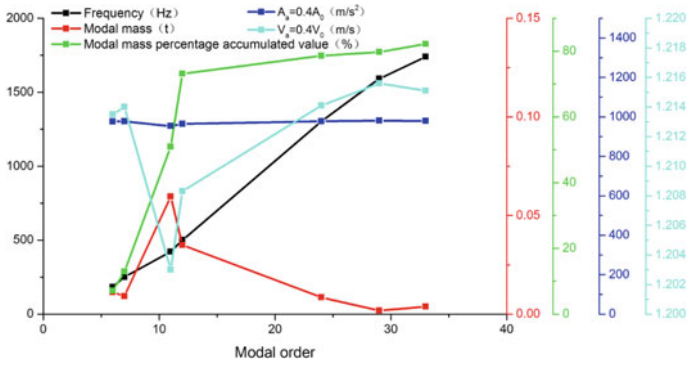


Fig. 4 X (transverse) direction modal parameters of turbocharger

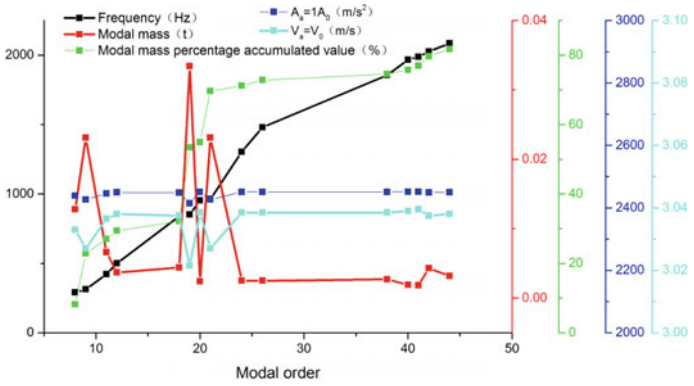


Fig. 5 Y (vertical) direction modal parameters of turbocharger

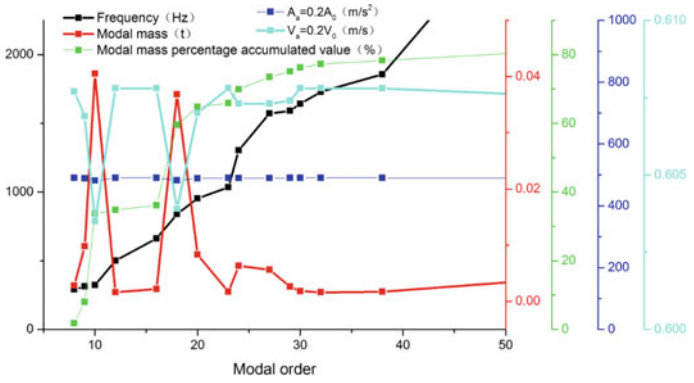


Fig. 6 Z (longitudinal) direction modal parameters of turbocharger

Table 2 The parameters of the three-fold line shock spectrum of turbocharger

Direction	A_a (m/s^2)	V_a (m/s)	D_a (m)
X (Transverse)	980	1.21	0.047
Y (Vertical)	2450	3.04	0.047
Z (Portrait)	490	0.61	0.047

the time-domain loading method for the design of impact spectrum into positive and negative triangular wave is shown in Fig. 7 [3].

The transformation formula is as follows:

$$\begin{aligned} &a_2 = 0.6A_a; V_2 = 0.75V_a; t_3 = 2V_2/a_2; t_2 = 0.4t_3; \\ &t_5 = t_3 + (6.3D_a - 1.6a_2t_3^2)/1.6a_2t_3; a_4 = -a_2t_3/(t_5 - t_3) \\ &t_4 = t_3 + 0.6(t_5 - t_3) \end{aligned} \tag{5}$$

Parameters of the three directions calculated are shown in Table 3.

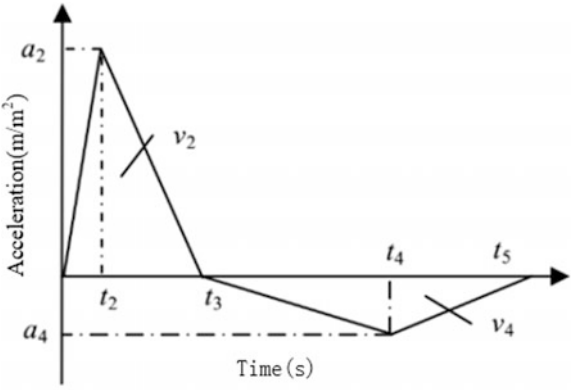


Fig. 7 Time domain load of positive and negative triangular wave

Table 3 Load parameters of three-fold line

Direction	a_2 (m/s^2)	t_2 (ms)	t_3 (ms)	t_4 (ms)	a_4 (m/s^2)	t_5 (ms)
X	588	1.23	3.08	62.50	−18.27	102.20
Y	1470	1.24	3.1	25.60	−121.52	40.6
Z	294	1.24	3.11	122.44	−4.60	202

3 Anti-shock Assessment

3.1 Anti-shock Assessment of Strength

For the equipment which the anti-shock is class-A, The maximum value of the shock response should not exceed the static yield limit for the materials of the parts. Define "failure factor" n as follows:

$$n = \frac{\sigma_{Mises}}{\sigma_s} \tag{6}$$

where, σ_{Mises} is the peak stress of the assessment, σ_s is the static yield limit. From the "failure factor" expression, it can be found that if n is less than 1, structural of part is safe, and also invalid.

By calculating and analysing the turbocharger stress shock response at maximum peak of the vertical, the horizontal and the vertical shocks respectively, Table 4 shows the maximum stress shock response values of the main components of the turbocharger respectively in calculating the vertical, the horizontal and the longitudinal. As can be seen from the table, the turbocharger components failure factors n are less than 1, which meet the standard requirements.

Table 4 Mises stress analysis of the critical components under the shock load

Number	Component names	Failure coefficient n		
		Y-direction	X-direction	Y-direction
1	Compressor shell	0.209	0.148	0.112
2	Diffuser	0.198	0.146	0.114
3	Nozzle ring	0.240	0.176	0.135
4	Rotor shaft	0.013	0.007	0.006
5	Gas inlet shell	0.199	0.148	0.115

3.2 Anti-shock Assessment of Shell and Blade Clearance

If the static clearance between the compressor or the turbine and leaf blade is d_1, in order to ensure the operational safety of the turbocharger when shocking, the relative displacement U_{rmax} between the shell and leaf blade should not exceed d_1, Define "failure coefficient" β as follows:

$$\beta = \frac{U_{r\max}}{d_1} \tag{7}$$

When β is larger than 1, the shell and the blade will not be collided, so it is safe, and then β will be bigger, indicating that the bigger value β is, the stronger the anti-shock of the assessment is; when β is greater than 1, the shell and the blade interfere with each other, even it leads to damage.

The turbocharger displacement shock response at the maximum peak of the vertical, the horizontal and vertical shocks respectively are calculated and analysed. The actual clearance of the compressor shell and the corresponding blade leaf is 0.79 mm, and the actual clearance of the turbine shell and the corresponding blade leaf is 0.89 mm, the representative points are shown in Table 5. From the data in Table 5, it can be seen that the failure coefficients β are less than 1, and results are the same between turbine shell and the corresponding blade leaf.

Table 5 Anti-shock analysis of compressor and turbine casing and blade tip clearance

Check node number	β Failure coefficient					
	Clearance between compressor shell and corresponding blade leaf			Clearance between turbine shell and corresponding blade leaf		
	Y-direction	X-direction	Z-direction	Y-direction	X-direction	Z-direction
1	0.900	0.372	0.120	0.865	0.231	0.072
2	0.897	0.372	0.119	0.850	0.229	0.071
3	0.897	0.372	0.118	0.864	0.228	0.077
4	0.895	0.371	0.116	0.878	0.215	0.079

4 Conclusion

1. The non-linear finite element model of turbocharger is established. The time-domain acceleration load in three-fold line is calculated and converted to the frequency-domain. Real-time Simulation Method is suited to analyse and evaluate the shock resistance of the turbocharger.
2. The results of the anti-shock calculation show that the stress response of the turbocharger is the greatest impact in the vertical shock, so the shock resistance design of the turbocharger should be meet the vertical requirements. According to the results of the calculation and the analysis, the strength of the turbocharger and the clearances between the shell and the leaf blade can meet the standard load requirements. Moreover, from the aspects of the structure and the deformation, the compressor shell is more susceptible to fail than the turbine shell.

References

1. BV0430-85 (1998) Shock safety: translation of the code for the construction of the federal defense forces
2. Shuqian C, Qian D, Yushu C (1999) Analysis on modeling steady rotor system with sliding bearings by using FEM. Turb Techn 41(6):347–350
3. Carter TL (1989) Correlation of the DDAM and transient shock analysis methods using the MK 45 LFSP shock test results. Shock Vibr Bull 60(1):205–234
4. Hai-long C, Xiong-liang Y, A-Man Z (2009) Shock resistance performance evaluation for the typical dynamic device used in ships. J Vibr Shock 28(2):45–50
5. FENG Linhan (2008) Research on anti-shock of characteristic of supercharged boiled on board
6. GENG Panpan (2012) Numerical simulation on shock resistance analysis of marine gas turbine
7. GJB1060.1-91 (1991) Requirements of ship environmental condition—mechanical environment
8. Tao J, Wei-li W, Xue-feng H (2010) Analysis and application of SRS from plus-minus triangular wave in marine equipment shock-resistant. J Naval Aeronaut Astronaut Univ 25 (2):45–48
9. LI Qun (2006) Studies on shock environment for equipments on surface ship and anti-shock characteristics of structures
10. Vinogradov O, Pivovarow I (1986) Vibrations of a system with non-linear hysteresis. J Sound Vibr 111:145–152
11. Wang L (2010) Research on shock-resistance of typical equipments interacted with ship structure
12. Wang Y (2005) Theory and application of modern ship shock
13. Zhong Y, Yanzhong HE, Wang Z (1987) Rotor dynamics

Analysis of Flexural Vibration of V-Shaped Beam Immersed in Viscous Fluids

Lu Hu, Wen-Ming Zhang, Han Yan and Hong-Xiang Zou

Abstract In this paper, the flexural vibration of the V-shaped beam with varying section and bending stiffness submerged in a quiescent viscous fluid is studied. Fluid-structure interactions are modelled through a complex hydrodynamic function that describes added mass and damping effects. A parametric study is conducted on the two-dimensional computational fluid dynamics of an oscillating V-shaped beam under harmonic base excitation to establish a handleable formula for the hydrodynamic function, which is in terms of the frequency parameter and the ratio between the gap of the beam cross-section and the width of the beam. The real part and imaginary parts of hydrodynamic function decrease with both the frequency parameter and the ratio. The frequency response of the V-shaped beam is obtained by solving the model numerically. Findings from this work can be used in the design of V-shaped structures of interest in marine applications such as the energy harvesting devices and micromechanical oscillators for sensing and actuation.

1 Introduction

The study of flexural vibrations of flexible plates submerged in viscous fluids is of fundamental importance in many research fields, such as atomic force microscopy [1], microscale energy harvesting [2] and biomimetic propulsion [3, 4]. Especially, rectangular and V-shaped cantilevers are commercially available and are most

L. Hu · W.-M. Zhang (✉) · H. Yan · H.-X. Zou
State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China
e-mail: wenmingz@sjtu.edu.cn

L. Hu
e-mail: hl_hulu@sjtu.edu.cn

H. Yan
e-mail: yanhan_mail@foxmail.com

H.-X. Zou
e-mail: zouhongxiang@sjtu.edu.cn

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_26

commonly used. A V-shaped cantilever can reduce lateral twisting compared to the rectangular cantilever [5] and V-shaped actuators have distinct advantages over the other commonly used folded-beam actuators [6]. In most research, the submerged structures undergoing forced oscillations are sharp-edged beams [7, 8]. Different from the analysis of sharp-edged beam vibrating in viscous fluid, the section and bending stiffness of V-shaped beam are variable and the hydrodynamic function depends not only on a frequency parameter, but also on the gap to width ratio of the beam cross-section. In this work, the hydrodynamic function of the local frequency parameter and the ratio between the gap and width of the V-shaped beam in cross-section is proposed and the frequency response of the beam is obtained.

2 Problem Statement

The oscillation of a slender beam of V-shape that is immersed in a quiescent viscous fluid is considered in this work. The geometry of the problem is shown in Fig. 1 along with pertinent nomenclature. A local $\{x, y, z\}$ coordinate system is attached to the beam, where the x coordinate is along the beam axis, the y and z coordinates are along the beam width and thickness directions. The beam has length L, width $b(x) = b_0(1 - x/L)$, thickness h and $\gamma(x)$ is the ratio between the gap of the beam cross-section and the width of the beam. The beam bending stiffness and mass per unit length are defined as $K(x)$ and $m_s(x)$. When $0 < x < l_1$, the width of the beam cross-section is $2w_d$, $K(x) = Ew_d h^3/6$, and $m_s(x) = 2\rho_s w_d h$, with E the Young's modulus of the beam material, and with ρ_s the mass density of the beam material per unit volume. When $l_1 < x < L$, $K(x) = Eb(x)h^3/12$ and $m_s(x) = \rho_s b(x)h$. Additionally, $w(x,t)$ and $W(t)$ are used to indicate the beam elastic deflection and the externally imposed base excitation.

Under the hypothesis of the classical Euler-Bernoulli beam theory, the equation of motion is given by

$$\frac{\partial^2}{\partial x^2}\left[K(x)\frac{\partial^2 w(x,t)}{\partial x^2}\right] + m_s(x)\frac{\partial^2}{\partial t^2}[w(x,t) + W(t)] = H(x,t) + s(x,t) \tag{1}$$

The imposed base excitation is given by the displacement of the clamped end $W(t) = W_0(\omega)\sin(\omega t)$, where $W_0(\omega)$ and ω represent the amplitude and the radian frequency of excitation. $H(x,t)$ describes the hydrodynamic loading exerted on the beam by the encompassing fluid and $s(x,t)$ represents the structural damping of the beam. In the frequency domain, Eq. 1 can be rewritten as

$$\begin{aligned}\frac{1+\mathbf{i}\eta}{m_s(x)}\frac{\partial^2}{\partial x^2}\left[K(x)\frac{\partial^2 \hat{w}(x,\omega)}{\partial x^2}\right] - \omega^2[W_0(\omega) + \hat{w}(x,\omega)] \\ = \frac{\pi\rho_f\omega^2 b(x)^2}{4m_s(x)}\Theta(\beta(x,\omega),\gamma(x))[W_0(\omega) + \hat{w}(x,\omega)]\end{aligned} \tag{2}$$

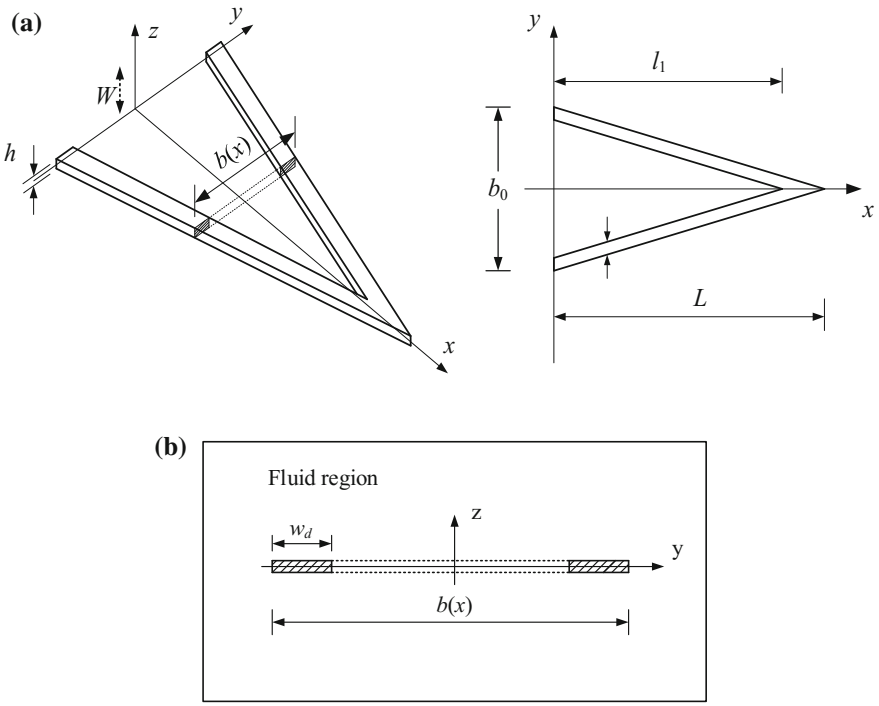


Fig. 1 **a** Nomenclature for the problem and **b** sketch of the associated 2D fluid problem

where, Θ is the complex hydrodynamic function which is in terms of the frequency parameter $\beta(x,\omega)=\rho_f \omega b(x)^2/(2\pi\mu)$ and the gap to width ratio $\gamma(x)$. η is a frequency-independent hysteretic damping.

3 Hydrodynamics

The hydrodynamic function $\Theta(\beta,\gamma)$ is estimated by studying the 2D problem of a rigid lamina which is divided into two parts by the gap (the ratio between the gap and the length of the lamina represents γ) oscillating in a quiescent viscous fluid. The lamina is regarded as a moving wall which harmonically oscillates in the z-direction with infinitely small vibration amplitude A_0 at the radian frequency ω. At any instant of time, the z-position z_C of its centroid is updated according to the time law $z_C(t) = A_0 \sin(\omega t)$. Additionally, the fluid is considered as common liquid water at room temperature for which $\rho_f = 998.2\,\text{kg m}^{-3}$ and $\mu = 1.003 \times 10^{-3}\,\text{Pa s}$.

The 2D Navier-Stokes equations governing the fluid velocity field $\mathbf{u}(y,z,t)$ and the pressure field $p(y,z,t)$ are given by,

$$\mathrm{div}\mathbf{u}(y,z,t) = 0 \tag{3}$$

$$\rho_f \frac{\mathrm{D}\mathbf{u}(y,z,t)}{\mathrm{D}t} = -\nabla p(y,z,t) + \mu\nabla^2\mathbf{u}(y,z,t) \tag{4}$$

The parametric study focuses on the ranges $\beta \in [50, 1000]$, $\gamma \in [0.1, 0.7]$. The range of the frequency parameter is included completely in the range $\beta \in [5, 5000]$, which is proved to be acceptable [7]. And the range of the gap to width ratio can exactly describe the geometry of the V-shaped beam. Simulations are run for about 3.3 cycles to extract velocity, pressure as well as the force along the z-direction per unit length in the x-direction exerted by the encompassing fluid on the oscillating rigid lamina.

Figures 2 and 3 show the representative velocity magnitude, pressure in the x-direction for the case of $\beta = 100$ and $\gamma = 0.3$.

In Figs. 2 and 3, snapshots are taken at time $t = nT/10$, with $n = 0, 1, \ldots, 9$ from left to right in the panel. Two black rectangles in every snapshot represent two parts of the rigid lamina. Here, T is the oscillation period, the value of T is 0.32 s. The size of the frame is about 1.75×1.75 lamina lengths. Figure 2 illustrates the evolution of the velocity fields during the oscillation. In the case shown, the highest velocity magnitude is concentrated in the neighbourhood of the lamina tips. And when the lamina crosses the $z_C = 0$ position with its maximum velocity, maximum value in the velocity field is attained. Figure 3 displays the evolution of the pressure fields during the oscillation. When the velocity of the lamina decreases, relatively large negative pressure is observed in the vicinity of the lamina edges. And relatively large negative and positive pressure regions are found upstream and downstream the motion of the lamina.

Besides the velocity and pressure fields in the computational domain, the CFD analysis provides the time history of the force along the z-direction which is the sought hydrodynamic force per unit length exerted on the vibrating lamina by the encompassing fluid. Figure 4 shows time histories of force and displacement in the simulation for $\beta = 100$, $\gamma = 0.3$. For each set of governing parameters, the force output $F(t)$ from the simulations is fitted with a harmonic signal of the type $F(t) = F_0 \sin(\omega t + \psi)$. Where F_0 is the force amplitude and ψ represents the phase shift between the force and the displacement.

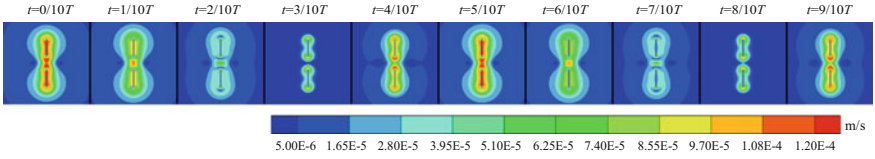


Fig. 2 Contours of velocity magnitude for $\beta = 100$, $\gamma = 0.3$

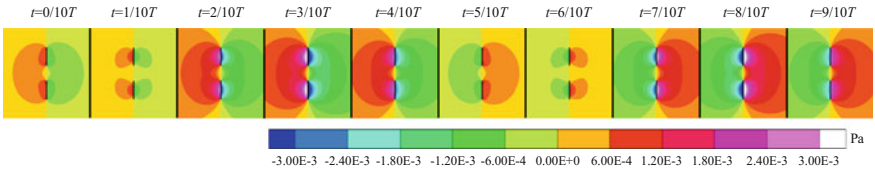


Fig. 3 Contours of pressure for $\beta = 100$, $\gamma = 0.3$

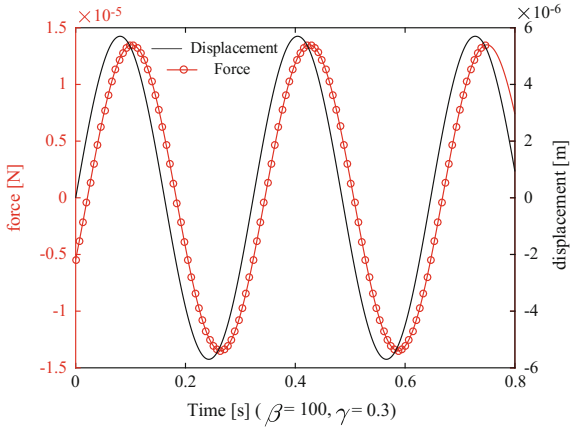


Fig. 4 Time histories of force and displacement in the simulation for $\beta = 100$, $\gamma = 0.3$

The hydrodynamic function $\Theta(\beta, \gamma)$ is consistently defined by following Eq. 5,

$$F_0 e^{i\psi} = \frac{\pi}{4}\rho_f \omega^2 b^2 \Theta(\beta, \gamma) A_0 \tag{5}$$

where A_0 is the oscillation amplitude.

Figure 5 shows the real and imaginary parts of the hydrodynamic function $\Theta(\beta, \gamma)$. The figure reports that $\mathrm{Re}[\Theta]$ and $-\mathrm{Im}[\Theta]$ decrease with both β and γ.

4 Results

After obtaining the formula of the hydrodynamic function $\Theta(\beta(x, \omega), \gamma(x))$, the frequency response of the V-shaped beam can be numerically solved by using Galerkin method. The parameters of the beam in Eq. 2 are $L = 0.05\,\mathrm{m}$, $l_1 = 0.037\,\mathrm{m}$, $h = 0.0001\,\mathrm{m}$, $E = 3.8\,\mathrm{GPa}$, $b_0 = 0.015385\,\mathrm{m}$, $w_d = 0.002\,\mathrm{m}$, $\rho_s = 1.23 \times 10^3\,\mathrm{kg\,m^{-3}}$. In the numerical implementation, the infinite summations are truncated to $m = 5$ modes of vibration. Figure 6 shows the frequency response of the tip of the beam.

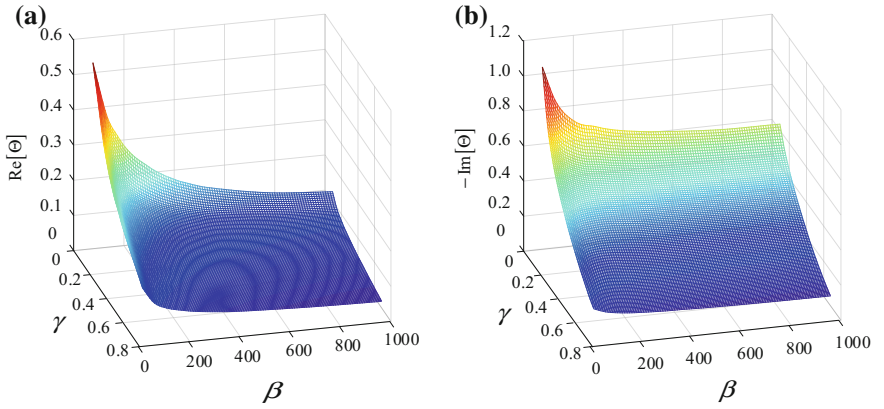


Fig. 5 **a** Real part and **b** imaginary part of hydrodynamic function $\Theta(\beta, \gamma)$ as obtained from the simulations

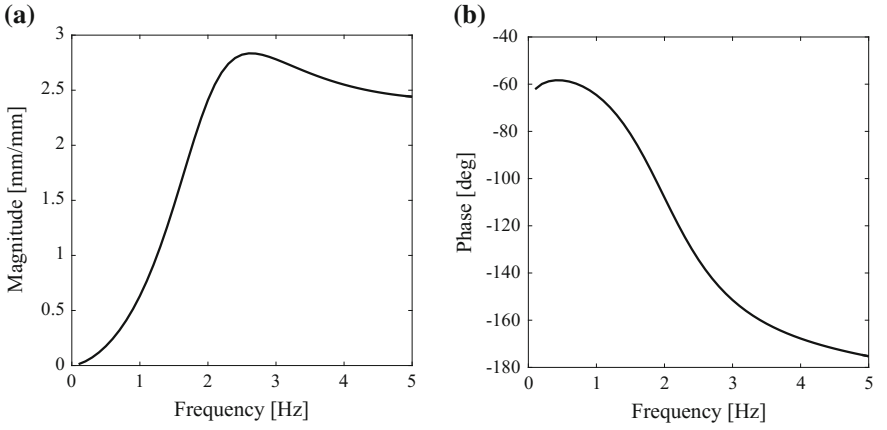


Fig. 6 Frequency response of the tip of the beam. **a** Magnitude of the elastic tip amplitude to the base amplitude and **b** phase of the beam under base excitation

5 Conclusion

In this paper, we developed a model to analyse and predict the vibration of the V-shaped beams vibrating under harmonic base excitation within an encompassing fluid. A hydrodynamic function depending on the oscillatory frequency parameter and the gap to width ratio of the beam cross-section was introduced to model the fluid-structure interaction. A formula for the hydrodynamic function is identified based on the results of a comprehensive 2D parametric CFD simulation campaign integrated by the analysis of unsteady Stokes hydrodynamics. And the frequency response of the V-shaped beam is obtained by solving the model numerically.

Further work will be devoted to the experimental validation of the problem and the exploration of the influence of the local Keulegan-Carpenter number on this problem.

Acknowledgements The authors gratefully acknowledge the supports by the National Natural Science Foundation of China (Grant No. 11572190), the National Science Foundation for Distinguished Young Scholars (Grant No. 11625208), and the National Program for Support of Top-Notch Young Professionals.

References

1. Raman A, Melcher J, Tung R (2008) Cantilever dynamics in atomic force microscopy. Nano Today 3(1):20–27
2. Cha Y, Verotti M, Walcott H, Peterson SD, Porfiri M (2013) Energy harvesting from the tail beating of a carangiform swimmer using ionic polymer–metal composites. Bioinspir Biomim 8 (3):036003
3. Aureli M, Kopman V, Porfiri M (2010) Free-locomotion of underwater vehicles actuated by ionic polymer metal composites. IEEE/ASME Trans Mechatron 15(4):603–614
4. Liu B, Powers TR, Breuer KS (2011) Force-free swimming of a model helical flagellum in viscoelastic fluids. Proc Natl Acad Sci 108(49):19516–19520
5. Lee HL, Chang WJ (2011) Sensitivity of V-shaped atomic force microscope cantilevers based on a modified couple stress theory. Microelectron Eng 88(11):3214–3218
6. Enikov ET, Kedar SS, Lazarov KV (2005) Analytical model for analysis and design of V-shaped thermal microactuators. J Microelectromech Syst 14(4):788–798
7. Aureli M, Basaran ME, Porfiri M (2012) Nonlinear finite amplitude vibrations of sharp-edged beams in viscous fluids. J Sound Vib 331(7):1624–1654
8. Grimaldi E, Porfiri M, Soria L (2012) Finite amplitude vibrations of a sharp-edged beam immersed in a viscous fluid near a solid surface. J Appl Phys 112(10):104907

Partners in Maintenance—Possibilities in Using Partnering-Based Maintenance Contracts for Swedish Railway

Anders Ingwald and Mirka Kans

Abstract Traditionally, procurement of maintenance has been based on fixed price on defined work. These types of contract forms can lead to negotiations on everything not included in the price, which may result in delays and a sense of distrust when all parties primarily look after themselves. Recently, forms for cooperation based on mutual trust and fairness for all parties have been developed; partnering. Partnering is characterized by common objectives and project organization, and risk analysis. The project is created and operated in cooperation with customers, contractors, suppliers and consultants. Transparency is a cornerstone of partnering and openness is what mainly differentiates partnering from traditional contractor projects. The Swedish railway uses traditional and very detailed contracting principles. This leads to several problems, such as lack of information and knowledge transfer between parties and life cycle phases, and no incentives for quality and productivity improvements. Is applying the concept of partnering for railway maintenance a solution to these problems? If so, why is the concept not already commonly applied and what are the barriers? This paper introduces the concept of partnering and reviews current research within the area. Thereafter, the possibilities in using partnering-based maintenance in Swedish railway are discussed.

1 Introduction

Procurement of maintenance has traditionally been based on fixed price on defined work, also for complex cooperation projects. These types of contract forms are inflexible and can easily lead to negotiations on everything that is not included in

A. Ingwald (✉)
Department of Physics and Electrical Engineering, Linnaeus University, Växjö, Sweden
e-mail: anders.ingwald@lnu.se

M. Kans
Department of Mechanical Engineering, Linnaeus University, Växjö, Sweden
e-mail: mirka.kans@lnu.se

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_27

the price, which may result in delays and a sense of distrust when all parties primarily look after themselves, Olsson [17]. Furthermore, because of the rigorous definitions in the agreements, there are little if no incentives for improving neither maintenance internal effectiveness nor its performance; see for instance Lingegård [13] and Abdi et al. [1]. The problem of optimizing maintenance effectiveness with regard to deciding proper maintenance intervals, the proper maintenance policy etc. has been extensively researched, see for example Sherwin [18], and several models for optimizing maintenance effectiveness have been developed. However, most models are developed for the situation where an organization also include maintenance resources and do not consider the more complex situation where the maintenance function is outsourced. For this kind of situation other models have been developed during the last twenty years. These models have different focus and purposes see for example Murthy and Asgharizadeh [14] where the profit of the contractor is optimized, and Godoy et al. [8], where the goal is to motivate continuous improvements. Still, many of the models developed for outsourced maintenance are based on a rather traditional view of maintenance, as a cost and service. This means that maintenance is used in a passive manner. A different view, where maintenance, is described as something contributing to value creation is described in, for example, Stenström et al. [20]. One concept that goes beyond the traditional seller-buyer relation is partnering where the problem of effective maintenance is supposed to become a common goal for the supplier and the buyer. However, so far, the partnering concept has not been used to a great extent in maintenance. One area, where traditional contracts for maintenance outsourcing are used and that currently is facing great problems related to these traditional contracts, is the Swedish railway, Lingegård [13], Abdi et al. [1], Ingwald and Kans [9]. The purpose of this paper is to introduce the concept of partnering as a feasible way for the effective procurement of maintenance. The paper is based on a structured literature review within the area of partnering. Research on partnering is found mainly in project management, supply chain management and construction literature but only a limited amount of publications is found within the area of maintenance. The experiences and lessons learned from previous studies will be utilized for expressing possibilities in using partnering-based maintenance in Swedish railway.

2 The Concept of Partnering

The concept of partnering has been discussed and used order to improve effectiveness and promote improvements in outsourcing, especially in the construction industry. Given that the concept of partnering is still relatively new, especially in some industries, there are still many areas related to partnering that need further research, Bresnen [3]. Formulating common goals and then combining resources, knowledge and competence from the contractor(s) and the customer in order to reach them is the core of partnering. However, even if partnering requires a certain degree of equality, it is primarily a client-contractor relationship. There is currently

no standardized definition of partnering. The concept is described as a dyadic relation, i.e. a contractor-client relation, but also as a multi-actor relation, see Bygballe et al. [4]. The duration spans from covering a specific project to strategic and long-term relationships. Partnering is about creating efficient multi-partner teams, social relationships and trust; see for example Kadefors [10] and Olsson and Esping [16], so activities that enable trust amongst involved people are important. Openness and access to information are crucial, and Bresnen [3] states that information technology often is seen as an enabler, especially at an early stage in the partnering relation. Kadefors [10] reports on several risks related to economic incentives, when used in partnering, since there is a risk of focusing on self-interest instead of what is beneficial for the partnering relation. Straub [19]describes a performance-based partnership for maintenance, and concludes that this method promise benefits compared to more traditional methods. Olsson and Esping [16] state that there must be financial incentives for all partners involved.

Vilkamo and Keil [24] report on lessons learned from studying technology partnering relationships in high-velocity environments, i.e. environments with rapid and discontinuous change in demand, competition, technology, and regulation. Even though the area of study differs from this paper, there are several similarities in the relationships built up between companies in terms of commitment and risk sharing. The authors conclude that the management of technology partnering in dynamic environments differs from traditional strategic alliances, which often are driven by efficiency through minimizing costs or maximizing outcomes. In contrast, a dynamic environment requires dynamics in the management too. The alliance should be *integrated into the technology strategy* (adequate level of competencies is needed on both sides), be *managed for flexibility* rather than for efficiency, reach a *balance between exploration and exploitation* (short-term benefits and long-term relationships), and be able to *manage different time paces* (different product life cycles) and *time scales* (short-term and lifecycle) simultaneously. Results from a cross-industry study in the telecom area are presented in Dodourova [6], including success and failure factors for company collaboration. Amongst success factors, the article lists *establishing specific goals* that help partners to remain committed to the relationship. Not agreeing upon goals might lead to negative impact on the company growth and strategic direction. *Ongoing management of the relationship* is also necessary: coordination, communication, transparency and commitment are essential for reaching success. *Cultural issues*, between companies as well as between departments, are one of the major problems for achieving this. The author also points out that developing partnerships is not always a linear process, and that changes in the relationships are required if the market conditions change. Partnerships that are long lasting often evolve into a merger. The main reason of failure is that one partner starts to *underperform*, which affect all partners.

2.1 Partnering in the Construction Industry

Ng et al. [15] studied six failed partnering projects within construction and identified in total fifteen problematic issues. Six were general issues: *lack of continuous and open communication, stakeholders not developing a "win-win" attitude, poor commitment to the partnering arrangement, lack of intimacy in the relationship, issues are allowed to escalate, and some partners unwilling to compromise*. These issues address mainly the basic characteristics of partnering and the core of the partnering idea. Thus, issues in these areas would directly affect the partnering relationship. Continuing, the authors identified four project-specific issues. One was that the *project was perceived as not suitable for partnering*, thus the initial conditions were not suitable. For public projects, the *competitive tendering process* was seen as problematic as well. Issues with whom to include and when also existed, leading to *problems with technical specifications* and commitment. Including contractors earlier in the design phase could be a way to reduce specification issues while *including subcontractors* could have contributed positively to the commitment of contractor. Another five issues were client or contractor specific. *Lack of empowerment, large bureaucratic organizations*, and *lack of technical knowledge* were hinders on the clients side. On the contractor side, *commercial pressures* were compromising the partnering attitude. Projects require financial flexibility to succeed with partnering. Also the knowledge about partnering is important. *Lack of training and guidance in the partnering concept* was mentioned as a problematic issue. Bresnen and Marshall [2] draw conclusions from two construction project on the nature of partnering. Firstly, it is *hard to attribute project performance with partnering*. Several other project related variables could be the cause for good or poor performance. Also, partnering is not necessarily the solution to project related problems. Continuing, there is *no one strategy for successful partnering* and *partnering is flexible and iterative* in its nature and is not only influenced by the project context, but also by wider organizational structures and cultures.

In Kadefors [10] theory of trust is used for understanding the building of trust in partnering projects. The article concludes that *team building processes* and *project-wide communication* in the early stages positively influenced the participants and contributed to building up trust in the long-term. Moreover, *systems for monitoring relations and managing conflicts* could help in preventing distrust. The contractor selection process was studied in Kadefors et al. [11]. In procurement of partnering based projects the focus is on *evaluating individuals* rather than companies, and their attitudes and team working abilities. The *bidding documents should be adapted* for capturing other aspects, for instance, visions and goals with the partnering itself, and not only the project deliverables and other traditional contract specifications. An extensive literature review of partnering in construction by Bygballe et al. [4] concludes that focus has been on the dyad rather on the multi-nature of construction, and on formal tools and outcomes rather than on social and informal aspects of the relationships. The authors suggest that construction could incorporate knowledge from other areas such as supply chain management.

Venselaar et al. [23] studied the social aspects of supply chain partnering in the construction industry from a work floor level, thus addressing the improvement areas identified in Bygballe et al. [4]. The main barrier found in the case study was the *lack of reaching a shared understanding of strategic needs*, which was affected by social factors of leadership and trust. The social aspects of construction partnering projects were addressed in Laan et al. [12] and Davies and Love [5] in terms of building and maintaining relationships through interview studies. Laan et al. [12] find that relationship-based contract forms change the conditions in the initiation stage; they *build trust and bring incentives to manage risks cooperatively* instead of sub-optimizing and distrusting each other. Partnering also helps in maintaining the trust because distrust is more costly than preserving the alliance. The *selection of employees* to actively work in the alliance is an important factor, according to the findings. Moreover, the development of trust is dependent on the *relational quality* rather than on traditional project control mechanisms such as budget and time planning. Davies and Love [5] propose a three-stage model for relationship development in construction. *Individual relationships, trust* and *organizational development* are cornerstones the relationship development process.

2.2 Partnering-Based and Performance-Based Contracts in the Maintenance Area

Olsson and Esping suggested already in 2004 partnering as a means to reach effectiveness in infrastructure maintenance outsourcing contracts. They note that there are several differences between construction and maintenance which should be taken into consideration when applying partnering in maintenance contracts. Partnering is potentially beneficial in situations with high complexity and a limited amount of contractors available in the market, and where activities are business critical. All these three characterize the Swedish railway maintenance. Thus, the authors conclude that the potential of partnering is very high; cost reduction to up to 30% in combination with high availability performance. A pilot project was conducted by the authors and reported in Esping and Olsson [7]. Some lessons learned were that in order to succeed, *financial incentives for all partners must exist* and a gain-sharing mechanics should be used. Moreover, the *top management must be committed* and the implementation should be *facilitated by a coordinator* which helps in creating an efficient team and promote openness and honesty in communication. In addition, the *objectives, as well as the strategy to reach them, should be clear and agreed upon*. The results from the pilot project were positive; all objectives were met or exceeded. For instance, the reduction in time delays was 19%, and the reduction in the number of breakdown repairs was 14%. Performance-based maintenance partnerships could provide benefits for housing maintenance according to Straub [19]. For achieving the benefits the partnering concept must be *adapted for maintenance contracts*. The author did not bring any

hard evidence on true benefits, but the housing companies included in the study expected improved quality of maintenance projects, improved performance and achieving budget certainty with performance-based maintenance partnerships. The main fear was that performance-based contracting would lead to less control, less flexible decision making and loss of knowledge about one's own properties.

Lingegård [13] studied the possibility to reach resource-efficiency in road and rail infrastructure by applying integrated product service offerings. Amongst the challenges mentioned were the lack of information and knowledge sharing between actors due to the use of traditional contract forms and conservative cooperate cultures. Actors tend to sub-optimize instead of cooperate. Just like Olsson and Esping [16] the author finds that rail and road infrastructure have characteristics well suited for holistic and cooperative contract forms such as partnering. Partnering potentially increases the information/knowledge sharing by building trust, and resource-efficiency could be reached by involving contractors already in the design phase. One obstacle identified is *risk-adversity amongst actors*, which implies a need for risk sharing models and that too speedy implementation of partnering should be avoided. In a study of maintenance contracts for road and railway it was concluded that contractors believe that high-quality work depends on close cooperation with clients, but that current tendering and contracting principles are not clear enough and result in lack of trust between contractor and client, Abdi et al. [1]. On the clients' side lack of expertise and reorganizations has affected the performance of maintenance activities negatively, even in case of well-formulated contracts. Partnering is seen as one possible way to over-come current problems. The authors point out the *ability to adjust the contracting during the contract period* and the need for *continuous dialogue between client and contractor* as two of the benefits.

3 Maintenance Related Problems of Swedish Railway

The railway in Sweden has undergone a deregulation process starting in the late 80s until 2011, when the traffic was fully deregulated. In 1988 around ten actors totally operated in the Swedish railway industry. Today, the number of actors are more than a thousand when taking into account all the contractors and subcontractors. The coordination of all actors is complex and the planning horizons for operations are long, Trafikverket [21]. This puts high demand on information handling as well as decision-making processes. An increase in the railway transports of 42% in recent years has led to very high capacity utilization and severe operational problems such as timetabling of bottlenecks, time delay sensitivity, reduced speed due to bottlenecks and delayed maintenance actions, which is a safety issue, Trafikverket [22]. All these problem areas are directly or indirectly affecting the railway maintenance. The problematic situation regarding maintenance services management of Swedish railway was in Ingwald and Kans [9] summed up in three distinct problem

areas: maintenance ineffectiveness, maintenance services not fulfilling the market needs, and lack of capacity. The causes are several and also interrelated.

Maintenance ineffectiveness is mainly caused due to a low level of competition, or a feeling that the market is not a real market, too detailed contracts, and poor cooperation amongst actors. The inability of maintenance services to fulfil the market needs is a result of low level of competition and poor co-operation amongst actors, and a poor quality assurance system. Reasons behind lack of capacity are inaccurate maintenance plans, complicated timetabling process, and high amount of degraded infrastructure. The root causes are found in three main areas: *information handling and management* (poor internal management, lack of suitable IT-support, poor information handling, and poor reporting structures), *regulation and control* (passive governmental management, vague contracts, conservative buyers' culture, poor quality charging system, and cost model not connected to real needs), and *resources* (low availability of rolling stocks, lack of appropriate maintenance resources, lack of investments in infrastructure, incomplete contractor abilities and competence, and inaccurate analysis models). The complicated timetabling process could be a resource matter too as it puts high demand on the operators' planning capabilities and understanding for the planning process. For more details, see Ingwald and Kans [9].

4 Is Partnering a Solution to the Problems in Swedish Railway Maintenance?

The concept of partnering has been used in order to overcome problems that arise when utilising more traditional methods for managing outsourcing. The Swedish railway suffers from severe operational problems directly or indirectly related to railway maintenance. In Table 1 current problems in railway maintenance are mapped against characteristics of successful partnering. Although partnering being a relatively new concept it has been used for ten-twenty years in the construction industry, and experience from this indicates several benefits that could be gained if the concept of partnering was introduced on a wider scale in Swedish railway maintenance. Several of the current problems described in Swedish railway could probably be reduced or at least managed in a better way. One large problem today is related to information handling. Relevant and complete information is not available for decision maker and maintenance planners. Partnering as a cooperative form of collaboration, building on trust and striving for agreed goals, creates a better condition to develop a suitable IT-solution that are beneficial for all involved actors.

Efficient information handling is identified as requirement for a successful partnering relation. Other identified problem areas, e.g. vague contracts, a conservative buyers' culture and poor quality charging system, could be managed better and more effectively if successful partnering could be established, because partnering requires agreed goals, cooperation and openness. Partnering is not the

Table 1 Mapping of problems areas in Swedish railway maintenance and characteristics of successful partnering

Information handling and management	
Poor internal management	There are no direct provisions for information handling and IT-systems in the description of partnering. However, result from case-studies shows that a suitable IT-solution is a must. Both development of management and an IT-solution is supported by the agreed common goals and also by the trust that must exist be-tween involved actors
Lack of suitable IT-support	
Poor information handling	
Poor reporting structures	
Regulation and control	
Passive governmental management	This could be an issue, since actors already involved in partnering projects, can have a substantial favour compared to other actors at the time of contracting. This may reduce competition
Vague contracts	In partnering, since the involved actors are together working towards agreed goals, it is possible during the contract period to adjust the agreements based on a continuous dialog among the actors. The alliance should be managed for flexibility rather than efficiency
Conservative buyers' culture	A win-win attitude among the involved actors is a characteristic of successful partnering
Poor quality charging system	The involved actors agree upon goals, and work together to reach this goal. Also the openness and continuous discussion required for a successful partnering-relation makes it easier to effectively pursue the goal
Cost model not connected to real needs	Conflicting goals among the actors could be a barrier. Financial flexibility is required to succeed. Several factors not traditionally related to project performance could be vital
Resources	
Low availability of rolling stocks	The concept of partnering includes nothing that directly relates to availability of rolling stock. However, depending on who the partners are, the utilization of available resources could be affected
Lack of appropriate maintenance resources	A successful partnering requires open cooperation and that all involved actors are working together towards agreed goals. This characteristic of partnering as being cooperative and more holistic and based on trust would increase the willingness to share knowledge and resources, in order to reach the objectives
Lack of investments in infra-structure	The concept of partnering will not have any direct impact on this problem
Incomplete contractor abilities and competence	The characteristics of partnering will not directly have an impact on the available abilities and competences. However, depending on type of partnering, strategic decisions could be managed in the partnering-relation what competencies are actually required. Also the cooperative work means that available abilities and competencies can be used more effectively
Inaccurate analysis models	Not directly influenced by partnering. However, the cooperative characteristics of partnering allow for a more holistic view of the situation

solution for everything; for example, lack of investment in infrastructure is probably not directly influenced by the type of relationship between business partners. Yet, the deeper cooperation that is required in partnering may influence the utilization of equipment and thereby decrease the problem. A barrier that must be considered is related to contracting. Partnering could place bidders, not already involved in a partnering relationship, in a negative position during bidding, and this is against governmental practices. Partnering seems to be a promising concept for solving several of the current problems in Swedish railway maintenance. Partnering might be applicable not only for the railway industry and the construction industry, but in all industries where maintenance is strategically important and long-term relationships are required with the maintenance subcontractor, for instance, the aircraft industry and process industries such as power generation, food and chemical industries. However, since partnering is a relatively new concept there are still many aspects that is in need of further research. Most of the experience reported regarding partnering is from other industries than railway maintenance, and the cultural differences between industries as well as between partners within an industry need to be considered, such as tendering rules and procedures for government/public procurement.

Acknowledgements This research forms a part of the project Future Industrial Services Management founded by the Swedish innovation agency Vinnova.

References

1. Abdi A, Lind H, Birgisson B (2014) Designing appropriate contracts for achieving efficient winter road and railway maintenance with high performance quality. International Journal of Quality and Service Sciences 6(4):399–415
2. Bresnen M, Marshall N (2002) The engineering or evolution of co-operation? A tale of two partnering projects. Int J Proj Manag 20:497–505
3. Bresnen M (2007) Deconstructing partnering in project-based organisation: seven pillars, seven paradoxes and seven deadly sins. Int J Proj Manag 25:365–374
4. Bygballe LE, Jahre M, Swärd A (2010) Partnering relationships in construction: a literature review. J Purch Supply Manag 16:239–253
5. Davies P, Love P (2011) Alliance contracting: adding value through relationship development. Eng Constr Archit Manag 18(5):444–461
6. Dodourova M (2009) Alliances as strategic tools. Manag Decis 47(5):831–844
7. Esping U, Olsson U (2004) Part II. Partnering in a railway infrastructure maintenance contract: a case study. J Qual Maint Eng 10(4):248–253
8. Godoy DR, Pascual R, Knights P (2014) A decision-making framework to integrate maintenance contract condition with critical spares management. Reliab Eng Saf 131:102–108
9. Ingwald A, Kans M (2016) Service management models for railway infrastructure, an ecosystem perspective. In: Proceedings of the 10th world congress on engineering asset management (WCEAM 2015), pp 289–303
10. Kadefors A (2004) Trust in project relationships—inside the black box. Int J Proj Manag 22:175–182
11. Kadefors A, Björlingson E, Karlsson A (2007) Procuring service innovations: contractor selection for partnering projects. Int J Proj Manag 25:375–385

12. Laan A, Noorderhaven N, Voordijk H, Dewulf G (2011) Building trust in construction partnering projects: an exploratory case-study. J Purch Supply Manag 17:98–108
13. Lingegård S (2014) Integrated product service offerings for rail and road infrastructure—reviewing applicability in Sweden. Ph.D. Dissertation, Linköping University, Linköping
14. Murthy DNP, Asgharizadeh E (1999) Optimal decision making in a maintenance operation. Eur J Oper Res 116(2):259–273
15. Ng ST, Rose TM, Mak M, Chen SE (2002) Problematic issues associated with project partnering—the contractor perspective. Int J Proj Manag 20:437–449
16. Olsson U, Esping U (2004) Part I. A framework of partnering for infrastructure maintenance. J Qual Maint Eng 10(4):234–24719
17. Olsson UKG (2012) Affärsmodeller för Partnering och Utökad samverkan. Former för samverkan, ekonomisk ersättning och upphandling, öppna böcker och gemensam kostnadsstyrning vid användning av Utökad samverkan eller Partnering. Fernia Consulting, Skurup
18. Sherwin D (2000) A review of overall models for maintenance management. J Qual Maint 6 (3):138–164
19. Straub A (2007) Performance-based maintenance partnering: a promising concept. J Facil Manag 5(2):129–142
20. Stenström C, Parida A, Kumar U, Galar D (2013) Performance indicators and terminology for value driven maintenacne. J Qual Maint Eng 19(3):222–232
21. Trafikverket (2014a) Att skapa tidtabeller för tåg - nu och i framtiden Från dialog och ansökan till samordnad tågplan. Trafikverket, Borlänge
22. Trafikverket (2014b) The Swedish transport administration annual report 2013. Trafikverket, Borlänge
23. Venselaar M, Gruis V, Verhoeven F (2015) Implementing supply chain partnering in the construction industry: work floor experiences within a Dutch housing association. J Purch Supply Manag 21:1
24. Vilkamo T, Keil T (2003) Strategic technology partnering in high-velocity environments—lessons from a case study. Technovation 23:193–204

Overhaul Decision of Repairable Systems Based on the Power-Law Model Fitted by a Weighted Least Squares Method

R. Jiang

Abstract The power law model has been widely used to analyze failure data from repairable systems and to optimize overhaul decision. However, it is not applicable for the situations where the empirical mean cumulative function displays complex shapes. It is noted that the failure observations in wear-out phase have larger influence on overhaul decision than the failure observations in early and normal use phases. This implies that the observations in different phases have different importance and the power-law model may be still appropriate for the overhaul decision optimization as long as the fitted model is mainly based on the observations in the wear-out phase. In this paper, we use a simple weight function, which is proportional to the normal density function, to reflect the importance of failure observations at different times. We propose a heuristic method to determine the parameters of the weight function so that it can appropriately stress the recent observations. The observed data are fitted to the power-law model using a weighted least squares method, and the fitted model is then used to optimize the overhaul decisions for the population and individual systems, respectively. The appropriateness of the proposed approach is illustrated by a real-world example.

1 Introduction

The power-law model has been widely used to model the failure processes of repairable systems when the mean cumulative function (MCF) is concave or convex. The resulting model can be used to predict failure and/or optimize maintenance decision. However, the empirical MCF is often neither concave nor convex (e.g., see [10]). In this case, the power-law model is inappropriate for modelling the corresponding failure process.

The failure intensity of a repairable system is often bathtub-shaped [5, 7, 14, 15]. In this case, the failure process consists of three phases: early failure phase, random

R. Jiang (✉)
Changsha University of Science and Technology, Changsha, China
e-mail: jiang@csust.edu.cn

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_28

failure phase and wear-out phase. Clearly, the failure observations in wear-out phase have larger influence on overhaul decision than the failure observations in early and normal use phases. This implies that the observations in different use phases have different importance and the power-law model may be still appropriate for the overhaul decision optimization as long as the fitted model is mainly based on the observations in the wear-out phase.

Jiang [8] deals with estimating the residual life distribution in a condition-based maintenance setting. He first derives the fractile curves of a degradation variable from the observed data and then fits these curves to a three-parameter power-law model using a weighted least squares method (LSM). The weight function used there is proportional to the normal density function with the mean parameter (μ) being set at the last observation time and the standard deviation parameter (σ) being simply taken as $\sigma = \mu/3$. This approach can be further improved. For example, the value of σ should be determined in a more objective way.

In this paper, we propose a heuristic approach to determine the value of σ. The power-law model fitted by the weighted LSM is then used to optimize the overhaul decisions for the population and individuals, respectively. It is shown by a real-world example that the model fitted by the weighted approach provides much better prediction of time to failure than the ordinary LSM and maximum likelihood method (MLM).

The paper is organized as follows. The maintenance policy is defined in Sect. 2. The proposed parameter estimation method is presented in Sect. 3 and illustrated in Sect. 4. The paper is concluded in Sect. 5.

2 Power-Law Model and Failure Counting Policy

2.1 Power-Law Model

Consider a failure point process $(t_i, i = 1, 2, \ldots)$. For a given time t, the number of cumulative failures over $(0, t)$, $N(t)$, is a discrete random variable. The MCF is the expectancy of $N(t)$ and denoted as $M(t)$. A typical model for representing the MCF of a failure point process is the power-law model, given by

$$M(t) = (t/\eta)^{\beta}; \beta, \eta > 0. \tag{1}$$

The corresponding failure intensity function is given by

$$m(t) = \beta M(t)/t. \tag{2}$$

2.2 Failure Counting Policy and Its Cost Model

Under the failure counting policy [11, 13], a system is replaced at the n-th failure with cost c_p, and the first $n-1$ failures are corrected by minimal repairs (with cost c_r per minimal repair). Let $t(n)$ denote the expected time to the n-th failure. The cost rate is given by

$$J(n) = [c_p + (n-1)c_r]/t(n). \tag{3}$$

Here, $t(n)$ is given by [6]:

$$t(n) = \sum_{i=0}^{n-1} \int_0^{\infty} \frac{M^i(t)e^{-M(t)}}{i!} dt. \tag{4}$$

When $M(t)$ is the power-law model given by Eq. (1), we have

$$t(n) = \frac{\eta}{\beta} \sum_{i=0}^{n-1} \Gamma(i+1/\beta)/\Gamma(i+1) \tag{5}$$

where $\Gamma(.)$ is the complete gamma function. Using Eq. (5) to Eq. (3) yields

$$\eta J(n)/c_r = (c_p/c_r + n - 1)\beta / \sum_{i=0}^{n-1} \Gamma(i+1/\beta)/\Gamma(i+1). \tag{6}$$

The optimal policy parameter is n^*, which makes $J(n)$ achieve its minimum. An Excel spreadsheet program is developed to find the optimal solution.

In this paper, we consider a variant of the failure counting policy: a system is overhauled (rather than replaced) with cost c_p at the n-th failure, and the first $n-1$ failures are corrected by minimal repairs. Under this policy, the cost model is still Eq. (3). However, the values of β and η may change after an overhaul if the overhaul is imperfect. In this case, according to Eq. (6) n^* can remain unchanged as long as the cost ratio (c_p/c_r) and β are unchanged. Since the time to the n^*-th failure for different system can be different, the overhaul decision has "condition-based" feature to some extent.

3 Parameter Estimation Methods

For a given cost ratio, the optimal solution depends on β, which describes the failure growth rate in the wear-out phase. The larger it is, the faster the failure grows.

The recent growth rate is more important than the past growth rate for the overhaul decision. We propose a weighted approach to reflect this point in this section.

3.1 Weighted Least Squares Method

Consider N nominally identical systems. Their failure processes are described by

$$\{t_{i1} \le t_{i2} \le \cdots \le t_{im_i} \le T_i, 1 \le i \le N\}. \tag{7}$$

Here T_i is censored time if $T_i > t_{im_i}$; otherwise, it is a failure time. The LSM requires calculating the empirical MCF. To do so, we pool and reorder all the observations (both failed and censored) as below:

$$t_{(1)}(s_1) \le t_{(2)}(s_2) \le \cdots \le t_{(j)}(s_j) \le \cdots \le t_{(K)}(s_K) \tag{8}$$

where K is the number of all the observations and s_j $(1 \le j \le K)$ is the number of systems under observation at $t_{(j)}$. The empirical MCF is given by

$$M_e(t_{(j)}) = [M^*(t_{(j-1)}) + M^*(t_{(j)}]/2 \tag{9}$$

where $M^*(t_{(0)}) = 0$, $M^*(t_{(j)}) = M^*(t_{(j-1)}) + I_j/s_j$, and $I_j = 1\,[= 0]$ for a failure [censoring] observation. For more details about the empirical MCF, see Leemis [12], Arkin and Leemis [1], Jiang and Guo [10].

Assume that the MCF of the population is represented by the power-law model. The least squares estimates of the parameters can be obtained by minimizing the sum of squared errors given by [9]

$$SSE = \sum_{j=1}^{K} [M(t_{(j)}; \beta, \eta) - M_e(t_{(j)})]^2. \tag{10}$$

Let $w(t)$ $(\in (0, 1])$ denote a weight function. The weighted least squares estimates of the parameters can be obtained by minimizing the following

$$SSE_w = \sum_{j=1}^{K} w(t_{(j)})[M(t_{(j)}; \beta, \eta) - M_e(t_{(j)})]^2. \tag{11}$$

In this paper, we consider the following form of weight function:

$$w(t) = \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] = \sqrt{2\pi}\sigma\phi(t;\mu,\sigma), 0 < t \leq \mu \tag{12}$$

where $\phi(.)$ is the normal pdf.

When modelling is for the population, we take $\mu = \max(T_1, \ldots, T_n)$ and σ is determined by solving the following equation

$$\sum_{j=1}^{K} w(t_{(j)}) = n_w \tag{13}$$

where n_w is a known value to be specified. Note that $n_w = K$ when $w(t_{(j)}) \equiv 1$. This implies that n_w can be viewed as an equivalent sample size. Since the power-law model has 2 parameters, n_w must be larger than 2. On the other hand, a small σ (which results in a small n_w) emphasizes the recent observations more. Taking account of these two aspects, we take $n_w = 5$ when modelling is for the population.

When modelling is for a single failure process, we take $\mu = T_i$. To make the modelling possible for the case of $m_i = 3$, n_w must be smaller than 3. Therefore, we take $n_w = 2.5$.

3.2 *Maximum Likelihood Method*

The maximum likelihood method is often preferred in parameter estimation. For the data given by Eq. (7), the log-likelihood function is given by [9]

$$\ln(L) = \sum_{i=1}^{N}\left\{-M(T_i) + \sum_{j=1}^{m_i}\ln[m(t_{ij})]\right\}. \tag{14}$$

The parameters are estimated by directly maximizing $\ln(L)$.

4 Illustration

The data shown in Table 1 deal with failure times of the hydra systems of five LHD machines and can be found from Attardi and Pulcini [2] and Blischke et al. [3].

Table 1 Failure times (in hours) of hydra systems of LHD machines

LHD01, $i = 1$	LHD03, $i = 2$	LHD09, $i = 3$	LHD11, $i = 4$	LHD17, $i = 5$
327	637	278	353	401
452	677	539	449	437
459	1074	1529	498	455
465	1110	1720	709	614
572	1164	1827	791	955
849	1217	1859	966	1126
903	1314	1910	1045	1150
1235	1377	1920	1162	1500
1745	1593	2052	1188	1572
1855	1711	2228	1192	1875
1865	1836	2475	1197	1909
1874	1861	2640	1257	1954
1959	1865	3094	1296	2278
1986	1966	3236	1331	2280
2045	2150	3274	1589	2350
2061	2317	3523	1686	2407
2069	2398	3735	1745	2510
2103	2444	3939	1748	2521
2124	2462	4121	1785	2526
2276	2494	4237	1793	2529
2434	2713	4267	2038	2673
2478	3118	4291	2117	2753
2496	3138	4323	2166	2806
	3386	4361	2197	2890
	3526	4371	2456	3108
		4682	2739	3230
		4743	2889	

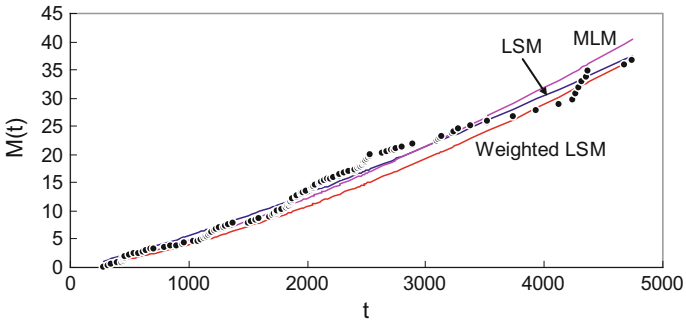


Fig. 1 MCF curve and power-law models fitted by different methods

Table 2 Parameter estimates and corresponding optimal policy parameters

Method	β	η	n^*
LSM	1.2187	242.6	42
Weighted LSM	1.4157	371.2	22
MLM	1.3775	323.8	24

4.1 Parameter Estimation for Population

Figure 1 shows the plot of empirical MCF (i.e., those dots). As seen, the empirical MCF is neither concave nor convex, implying that the power-law model is not appropriate for modelling the data on the whole. However, to avoid developing a complex model, we may use the power-law model to approximately describe the right-hand-side part of MCF so as to obtain a relatively reasonable overhaul decision.

For $\mu = T = 4743$ and $n_w = 5$, solving Eq. (13) yields $\sigma = 345.3$. Using the ordinary and weighted LSMs, respectively, we obtained the estimates of model parameters shown in the second and third rows of Table 2. As seen from the table, the parameters estimated from the two LSMs are quite different. The maximum likelihood estimates of the parameters are shown in the last row of Table 2.

4.2 Population-Oriented Overhaul Decision

Assume that $c_r = 1$ and $c_p = 10$. The minimal repair numbers associated with the failure counting policy are shown in the last column of Table 2. As seen, different estimation methods lead to significantly different values of the decision variable (i.e., n^*). This confirms the importance to use an appropriate parameter estimation method.

According to the optimal solution obtained from the weighted LSM (i.e., $n^* = 22$), the overhaul times for systems 1–5 are 2478, 3118, 4291, 2117 and 2753 h, respectively. This illustrates that the failure-counting policy is a "condition-based" policy.

4.3 Accuracy of Failure Prediction

According to Coetzee [4], the fitted power-law model can be used to

1. understand the failure behaviour of the repairable system,
2. forecast future failures, and
3. optimize the maintenance strategy for the repairable system.

In this subsection, we illustrate the application of the fitted model in predicting the overhaul times of systems.

Table 3 Model parameters estimated by different methods

	i	1	2	3	4	5
LSM	β	1.5757	1.5497	1.4327	1.5374	1.5926
	η	365.3	392.8	521.3	276.6	418.4
Weighted LSM	σ	147.2	371.3	52.79	173.7	122.3
	β	1.2589	0.7846	7.4168	1.0033	2.0176
	η	218.2	61.29	2838.4	99.69	600.3
MLM	β	1.5751	1.5841	1.5884	1.6381	1.6660
	η	348.2	443.1	612.9	320.8	430.6

Table 4 Actual and predicted overhaul times and averages of their relative errors

i	1	2	3	4	5	ε
Δt_i	202	624	54	324	224	
$\Delta t_{p,i}$ from LSM	153.0	172.7	290.8	124.5	169.8	1.2418
$\Delta t_{p,i}$ from weighted LSM	185.5	361.0	56.0	196.7	128.9	0.2713
$\Delta t_{p,i}$ from MLM	146.0	182.7	250.7	120.1	153.7	1.1141

We consider the failure counting policy with $n^* = 22$. Assume that we want to predict the overhaul times (i.e., $t_{i,22}$) for all systems at $t = t_{i,20}$. We first fit the first 20 observations of each failure point process to the power-law model using various parameter estimation methods. The results are shown in Table 3. As seen, the estimates of β obtained from the weighted methods are significantly different from the estimates of β obtained from LSM and MLM. This confirms the necessity of introducing the weighted approaches.

The actual residual time to overhaul for the i-th system is given by $\Delta t_i = t_{i,22} - t_{i,20}$, which is shown in the second row of Table 4. For the power-law model associated with the i-th system, the predicted residual time to overhaul is given by $\Delta t_{p,i} = t_i(22) - t_i(20)$, which can be evaluated using Eq. (5). The results are shown in the last three rows of Table 4.

The performance of a parameter estimation method can be evaluated using the average of relative errors

$$\varepsilon = \frac{1}{5}\sum_{i=1}^{5} |1 - \Delta t_{p,i}/\Delta t_i|. \tag{15}$$

The last column of Table 4 shows the values of ε. As seen from the table, the power-law model obtained from the weighted LSM provides the best prediction accuracy; and the power-law models obtained from the LSM and MLM provide much poorer prediction accuracies. This further confirms the advantage of the weighted approach.

5 Conclusions

In this paper we have used the failure counting policy for determining the overhaul time of a repairable system. We have shown that the optimal solution only depends on the cost ratio and the value of β without a need to assume that the overhaul is perfect.

We have proposed the weighted LSM to fit a failure point process to the power-law model. Since the weighted approach stresses recent observations of the failure process, the fitted power-law model can be used for situations where the empirical MCF has a complex shape.

Through a real-world example we illustrated that the predicted times to overhaul based on the model fitted by the weighted LSM are much more accurate than the overhaul times predicted based on the models fitted by LSM and MLM.

Topics for future research include (a) developing other form of weight functions and (b) developing other weighted methods that contain the element of MLM.

Acknowledgements The research was supported by the National Natural Science Foundation (No. 71371035).

References

1. Arkin BL, Leemis LM (2000) Nonparametric estimation of the cumulative intensity function for a nonhomogeneous Poisson process from overlapping realizations. Manag Sci 989–998
2. Attardi L, Pulcini G (2005) A new model for repairable systems with bounded failure intensity. IEEE Trans Reliab 572–582
3. Blischke WR, Karim MR, Murthy DNP (2011) Warranty data collection and analysis. Springer, New York
4. Coetzee JL (1997) Role of NHPP models in the practical analysis of maintenance failure data. Reliab Eng Syst Saf 161–168
5. Guida M, Pulcini G (2009) Reliability analysis of mechanical systems with bounded and bathtub shaped intensity function. IEEE Trans Reliab 58:432–443
6. Jiang R (2013) Life restoration degree of minimal repair and its applications. J Qual Maint Eng 19(4):413–428
7. Jiang R (2015) An approximation to mean time to the next failure for repairable systems. In: The ninth international conference on mathematical methods in reliability, Tokyo, Japan, pp 436–442
8. Jiang R (2015) Estimating residual life distribution from fractile curves of a condition variable. In: 2015 prognostics and system health management conference, Beijing, RP0294
9. Jiang R (2015) Introduction to quality and reliability engineering. Springer, Berlin, pp 92–93, 205–206
10. Jiang R, Guo Y (2014) Estimating failure intensity of a repairable system to decide its preventive maintenance or retirement. Int J Perform Eng 577–588
11. Jiang R, Zhou Y (2010) Failure-counting based health evaluation of a bus fleet. In: 2010 Prognostics and system health management conference, Macau, China, MU3013
12. Leemis LM (1991) Nonparametric estimation of the cumulative intensity function for a nonhomogeneous Poisson process. Manag Sci 886–900

13. Makabe H, Morimura H (1963) On some preventive maintenance policies. J Oper Res Soc Jpn 17–47
14. Mun BM, Bae SJ, Kvam PH (2013) A superposed log-linear failure intensity model for repairable artillery systems. J Qual Technol 45:100–115
15. Pulcini G (2001) Modelling the failure data of a repairable equipment with bathtub type failure intensity. Reliab Eng Syst Saf 71:209–218

Process Characteristics and Process Performance Indicators for Analysis of Process Standardization

Achim Kampker, Maximilian Lukas and Philipp Jussen

Abstract Industrial service companies deliver technically complex services (inspection, maintenance, repair, improvement, installation) for an enormous variety of technical assets in the chemical, steel, food and pharmaceutical industry. This variety of assets leads to a corresponding variety of service processes. To ensure competitiveness, the management of industrial service companies aims to increase the service process efficiency, especially through service process standardization. However, decision-makers struggle to make knowledge-based decisions on service process standardization because ex-ante the cost-benefit ratios of process standardization are unknown. The missing understanding of cost-benefit ratios of process standardization is caused by a missing understanding, which interdependencies exist between process characteristics and process performance indicators. Thus, the objective of this paper is to determine suitable characteristics and performance indicators to measure the way service provision processes are executed in the industrial service sector. The results represent the basis for executing an empirical questionnaire study focusing on the execution of service provision processes and identifying the cause-effect relations of process standardization.

1 Introduction and Objective

Industrial service companies deliver complex technical services (inspection, maintenance, repair, improvement, installation) in the process industry (chemical, steel, food and pharmaceutical industry) [36]. These service providers, mostly small and medium-sized companies (SMEs), have to cope with a high variety of installed

A. Kampker · M. Lukas (✉) · P. Jussen
FIR at RWTH Aachen, Campus-Boulevard 55, 52074 Aachen, Germany
e-mail: Maximilian.Lukas@fir.rwth-aachen.de

A. Kampker
e-mail: achim.kampker@fir.rwth-aachen.de

P. Jussen
e-mail: Philipp.Jussen@fir.rwth-aachen.de

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_29

technical assets, paired with diverse asset conditions and degrees of wear [15]. This variety of service objects leads to heterogeneous demands for service provision processes, e.g. complexity of diagnosis, qualification of service technicians, and integration of customer's employees in service processes. The industrial service sector in Europe is characterized by an intense competition between mostly SMEs [1]. The market place for industrial services constitutes a buyer's market. A little total number of customers, to a large extent large process industry companies, is capable to stipulate market conditions, e.g. service pricing systems and service product portfolios [4]. In order to stay ahead of their competition, industrial service companies have to enhance their service provision process efficiency to improve their competitiveness.

Decision-makers of industrial service companies aim to enhance their service efficiency by increasing the standardization level of their service processes. Process standardization has a long successful history in the manufacturing industry. Starting with Taylorism and Fordism and leading to quality control, six sigma and lean management the reduction of process variance has led, inter alia, to substantial increases in manufacturing process efficiency [18]. Studies among industrial service providers show that, successful large companies standardize their processes to a high degree (cf. Fig. 1).

Standardization of service provision processes aims at minimizing process variance and thus increasing process control [31]. Thus, it helps to increase customer satisfaction and reduce waste [20]. High levels of process control result in process stability and allow precise process calculations, predictable profitability, and deriving process optimization measures purposefully [23, 32]. However, standardization of service provision processes in the industrial service sector poses multiple challenges. First, the high process variety leading to low repetition frequencies of individual service provision processes makes it difficult to determine well-suited processes to standardize. Second, decision-makers mostly do not know which quantitative effects process standardization measures will have on process performance. Thus, decision-makers are usually not capable to determine cost-benefit ratios of standardization measures ex-ante easily. This leads to subjective estimations of standardization effects based on individual experience and poses the risk to take bad decisions. Although scientific findings show that, an increasing level of process standardization leads to increasing benefits and costs, with a decrease of marginal utility and marginal costs, the precise quantitative

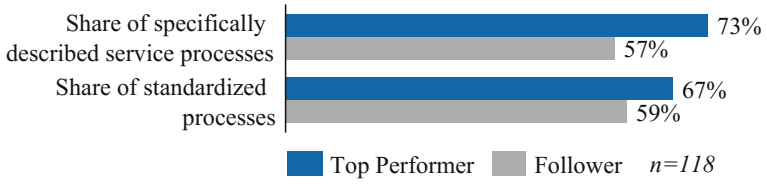


Fig. 1 Definition and standardization of service processes (mean values based on companies' statements) [30]

effects of process standardization remain unknown [28]. Due to their limited financial funds, especially SMEs have to ensure that investments into process standardization projects have a positive return on investment. In order to ensure financially sensible decisions resulting in a net benefit of standardization, often-extensive process cost analyses would be necessary. This enormous effort, however, discourages companies from promoting process standardization and improving their competitive situation.

In practice, the missing understanding of cost-benefit ratios of process standardization keeps industrial service companies from implementing it. The objective of this paper is to present a methodology to identify cause-effect relations regarding standardization of service provision processes in the industrial service sector. Those findings will support industrial service companies, especially SMEs, to derive financially sensible decisions regarding process standardization. To do so, this paper will outline the following:

1. Suitable process characteristics to describe the relevant process components that are subject to standardization for industrial service provision processes
2. Suitable process performance indicators to describe the level of performance for industrial service provision processes.

2 Theoretical Backgrounds

The Lean Services Framework developed at the Institute for Industrial Management (FIR) at RWTH Aachen University represents a well-established framework for industrial service management. It utilizes existing Lean Management principles and adapts them to industrial service management. By that, it enables industrial service companies to introduce Lean Management to industrial service business in a structured and holistic way [34]. Based on the principle of continuous improvement, the framework lays out an iterative application of its five phases and fifteen action-guiding principles. Each phase comprises three action-guiding principles, respectively (Fig. 2).

The Lean Services Framework acts as a scientific framework for the presented research. This paper refers to the third phase “Designing Service Provision”. This phase focusses on processes, resources and costs. It recommends defining standardized, repetitive processes based on the value creation for the customer to ensure a controllable level of process variance. Further, the third phase focusses on qualitative and quantitative dimensioning of the required resource capacities as well as securing service profitability by means of capturing customer’s price acceptance and calculating service provision costs. Specifically, this paper aims to substantiate the seventh action-guiding principle “Value Stream Definition with repetitive reference processes” by providing findings to design standardized reference processes for service provision evidence-based.

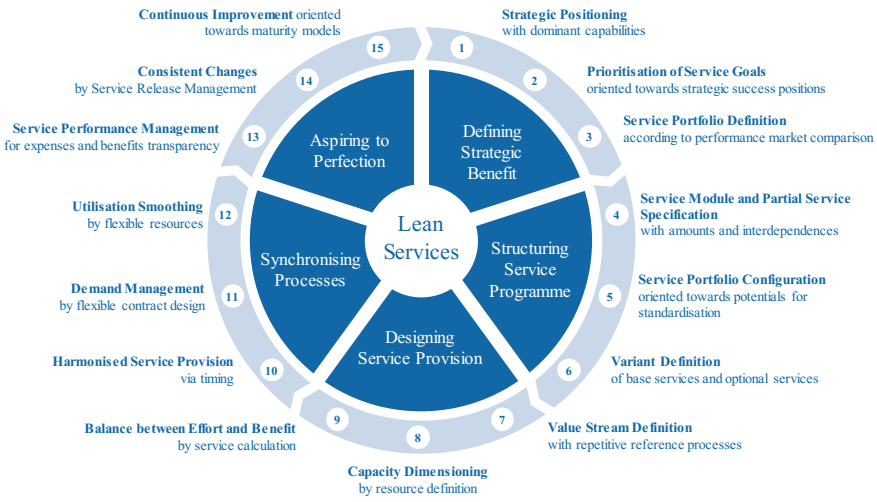


Fig. 2 Lean Services Framework (according to [29])

3 Methods

To explain the cause-effect relations of process standardization and its impact on process performance, the first step was to specify the scope of research through reviewing current industry surveys, literature and conducting interviews with industry experts. The review focused on studies targeting external industrial service companies and internal maintenance departments of manufacturing companies, which perform maintenance works according to DIN 31051:2012-09 as well as cleaning of assets [10]. Industry studies showed that the actual service provision processes and related activities lack standardization due to an enormous variety of assets, respective asset conditions and customer requirements [3, 11, 13, 26, 35]. The literature review showed that practicable approaches are missing to determine the sought after cause-effect relations for the actual service provision processes and related activities [5, 6, 16, 17, 25]. Expert interviews were conducted with senior industrial service experts during working group meetings with five industrial service companies as well as with one maintenance department of a chemical process company and a chemistry park operating company. This preparatory study determined service provision processes, e.g. repair, maintenance, inspection, modernization and cleaning of assets and equipment at the customer's site with an integration of the customer in the process as the research object in scope. Following, process characteristics were defined describing dispersion in process execution and determining the variance in the respective processes. The validation of these characteristics took place through iterative review by industry and scientific experts to ensure suitability. Second, process performance indicators, suitable to measure process performance from a customer's—and provider's perspective were

collected, reviewed and selected. Considering both scientific and industry opinion ensures the practicality of the selected performance indicators.

4 Results

An analysis of service processes in the industrial service sector shows that service processes largely possess the same structure. Processes start with order intake, followed by problem clarification, order planning, order control, order execution, order feedback and end with invoicing. This corresponds with well-established generic process models [19]. Further investigation reveals that the enormous process variety caused by a huge variety of technical assets of customers mostly arises in the execution of processes, meaning individual service orders. Therefore, better understanding standardization of industrial service processes demands analyzing the process execution, i.e. service orders. Consequently, repair, maintenance, inspection, modernization and cleaning of assets at the customer's site with an integration of the customer in the process represent the research object. Process steps such as order planning, order control, order feedback, and order invoicing are not the direct object of research.

4.1 Process Characteristics

The objective of process standardization is to achieve a certain level of process control by determining the way in which a process is executed [32]. The level of standardization refers to how wide or narrow process execution is determined. A low level of process standardization leaves a large leeway and vice versa. Understanding how process standardization works requires the identification of service process components targeted by process standardization. This requires identifying which components of service provision processes underlie an uncontrolled dispersion during process execution. In this context, dispersion refers to different ways of executing service orders. For example, dispersion can relate to different tools used during execution, different sequence of performed activities, different level of customer integration and different information available to service employees. It is important to consider dispersion of process execution for identical or equivalent processes. Thus, for example, the execution of two electric motor repair orders by service employees with different levels of qualification represent a dispersion in process execution, specifically the level of service employee qualification. Different levels of qualification of service employees for an electric motor repair order and a pipeline maintenance order however do not constitute dispersion in process execution.

The process step input, activity, service technician or executing system, service object, tools and systems, information, required material (spare parts, working

material), the external factor (customer), and process output represent service process components and allow describing service provision processes on a generic level (cf. Fig. 3).

In order to determine variance of characteristics during process execution the service process components that underlie dispersion in industrial services need to be specified and translated into measureable process characteristics. Based on [8, 12, 17, 22, 32] the following eight process characteristics were defined. Subsequently, the determined set of process characteristics were discussed with industry experts from five industrial service companies during a working group meeting and with two additional industry experts during individual workshops. All eight process characteristics were clarified during this validation and were confirmed (Fig. 4).

The determined process characteristics are divided into four categories; each category comprises two characteristics. The first category, *Level of standardization of process step content and sequence* focusses on how much leeway service employees have during process execution. The characteristic *Execution of activity* determines to what extent instructions are specifying how activities are executed (e.g. how a diagnosis of the customer's asset should be executed). The characteristic *Sequence of activities* determines to what extent the sequence of activities is predetermined (e.g. at which point a diagnosis takes place during order execution). The second category, *Standardization of operating materials and information* focusses on how much freedom employees have in terms of choosing operating materials and tools and which information is available for order execution. The characteristic *Operating material and tools* determines to what extent material and tools are predetermined and standardized (e.g. the use only one diagnosis software). The characteristic *Quality and form of order information* determines how order information is presented to the service employee (e.g. verbally, paper-based in writing, via an IT-system) and to what extent the provided order information is complete, matching and up-to-date (e.g. location of the customer's asset). The third category, *Homogeneity of employee skills* focusses on how the employees' skills regarding order execution vary. The characteristic *Experience* determines to what extent the employee who executes an order has experience in working on the asset and executing the type of order (e.g. maintenance, repair, cleaning). The characteristic *Qualification* determines if the employees has the required formal qualification to execute an order and to what extent the employee sees a need for further qualification. The forth category, *Integration of the external factor* focusses on the

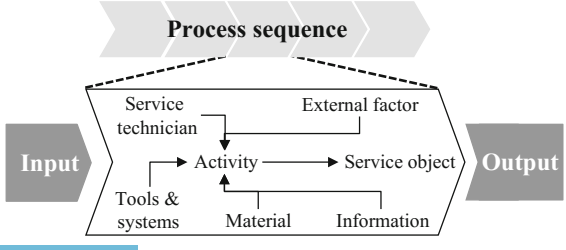


Fig. 3 Service process components describing service provision processes following [9]

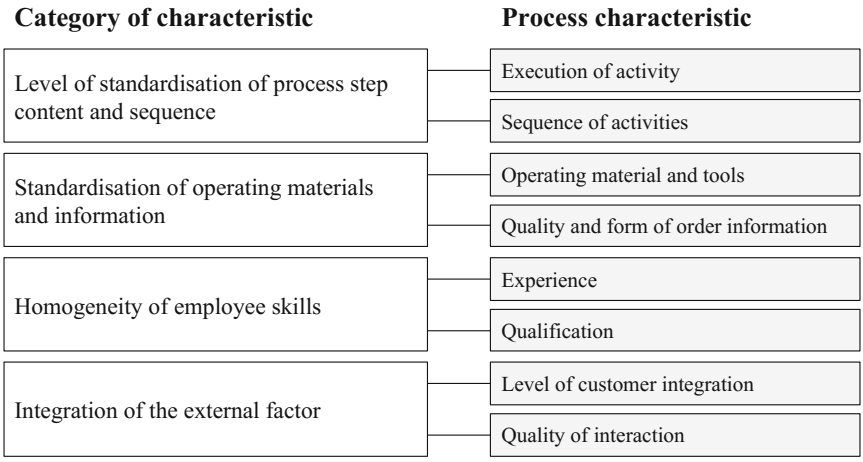


Fig. 4 Process characteristics for recording dispersion in process execution (author's own graph)

integration of the customer's employees into service provision. The characteristic *Level of customer integration* determines to what extent customers' employees take part in the order execution (e.g. disconnecting the asset from power supply, supervision of order execution, answering of queries). The characteristic *Quality of interaction* determines to what extent customer integration into the order execution meets quality requirements of the service provider (e.g. waiting time to answer queries, accuracy of provided information).

Because service objects, i.e. technical assets of customers, underlie a high dispersion and lead to heterogeneous processes in the first place, the service object does not represent a process characteristic.

4.2 Process Performance Indicators

Industrial service companies, especially SMEs, lack specific knowledge, which effects standardization of service provision processes has. In order to illustrate these effects, changes in process variance need to be related to factors, which describe the process performance. Service provision processes are specifically characterized by the integration of the external factor into the value creation process. Thus, process performance indicators need to be selected, which measure process quality from both the customer's—and provider's perspective. Based on Stich [33], Neumann et al. [24] eight process performance indicators were selected and validated by industry experts in a similar manner as the validation of the process characteristics took place. Four indicators respectively focus on internal and on external process quality (cf. Fig. 5).

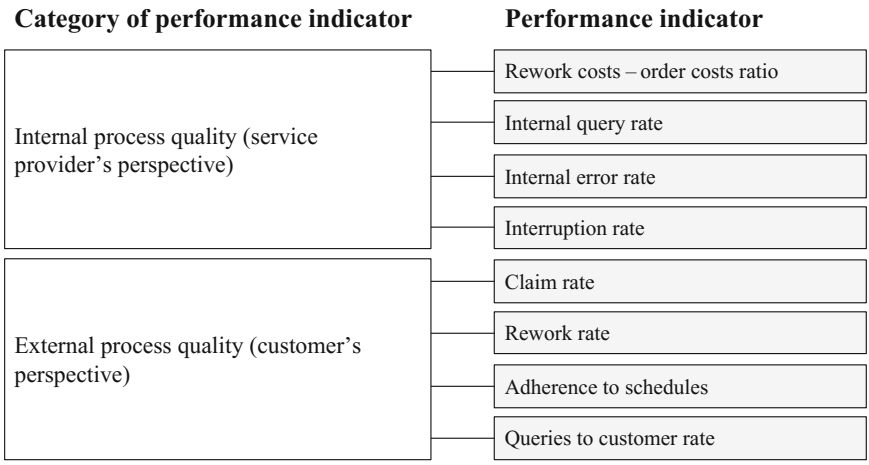


Fig. 5 Process performance indicators

For both the customer's—and the provider's perspective four performance indicators were selected. The internal process quality is determined by *Rework costs − order costs ratio*, *Internal query rate*, *Internal error rate* and *Interruption rate*. The external process quality is determined by *Claim rate*, *Rework rate*, *Adherence to schedules* and *Queries to customer rate*. The following definitions describe the selected performance indicators (Table 1).

Table 1 Definition of performance indicators

Performance indicator	Definition
Rework costs − order costs ratio	Mean value of rework costs divided by total costs of an order
Internal query rate	Mean value of number of unplanned internal queries divided by planned number of queries
Internal error rate	Number of internal errors divided by total number of orders
Interruption rate	Mean value of duration of waiting time divided by order duration
Claim rate	Number of claims by customers divided by total number of orders
Rework rate	Number of orders which required rework divided by total number of orders
Adherence to schedules	Number of order on time divided by total number of orders
Queries to customer rate	Mean value of number of unplanned queries towards the customer divided by planned queries

4.3 Application of Process Characteristics and Indicators in the Industrial Service Industry

Based on the developed process characteristics and process performance indicators the research team performed a questionnaire-based study at a large industrial service provider in Germany. The service provider operates a large chemical business park in which numerous customers from the chemical industry manufacture chemical intermediate and final products. Operating the business park comprises industrial services such as planned and unplanned maintenance, pipeline construction, plant construction, montage, planning and execution of large shutdown, engineering services and management of external contractors. The questionnaire-based study served four goals. The first goal was to investigate whether the developed process characteristics are suitable to measure dispersion in process execution for comparable service orders. The second goal was to examine whether the developed process performance indicators are suitable to measure process performance from both the customer's and the service provider's perspective. The third goal was to investigate which method of multivariate analysis is suitable to analyze the collected data. The final goal was investigate if the assumed relations between a high degree of process variance and a low process performance do in fact exist.

Prior to performing the questionnaire-based study the research object, meaning specific service processes were to be determined. Due to a comparatively low process efficiency and a high repetition rate, maintenance processes were selected as the research object. In order to ensure the comparability of the captured service orders, the research team narrowed the research object down based on seven key process properties, which were identified in cooperation with the industrial service provider. Table 2 illustrates the selected process property specifications.

In cooperation with employees from the industrial service provider, the research team developed a questionnaire in order to capture the specific process characteristics and process performance indicators of individual service orders. In order to capture the variance in process execution for each process characteristic closed questions and ordinal-scaled answers were developed. In order to capture detailed information regarding process characteristics the researchers developed fourteen questions for the eight specified process characteristics. Process performance indicators were captured by closed questions and nominal-scaled yes/no answers. In total fourteen process characteristics questions and eight process performance indicator questions were developed and tested with the service personnel. Figure 6 illustrates excerpts from the questionnaire for both the process characteristics and process performance indicators.

Following, 84 paper-based questionnaires for individual maintenance service orders were collected. The questionnaires were digitized and the questionnaire completion rate was evaluated. Questionnaires with a completion rate less than 70% were eliminated [27]. The remaining 76 service orders were analyzed by the following approach. First, the coefficient of variation was calculated for all process

Table 2 Narrowing of research object based on process properties

Process property	Property specification				
Location of process execution	*Service object at customer's site*	At centralized customer's site		At centralized workshop	Location-independent execution
Use of tools	No tools required			*Tools are required*	
Planning horizon	*Unplanned*			*Planned*	
Responsible maintenance group	*Mechanical group*	Electrical group	Hydraulic group	Instrument and control group	Construction group
Involvement of external contractors	*No involvement*	External contractor as supplier		Parallel joint service provision	Sequential joint service provision
Order duration	2 h or less	5 h or less	*10 h or less*	More than 1 day	More than 5 days
Personnel requirement for order execution	1 employee	2 employees or less	*4 employees or less*	More than 4 employees	More than 10 employees

Process characteristics (*excerpt*)

Quality and form of order information					
Was the description of the service object location correct?	Plant, component, part correct (100%)	Plant & component correct (75%)	Plant correct (50%)	Plant complex correct (25%)	Location incorrect (0&)
To what extend was the given order information correct?	Fully correct (100%)	Mostly correct (75%)	Partially correct (50%)	Hardly correct (25%)	Completely wrong (0%)

Process performance indicators (*excerpt*)

External process quality (customer's perspective)	**Claim rate**	Did the customer file a complaint?	yes	no
	Rework rate	Was any rework necessary?	yes	no
	Adherence to schedule	Was the order time limit exceeded?	yes	no
	Queries to customer rate	Were there queries to the customer required?	yes	no

Fig. 6 Questions for collection of process characteristics and process performance indicators (translated excerpt)

characteristics [7]. Process characteristics questions with a coefficient higher than 0.2 were selected for further analysis since they showed a significant degree of process variation. Second, process performance indicators were calculated as percentages by dividing the number of service orders with the answer "no" by the total number of service orders. Process performance indicators with a deviation higher than 5% from the desired value (e.g. 0% claim rate, 100% adherence to schedule) were selected for further analysis. By that, nine process characteristics questions and six process performance indicators were selected. Third, for all 54 combinations of one process characteristic and one process performance indicator each the multivariate analysis method of contingency analysis was performed. This method was chosen because it allows analyzing nominal-scaled and ordinal-scaled variables

Table 3 Key results of questionnaire-based study

Process performance indicator	Process characteristic question (associated process characteristic)	Significance level	Cramer's V	Contingency coefficient
Queries to customer rate	Means of communication (quality and form of order information)	1.00	0.65	0.55 (max. 0.79)
Interruption rate	Information quality regarding tools (quality and form of order information)	0.98	0.40	0.37 (max. 0.79)
Adherence to schedule rate	Information quality regarding order description (quality and form of order information)	1.00	0.79	0.62 (max. 0.79)
Internal query rate	Information quality regarding tools (quality and form of order information)	1.00	0.53	0.47 (max. 0.79)

regarding statistical correlations by means of the chi square test [21]. The analysis revealed correlations with a significance level higher than 0.90 for ten pairs of variables. In order to evaluate the correlations the Cramer's V coefficient and the contingency coefficient were calculated [2, 14]. Both, the Cramer's V coefficient and the contingency coefficient allow assessing the intensity of a correlation. Cramer's V takes on values between zero and one. A value above 0.3 represents a non-trivial correlation. The maximum values of the contingency coefficient depend on the minimal degree of freedom of the analyzed variables (minimum number of possible values of the analyzed variables). The contingency analyses resulted in four correlations that showed a medium to strong statistical correlation between a high degree of process variance and a low level of process performance. Table 3 illustrates the key results of the questionnaire-based study.

Based on the identified correlations the research team conducted expert interviews with managers from the industrial service provider in order to interpret the findings. The interviews revealed that the identified correlations represent causal relations between the process characteristics and the process performance indicators. This in fact means that for the analyzed processes a high degree of process variance lead to a low degree of process performance.

5 Conclusion

This paper presents three results of an ongoing research project focusing on the effect of process standardization in the European industrial services sector. The first result are process characteristics that allow recording the dispersion in the execution of service provision processes. The second result are process performance indicators that allow recording the internal and external quality of the process execution. The third result is the application of the developed process characteristics and process performance indicators at a large industrial service company. In this paper, the research team could show that the developed process characteristics and process performance indicators are suitable to measure process variance and process performance. Additionally the application showed that collected data could be analyzed by means of the contingency analysis and that correlations between a high degree of process variance and a low degree of process performance could be detected.

With regard to further understanding how process standardization works, which cause-effect relations exist between process characteristics and performance indicators further research is required. Therefore the developed approach to measure process variance and process performance should be applied to additional companies from the industrial services sector.

Acknowledgements This work has been funded by the "Allianz von Forschungsvereinigungen" (AiF) within the program "Förderung der industriellen Gemeinschaftsforschung und -entwicklung" (IGF) by the German Federal Ministry of Research and Technology (support code: 19388N). The authors wish to acknowledge the AiF for their support.

References

1. Abegglen (2012) Lean Management in Dienstleistungsorganisationen: Wie Dienstleister künftige Herausforderungen meistern: ACE—Allied Consultants Europe Studie 2012
2. Backhaus K, Erichson B, Plinke W, Weiber R (2016) Multivariate analysemethoden. Springer, Berlin
3. Ball T (2015) Standardisierte Facility Services für die Industrie. Neue Dienstleistungsmodelle für industrienahe Sekundärprozesse. Facil Manag 3
4. Ball T, Hossenfelder J (2015) e | Site Services: Standorteffizienz mit engineered Site Services: Wertschöpfungssteigerung durch professionelle Service- und Technik-Partner. Themendossier, Kaufbeuren
5. Bogajewkaja J, Jacob F, Michaelis K (1998) Prozeßkostenrechnung im Projektgeschäft - Ein Instrument zum Controlling der Kundenintegration. Business to business marketing, vol 11, Inst. für Allg. Betriebswirtschaftslehre, Weiterbildendes Studium Techn. Vertrieb, Freie Univ, Berlin-Dahlem
6. Braunwarth KS (2009) Wertorientiertes Prozessmanagement von Dienstleistungsprozessen. Available at: http://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/index/index/docId/1563. Accessed 13 July 2016

7. Brown CE (2012) Coefficient of variation. In: Brown CE (ed) Applied multivariate statistics in geohydrology and related sciences. Springer, Berlin (Place of publication not identified), pp 155–157
8. Carlborg P, Kindström D (2014) Service process modularization and modular strategies. J Bus Ind Mark 29(4):313–323
9. Deuse J, Wischniewski S, Birkmann S (2009) Knowledgebase für die kontinuierliche Innovationsarbeit im Technischen Kundendienst. In: Herrmann TA, Kleinbeck U, Ritterskamp C (eds) Innovationen an der Schnittstelle zwischen technischer Dienstleistung und Kunden 2. Physica-Verlag HD, Heidelberg, pp 155–176
10. DIN (2012) Grundlagen der Instandhaltung No. 31051, Beuth Verlag, Berlin. Available at: http://www.beuth.de/de/norm/din-31051/154459920
11. DocuWare (2015) Bilfinger Berger industrial services case study
12. Fabry C (2014) Synchronisation der Dienstleistungsproduktion mittels Takt, Schriftenreihe Rationalisierung, vol 128, 1. Aufl., Apprimus Verlag, Aachen
13. Fabry C, Honné M, Lukas M (2014) Lean services consortial benchmark: successful practices from European Industrial Service Organizations, Aachen
14. Fleiss JL, Levin B, Paik MC (2003) Statistical methods for rates and proportions. Wiley series in probability and statistics, 3rd edn. Wiley-Interscience, Chichester
15. Fleiter T (ed) (2013) Energieverbrauch und CO_2-Emissionen industrieller Prozesstechnologien: Einsparpotenziale, Hemmnisse und Instrumente. ISI-Schriftenreihe "Innovationspotenziale". Fraunhofer-Verl, Stuttgart
16. Fliess S (2009) Dienstleistungsmanagement: Kundenintegration gestalten und steuern. Lehrbuch, Gabler, Wiesbaden
17. von Garrel J, Tackenberg S, Seidel H, Grandt C (2014) Dienstleistungen produktiv erbringen: Eine empirische Analyse wissensintensiver Unternehmen in Deutschland. Springer Gabler, Wiesbaden
18. Honné M (2016) Erklärungsmodell ausgewählter Lean-Prinzipien für industrielle Dienstleistungen: explanatory model of selected lean principles for industrial services, Edition Wissenschaft Apprimus, vol 143. Apprimus Verlag, Aachen
19. Kallenberg R (2002) Ein Referenzmodell für den Service in Unternehmen des Maschinenbaus. Schriftenreihe Rationalisierung und Humanisierung, vol 44. Shaker, Aachen
20. Kasiri LA, Guan Cheng KT, Sambasivan M, Sidin SM (2017) Integration of standardization and customization. Impact on service quality, customer satisfaction, and loyalty. J Retail Consum Serv 35:91–97
21. Kateri M (2014) Contingency table analysis: methods and implementation using R. Statistics for industry and technology, Birkhäuser, New York
22. Matyas K (2010) Taschenbuch Instandhaltungslogistik: Qualität und Produktivität steigern. Praxisreihe Qualitätswissen, 4., überarb. Aufl., Hanser, München, Wien
23. MÍkva M, Prajová V, Yakimovich B, Korshunov A, Tyurin I (2016) Standardization—one of the tools of continuous improvement. Procedia Eng 149:329–332
24. Neumann S, Probst C, Wernsmann C (2012) Kontinuierliches Prozessmanagement. In: Becker J, Kugeler M, Rosemann M (eds) Prozessmanagement. Springer, Berlin, pp 303–325
25. Reichel J, Müller G, Mandelartz J (eds) (2009) Betriebliche Instandhaltung. Springer, Berlin
26. Roland Berger (2010) Industrieservices in Deutschland: Status Quo und zukünftige Entwicklung
27. Roth PL, Switzer FS III (1995) A Monte Carlo analysis of missing data techniques in a HRM setting. J Manag 21(5):1003–1023
28. Schuh G (2006) Change management—prozesse strategiekonform gestalten. Springer, Berlin
29. Schuh G, Fabry C, Jussen P, Stüer P (2013) Mit Lean services Dienstleistungsorganisation wertorientiert und verschwendungsfrei gestalten
30. Schuh G, Honné M, Lukas M (2014) Konsortial-Benchmarking. Lean Services 2014, Aachen
31. Shewhart WA (1931) Economic control of quality of manufactured product. The Bell Telephon Laboratories Series. D. Van Nostrand Company Inc., New York

32. Sihn W, Matyas K (2011) Lean Maintenance. Standardisierung und Optimierung von Instandhaltungsprozessen. In: Biedermann H (ed) Lean maintenance: Null-Verschwendung durch schlanke Strukturen und wertsteigernde Managementkonzepte, Praxiswissen Instandhaltung, TÜV Media, Köln, pp 59–77
33. Stich V (2015) Nachhaltige Effizienzsteigerung im Service: Verschwendungen vermeiden. Prozesse optimieren, innovation, 1st edn. Beuth, Berlin (u.a.)
34. Stüer P (2015) Gestaltung industrieller Dienstleistungen nach Lean-Prinzipien. Edition Wissenschaft Apprimus, Band 132, 1. Aufl., Apprimus Verlag, Aachen
35. T.A. Cook & Partner Consultants GmbH (2013) Maintenance efficiency report 2013: Internationale Studie zur Entwicklung der Instandhaltungseffizienz in der Prozessindustrie
36. T.A. Cook & Partner Consultants GmbH (2014) Jahrbuch Maintenance 2014, 1. Jahrgang, Berlin

Modular-Based Framework of Key Performance Indicators Regulating Maintenance Contracts

Mirka Kans and Anders Ingwald

Abstract Key performance indicators (KPI) are necessary for regulating maintenance performance, setting goals as well as for follow up and improvement. Several standards and models for measuring maintenance performance exist today, but these are mainly developed for in-house maintenance. For outsourced maintenance, which is regulated in a service contract, other kinds of KPIs are needed. The procurement of maintenance and contract forms is also changing; an alternative to the traditional maintenance contract based on fixed price and predetermined activities are the performance-based contracts. Cooperation contracts based on mutual trust and fairness for all parties are also available. New KPI models for regulating maintenance service contracts are therefore needed. In this paper, a KPI framework for maintenance contracts based on the concept of modular maintenance offerings is proposed. Modular maintenance offerings is a way to classify maintenance services offerings with increasing integration of the offering, and increasing focus on utility for the customer and the business ecosystem. The KPI framework proposes indicators for regulating contracts on three levels (resource, performance and utility level) and includes six categories of indicators: economic, technical, organisational, quality, safety and health, and relationship between actors.

1 Introduction

Maintenance key performance indicators support the management in achieving excellence and high asset utilisation, SIS [22]. These indicators (KPI) reflect the hard, measurable, aspects as well as the soft aspects that are more difficult to measure. Hard indicators include for instance cost aspects, time measures, and

M. Kans (✉)
Department of Mechanical Engineering, Linnaeus University, Växjö, Sweden
e-mail: mirka.kans@lnu.se

A. Ingwald
Department of Physics and Electrical Engineering, Linnaeus University, Växjö, Sweden
e-mail: anders.ingwald@lnu.se

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_30

technical performance, and are typically quantitative of ordinal, interval or ratio scales. Soft indicators normally use nominal or ordinal scales to measure dimensions such as relationships or quality and are extracted from textual descriptions rather than from numerical figures. Even typical quantitative values can sometimes be difficult to measure because of low data quality or lack of accurate data. Åhrén and Parida [2] for example found that only 38% of the total of 137 KPIs for measuring the effectiveness of railway infrastructure maintenance could be verified using available data. Economic data and safety data was easiest to access, while data on quality, environment and labour were difficult to obtain. In the development of key performance indicators it is thus important to address data quality and data availability issues, Al-Najjar and Kans [4]. It is also important to define a relevant set of data, depending on the situation and decision context, Kans and Ingwald [9].

A number of models describe key performance indicators for maintenance. The standard *EN 15341:2007* includes economic, technical and organizational indicators for monitoring maintenance performance on the strategic, tactical and operational business levels, SIS [22], Komonen [11] points out that key indicators at different levels are interlinked and connected to overall strategy. The author therefore describes a framework of indicators for monitoring the maintenance, with indicators describing business impact in focus grouped according to effects on production, quality of production and overall equipment effectiveness. In a comprehensive review Kumar et al. [13] describe the state-of-the-art of maintenance related key performance indicators and classify these into two main categories: process-related indicators and performance-related indicators. Process-related indicators are connected to the core maintenance work while the performance-based indicators are related with the outcome of the maintenance work.

The previous research on KPIs for regulating maintenance contracts is rather extensive, but holistic models such as the ones described above are missing. This paper therefore focuses on indicators for regulating maintenance contracts. While the general maintenance performance indicators to some extent are applicable also for regulating maintenance contracts, others are context specific and require different logics than for in-house maintenance. The standard *EN 13269:2006* contains a checklist of key elements for the development of maintenance contracts and examples of factors that affect the contract [21], but these are not classified neither grouped with respect to contract type. The lack of classification of KPIs in the standard makes it hard to know which indicators are suitable for different types of contracts. Some indicators are the same regardless of contract type, while others are tightly connected to the agreement type. The purpose of this paper is therefore to develop a KPI framework for maintenance contracts based on the concept of modular maintenance offerings, described in the next section.

2 Maintenance Contracts and Modular Maintenance Offerings

Maintenance services are regulated in agreements and contracts. Resource-based contracts are cost based, i.e. the supplier is compensated for direct resources used such as personnel costs and spare parts, and a profit margin. Resource-based contracts could be in form of holistic contracts covering all maintenance activities for a particular system, Norden et al. [16]. In performance-based contracts, the regulation is based on a predefined performance, Abdi et al. [1]. The supplier guarantees a certain level of performance and the supplier, in addition to coverage of direct and indirect expenses, is compensated for the increased risk taken, Lawther and Martin [14]. Performance-based contracts often cover total maintenance needs to ensure uptime or function, either as service contracts or leasing contracts, where the supplier owns the system and leases the function or the system to the customer. Advanced form of performance-based contracts guarantee the customer's ongoing operations. In these utility-based contracts, the supplier takes a greater risk, in that it ensures the business performance and not only a specific system's or product's operation. This risk must carefully be regulated in the contract, and be soundly priced. Lately, extensive research has been conducted in the area of performance-based and utility-based maintenance contract forms; see for instance Famurewa et al. [7], Sinkkonen et al. [20], Abdi et al. [1], Lawther and Martin [14], Lieckens et al. [15].

In contrast, the resource-based contract form is still most commonly used in practice. One reason for this is the complexity and risk, either real or perceived, that comes with other types of contracts. Another is the need for a mental shift from product-centration to utility-thinking that more advanced contract forms require, Kans and Ingwald [10].

The concept of modular maintenance offerings depicted in Fig. 1 applies ecosystem reasoning on the maintenance contracts, and an integrative approach inspired by the product-services systems concept, see Kans and Ingwald [10]. Modular maintenance offerings is a way to classify maintenance services offerings with increasing integration of the offering, and increasing focus on utility for the customer and the business ecosystem. The value proposition is based on the core competencies of the service provider expressed as key activities and resources (for a more elaborate discussion of key competencies, see [18]). The basic value proposition is maintenance as a resource, based on the supplier's maintenance skills, planning skills and spare parts management. Resources in form of tools and equipment also affect the content of the offering. The next level considers maintenance from a performance perspective, i.e. connected to the system functionality. Key competencies for performance-based contracts are, in addition to the ones for resource-based contracts, the ability to monitor and predict system behaviour. External expertise, e.g. oil analysis, and speedy spare parts delivery, are examples of additional competencies the supplier might acquire for reaching the defined performance standards. Utility-based contracts focus on the benefits the

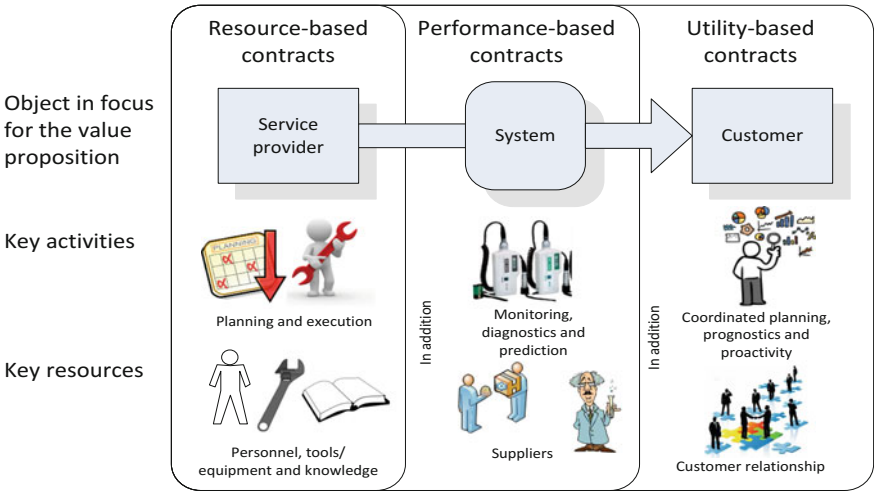


Fig. 1 Modular maintenance offerings

maintenance offering provides to the customer and is thus linked to the customer's results. Key competencies are, in addition to the ones described above, coordinated planning of the production and maintenance for maximum benefit, i.e. an asset management rather than a maintenance perspective. The relationship to the customer becomes an important resource—without willingness to share operational data and production plans, the utility goal might be at risk. The relationship is characterised by mutual benefits, for instance in the form of profit sharing mechanisms. One maintenance service provider can offer contracts of all three types simultaneously to different customer segments, if the contract portfolio is soundly built upon the key competencies of the provider.

3 Key Performance Indicators for Regulating Maintenance Contracts

Maintenance contracts, as described above, differ with respect to business logic and must thus be regulated using different types of indicators. In this section the nature and type of indicators used for respective contract type is described.

3.1 Indicators for Resource-Based Contracts

In resource-based contract the execution of maintenance is in focus. What to be done where, how, when and why this is done should be indicated, and the

objectives must be measurable, SIS [21]. You should also determine the frequency of activities, reporting time and end time, and how changes in the agreed plans are managed. The key indicators are therefore directly linked to what is done, the time consumed, and by whom maintenance is carried out using which resources. Key economic indicators are directly related to the costs incurred for the execution, such as maintenance costs relative to replacement cost, which are connected to pricing strategies such as lump sum, hourly-based or performance-based pricing (performance of the maintenance, not system performance). A number of indicators in the EN 15341:2007, especially on strategic and tactical levels, can be used for follow-up of maintenance contracts, SIS [22]. For resource-based contracts the indicator T1, which describes maintenance related availability or E1, total maintenance costs relative to replacement value, are suitable.

3.2 Indicators for Performance-Based Contracts

Performance-based maintenance contracts the parties need to regulate both the pricing and the technical performance of the system, Abdi et al. [1]. In addition to direct maintenance costs, the customer pays for the risk transfer that takes place when the supplier guarantees a certain level of performance. Incentives for the supplier are to find an optimal balance; not over-maintain, as it directly affects the cost, but also not to under-maintain, as it leads to fines. In the first hand, the variable costs are affected, but fixed costs can also be affected because the total resource utilization must be balanced. The customer can benefit from performance-based contracts through reduced maintenance costs, but also in form of extended system lifetime, Damnjanovic and Zhang [5]. The purchaser of maintenance can follow up maintenance cost (what the customer pays) in relation to availability, while the seller in addition to traditional fixed and variable maintenance costs also must have indicators reflecting the risk transfer margin, i.e. the margin the maintenance provider requires in order to cover costs in the event of maintenance-related production disruptions. The opposite risk transfer strategy is also possible, i.e. that the buyer takes all the performance based risk, in form of availability payments, Lawther and Martin [14]. Performance is controlled using key technical indicators reflecting the quantity and, if available, the quality, of performance, which is directly linked to the system's function. It is therefore important to specify which level of functionality to obtain. Key indicators that reflect the uptime or availability are often used, but the definition varies depending on situation and context, such as *total time*, i.e. the time the system is switched on whether it is used or not, *effective hours* when the system's engine is running, or *engine hours* when the system is running (in motion), Kowalkowski and Kindström [12].

The consequences of functional failure is normally regulated in the form of fines or other penalties, based on simple logic such as met/not met, or by means of a scale, where the penalty depends on the length or extent of the lost function, Lawther and Martin [14]. Decisions regarding the pricing and performance are

interlinked: the supplier can ensure high uptime (or other type of performance) if sufficient resources are available when needed, which in turn depends on factors such as inventory management and the capacity maintenance of the organization (such as availability of technicians), Lieckens et al. [15]. Consequently, the supplier needs internal key indicators describing the capacity of the maintenance organization. Lieckens et al. [15] suggest that the contract is determined on the basis of price, performance (in form of guaranteed uptime), and maintenance intervals, while maintenance capacity is determined by the capacity of the technicians, inventory levels, and, if the possibility of renovating components exists, also the location and capacity of repair facilities. The link between technical and economic perspectives can also be reflected in the form of key indicators, such as maintenance related availability in relation to the total maintenance costs, SIS [22] or the annual cost of maintenance per performance unit, Shohet [19]. While the quantitative performance aspects are the focus of the performance-based contracts, key indicators reflecting quality could be used as well. Qualitative dimensions are according to Grönroos [8]:

- Tangibles—the natural way of presenting oneself
- Reliability—supply security
- Responsiveness—willingness to help the customer
- Assurance—expertise and trust
- Empathy—consideration and individualization

3.3 Indicators for Utility-Based Contracts

Key indicators for these types of contracts have to reflect the benefit of maintenance services for the customer, or the customer's customer. This is reflected either as qualitative business impact or in economic terms. Al-Najjar and Alsyouf [3] connects maintenance with business impact using indicators such as cost of rework of defective products, penalties due to incorrect delivery (low quality, late delivery, etc. which can be related to maintenance), and environmental costs. In Sinkkonen et al. [20] a cost model for partnering projects is presented, covering the cost categories operating costs, sub-contractor costs, machinery and tools, spare parts costs, logistics costs, environmental costs, quality costs, and other costs. The model can be used distribution of project cost/profit. Moreover, the ability to measure and assess quality is important for utility-based contracts. The quality aspects could be qualitative, such as in the study of the client end user (i.e. the residents) perceptions of property maintenance reported in Straub [24]. The client wanted delivery assurance and flexibility in the offer; the supplier should possess broad competence and a wide range of offerings, and the ability to lead and coordinate all maintenance services, including subcontractors. In addition, the attitude and behaviour of the service provider was seen as important; the supplier is the client's face in the maintenance context and must comply with applicable codes, acting considerate

towards the residents and being regarded as proper and serious, including proper equipment and clothing. The residents believed that the quality of execution, the skills of maintenance workers and that the maintenance activities are not prolonged are the most important quality factors. Quantitative quality indicators could also be used, such as the quality indicator for railway tracks proposed in Stenbeck [23]. The indicator reflects the passenger quality experience, expressed as minimal delay and maximum comfort compared to maintenance cost. An interesting finding from this study is that the quality was raised in the case studied without cost increase. The maintenance provider made no technical improvements during the period, so the quality improvements seem to be connected to organizational and human factors.

Famurewa et al. [7] suggest key indicators for utility-based contracts in railway infrastructure, divided into the areas of security, economy, technology and customer. These key indicators reflect both the direct customer benefit and customer's customer, in this case passengers. Moreover, the indicators are linked to the national objectives set for railway transportation. The other goals are relating to safety, environment and health. While traditional contracts mainly are focused on hard targets the collaborative contracts focus on both hard and soft goals, Olsson and Esping [17]. Utility-based contracts postulate trust and closeness between the supplier and the customer, and should thus be regulated using both quantifiable technical and economic key indicators as well as qualitative key indicators that regulate e.g. quality aspects and the relationship. In a pilot project on partnering for railway infrastructure described in Esping and Olsson [6] following overall key indicators were used:

1. Costs (economic indicator)
2. Train delays (technical indicator)
3. Number of corrective actions (technical indicator)
4. Number of errors reported during inspections (technical indicator)
5. Track Quality Index (technical indicator)
6. Relationship between client and supplier (soft indicator).

4 A Modular-Based Key Performance Indicator Framework for Maintenance Contracts

Reviewing the literature on indicators for maintenance contracts, it seems like it would be hard to find any commonalities between different contract forms. In this section, an attempt to bring some structure is made utilising the concept of modular-based maintenance offerings. The basic assumption is that a maintenance service provider could offer different types of contract forms according to customer requirements. The main differences between contract types is found in the economic and technical indicators, as the different contract types apply different business logic. These differences also affect the definition of quality, while other factors such

as effects on safety and health are the same disrespect of contract type. In the framework, see Table 1, six performance indicator categories are included: economic, technical, organisational, safety and health, and finally quality of performance and relationship. The performance quality covers quantitative measures while relationship quality is assessed using in first hand qualitative measures.

This means that for each specific agreement, indicators and how they should be measured have to be defined. Thus, Table 1 lists a number of generic indicator types, while the specific indicators reflecting these differ. This was evident when applying the framework in two case studies, one within railway maintenance and

Table 1 Key performance indicator framework for maintenance contracts

Category	Resource-based contracts	Performance-based contracts	Utility-based contracts
Economic	- Total cost during the contract period - Costs per maintenance object - Penalty for backlog	- Total cost during the contract period - Costs in relation to system availability - Penalty for lost performance	- Total cost during the contract period - Costs in relation to production output - Penalty for lost production
Technical	- Agreed maintenance performance (activities, level of execution, frequency)	- Agreed system performance (functionality, reliability)	- Agreed production output (productivity, reliability)
Organisational	- Maintenance organisational capacity (capacity of personnel, inventory level, repair facilities)	- Maintenance organisational capacity	- Maintenance organisational capacity
Safety and health	- Maintenance related incidents - Maintenance related accidents - Maintenance related environmental impact	- Incidents - Accidents - Environmental impact	- Incidents - Accidents - Environmental impact
Performance quality	- Mean waiting time (MWT) - Mean time to repair (MTTR) - Backlog - Delivery assurance[a] in terms of performed maintenance	- MWT - MTTR - Delivery assurance[a] in terms of performance realisation	- MWT - MTTR - Delivery assurance[a] in terms of utility realisation
Relationship quality	- Service quality (tangibles, responsiveness, assurance, empathy)	- Service quality - Number of relationship contacts and interfaces	- Service quality - Number of relationship contacts and interfaces

[a]Service reliability

one within power generation. The KPI models developed for each case followed the framework proposed, but with case specific adaptations. For instance, the technical performance for railway maintenance is defined as availability of trains according to agreement, while for the power industry performance of the turbine according to agreement is measured. The two cases also differed with respect to performance and relationship quality indicators. An interesting reflection is that the level of adaption increased from the resource-based to the utility-based contract form.

5 Conclusions

In this paper, a KPI framework for maintenance contracts based on the concept of modular maintenance offerings was proposed. The framework supports a holistic understanding of performance indicators for regulating maintenance contracts, and builds upon the premise that different contract types could be logically built up on the core competencies of the maintenance service provider, and with increasing focus on the extended business ecosystem (i.e. moving from the supplier perspective via the system to be maintained to the customer needs, and even the customer's customer). Moreover, the framework includes both hard and soft measures for the regulation of contracts. Focusing not only on the quantifiable, but also on how the relationship between the provider and the customer should be regulated, is needed for successful implementation of maintenance contracts, and especially for utility-based contract forms. The framework has been utilised in the project Future Industrial Services Management, together with the concept of modular-based maintenance offerings, as a basis for developing strategies for maintenance contracts within two case studies, but further studies are required for testing and verifying the concept in real industrial settings.

Acknowledgements This research is partly founded by the Swedish innovation agency Vinnova within the project Future Industrial Services Management.

References

1. Abdi A, Lind H, Birgisson B (2014) Designing appropriate contracts for achieving efficient winter road and railway maintenance with high performance quality. Int J Qual Serv Sci 6 (4):399–415
2. Åhrén T, Parida A (2009) Maintenance performance indicators (MPIs) for benchmarking the railway infrastructure. Benchmark Int J 16(2):247–258
3. Al-Najjar B, Alsyouf I (2004) Enhancing a company's profitability and competitiveness using integrated vibration-based maintenance: a case study. Eur J Oper Res 157(3):643–657
4. Al-Najjar B, Kans M (2006) A model to identify relevant data for problem tracing and maintenance cost-effective decisions. Int J Prod Perform Manag 55(8):616–637
5. Damnjanovic I, Zhang Z (2008) Risk-based model for valuation of performance-specified pavement maintenance contracts. J Constr Eng Manag 134(7):492–500

6. Esping U, Olsson U (2004) Part II. Partnering in a railway infrastructure maintenance con-tract: a case study. J Qual Maint Eng 10(4):248–253
7. Famurewa SM, Asplund M, Galar D, Kumar U (2013) Implementation of performance based maintenance contracting in railway industries. Int J Syst Assur Eng Manag 4(3):231–240
8. Grönroos Christian (2008) Service management och marknadsföring: kundorienterat ledarskap i servicekonkurrensen. Liber, Malmö
9. Kans M, Ingwald A (2008) Common database for cost-effective improvement of maintenance performance. Int J Prod Econ 113:734–747
10. Kans M, Ingwald A (2016) A framework for business model development for reaching service management 4.0. J Maint Eng 1:398–407
11. Komonen K (2002) A cost model of industrial maintenance for profitability analysis and benchmarking. Int J Prod Econ 79(1):15–31
12. Kowalkowski C, Kindström D (2012) Tjänster och helhetslösningar: nya affärsmodeller för konkurrenskraft. Liber, Malmö
13. Kumar U, Galar D, Parida A, Stenström C, Berges L (2013) Maintenance performance metrics: a state-of-the-art review. J Qual Maint Eng 19(3):233–277
14. Lawther WC, Martin L (2014) Availability payments and key performance indicators: challenges for effective implementation of performance management systems in transportation public-private partnerships. Publ Works Manag Policy 19(3):219–234
15. Lieckens KT, Colen PJ, Lambrecht MR (2015) Network and contract optimization for maintenance services with remanufacturing. Comput Oper Res 54:232–244
16. Norden C, Hribernik K, Ghrairi Z, Thoben K-D, Fuggini C (2013) New approaches to through-life asset management in the maritime industry. Procedia CIRP 11:219–224
17. Olsson U, Esping U, Part I (2004) A framework of partnering for infrastructure maintenance. J Qual Maint Eng 10(4):234–24719
18. Osterwalder A, Pigneur Y (2010) Business model generation: a handbook for visionaries, game changers, and challengers [electronic resource]. Wiley, Hoboken
19. Shohet IM (2003) Key performance indicators for maintenance of health-care facilities. Facilities 21(1/2):5–12
20. Sinkkonen T, Marttonen S, Tynninen L, Kärri T (2013) Modelling costs in maintenance networks. J Qual Maint Eng 19(3):330–344
21. SIS (2006) SS-EN 13269:2006, Maintenance—guidelines for the preparation of maintenance contracts
22. SIS (2007) SS-EN 15341:2007, Maintenance key performance indicators
23. Stenbeck T (2008) Quantifying effects of incentives in a rail maintenance performance-based contract. J Constr Eng Manag 134(4):265–272
24. Straub A (2010) Competences of maintenance service suppliers servicing end-customers. Constr Manag Econ 28:1187–1195

Smart Asset Management for Electric Utilities: Big Data and Future

Swasti R. Khuntia, Jose L. Rueda and Mart A. M. M. van der Meijden

Abstract This paper discusses about future challenges in terms of big data and new technologies. Utilities have been collecting data in large amounts but they are hardly utilized because they are huge in amount and also there is uncertainty associated with it. Condition monitoring of assets collects large amounts of data during daily operations. The question arises "*How to extract information from large chunk of data?*" The concept of "*rich data and poor information*" is being challenged by big data analytics with advent of machine learning techniques. Along with technological advancements like Internet of Things (IoT), big data analytics will play an important role for electric utilities. In this paper, challenges are answered by pathways and guidelines to make the current asset management practices smarter for the future.

1 Introduction

Asset management, also described as mid-term planning, forms an important activity in the electric transmission and distribution system, the other two being long-term planning (or grid development) and short-term planning (or system operation) [4]. Under asset management, maintenance of physical assets and devising maintenance policies form the crucial part. For a recent detailed study on asset management, ref. [5] is recommended. The aims of asset management are to optimize the asset life cycle, improve predictive maintenance and prepare an efficient business plan for investments on new assets. And, with the recent advancements in technology when data and communication play important roles, asset management has evolved from a more traditional to a smarter decision making

S. R. Khuntia (✉) · J. L. Rueda · M. A. M. M. van der Meijden
Department of Electrical Sustainable Energy, Delft University of Technology, Delft, The Netherlands
e-mail: s.r.khuntia@tudelft.nl

M. A. M. M. van der Meijden
TenneT TSO B.V., Arnhem, The Netherlands

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_31

process. A two-way communication is built, i.e., mechanisms to process the data and the actions taken on the basis of data that may eventually lead to some form of smart function is achieved. Connectivity and data integration are important and necessary as well but not yet sufficient for a smarter asset management. This can be achieved by designing better information management systems that cannot only handle data archiving and retrieval but also help towards data analysis tools.

Using efficient predictive maintenance strategies, utilities can make smarter decisions about when and where maintenance should be performed, which results in reduced maintenance costs with better planning. Moreover, with the advent of advanced computational tools and smart meters, a huge amount of data is collected by the utilities that can be used later for improving the performance of commissioned assets and/or enhancing maintenance policies. Utilizing condition monitoring data for reliability estimation has some similarities with fault diagnosis. The severity of a fault is a key issue for diagnosis, as it presents insight into how long the asset performs before it fails. Identification of the severity of faults is analogous to the estimation of asset reliability. Data mining is useful in this type of study where the aim is to train mathematical failure models and later use them for extended studies like reliability and maintenance optimization. Different data mining techniques serve different purposes, each offering its own advantages and disadvantages. When most of the utilities are trying to follow the condition based preventive maintenance, it is obvious they are moving towards a data centric asset management. Condition-based maintenance of asset health lends itself to large quantities of data. With the use of big data techniques, utilities have to find their way to effectively uncover the patterns, trends, and subtle connections that lay at the root of asset status and component failures. It is important to emphasize that utilities collect large amounts of data every year, but it is not efficiently used because of a number of reasons starting from unavailability of valid mathematical techniques to score of bad/missing data. This work will focus on data collection for asset condition monitoring and the challenges of big data in maintenance including capturing, accessing, and processing information.

The rest of this paper is organized as follows: Sect. 2 describes the condition monitoring in asset management. Section 3 presents predictive analytics and big data in asset management, how technologies like Internet of Things (IoT) and cloud computing can change the way asset management is studied. This study is concluded in Sect. 4 with recommendations and discussions on future smart asset management.

2 Role of Data in Asset Condition Monitoring

Asset condition monitoring includes data collection, condition detection, asset audit, diagnosis, and decision making. It is, therefore, important to focus on collecting right data and improving data quality at the same time. Condition monitoring systems are becoming financially attractive, and may even come built-in in

the future assets purchased by the utilities (no matter whether they want it or not). However, asset managers and planners still lack straightforward strategies or frameworks to know what information is required in guiding the selection of an appropriate condition monitoring regime. The data requirement for condition monitoring can be divided into two categories: static and dynamic data. Static data refers to asset data that defines the asset itself in normal operating condition. As the name suggests, it does not change during the asset life cycle and is the registered data type often captured. Examples of static data include asset datasheet, location of installation and installation data. Dynamic data refers to asset data that is recorded during the life cycle of asset, mostly operational and maintenance data. This type of data changes during the asset life cycle. Examples are failure rate, maintenance history and other diagnostic data. Due to their volatile nature, dynamic data is often difficult to collect in terms of quantity and quality. So, a rich history of data is often required to perform any kind of maintenance activities. It might be expensive to collect the data as compared to static data, but the advantages overcome the cost.

The interpretation of data coming from condition monitoring systems, the reliability mismatch of diagnostic systems with the equipment being monitored and the volume of data (big data challenge as discussed in Sect. 3) damped the application of condition monitoring systems. Another important issue is the timeliness with which the acquired condition data can be provided and the relationship with the time to failure of this specific asset [1]. For instance, it is vital to differentiate between condition indicators and health indices though both provide information about asset failure rate. While condition indicators relates to more real-time data indicator, health indices generate information in monthly or yearly time-span. Nowadays, most of these challenges remain and form an obstacle for large scale applications of condition monitoring, especially in combination with the costs for setting up a condition monitoring program.

With ambitious project like GARPUR [2], the concept of smart monitoring is introduced which aims to convert condition monitoring of data collected by sensors to visual diagnosis information automatically based on signal processing, signal classification and knowledge rules or expert-based diagnostics systems. To illustrate this, consider the example of partial discharge. Partial discharge is one of the widely used diagnosis methods in power industry to assess cable conditions and can provide an early warning of cable components. It enables maintenance to plan repair and replacement of cables to be carried out timely. For many years, incipient partial discharge faults in power cables have been identified through offline investigation techniques. In the smart grid era, online monitoring will be required for strategically important cable circuits to allow proactive asset management of the cable network to be carried out. Accurate diagnosis of cable conditions, till date, relies on knowledge rules which are based on intensive analysis by human experts. This expert knowledge has, so far, been exclusively obtained from offline partial discharge tests. The data acquired, and the rules derived, are often not applicable to online partial discharge monitoring because emphasis of off-line and on-line are substantially different [10]. As a result, using online smart monitoring for automatically performing (or deriving or producing or similar) accurate diagnosis can

save labor cost and eliminates errors caused by human. However, to diagnose possible faults by partial discharge, there are some challenges discussed in ref. [7], like:

- developing diagnostic techniques to identify partial discharge signals accurately,
- developing pattern recognition techniques to classify fault signals,
- mining large volume of data and transmitting the data from front-end processor to the control center,
- identifying and localizing possible faults when partial discharge activity occurs in cables.

The above discussed techniques are essential to enable smart monitoring by recognizing partial discharge and localizing faults in an autonomous manner. Only by the identification of features, condition monitoring can be applied as smart monitoring in asset management decision making. By generating asset health features or also called health indices in a much more timely and automated fashion, utilities will have better visibility into the overall health of their assets. Health indexing helps in bridging the gap between short-term corrective work driven by condition-based maintenance, and longer-term capital planning.

In such cases, utilities have to find methods that will allow them to assess and monitor the condition of the assets in order to prolong their lifetimes (where possible) because they cannot afford (they do not have the resources and capacity) to refurbish/reinvest the whole network at the same time. Advent of new technologies like Internet of Things (IoT) and deep learning in machine learning will bridge the gap between short-term corrective work driven by condition-based maintenance, and longer-term capital planning, as a result of which acceptance of condition monitoring will play an important role. With advances in data mining and machine learning techniques, deep learning is gaining popularity in the field of asset management and condition monitoring. Deep learning, a branch of machine learning, differs from machine learning in many forms, such as large amount of training data equipped with high performance hardware. Larger the data volume, more efficient the process is. It uses a hierarchical approach of determining the most important characteristics to compare. Based on learning multiple layers of neural network structures, deep learning's success in areas of language modelling, speech and image recognition has made researchers to focus on its application in asset management. One of the advantages of deep learning over various machine learning algorithms is its ability to generate new features from limited series of features located in the training dataset. A deep learning algorithm comprises of two main parts: training and inferring, which is quite similar to neural networks. The training part involves labeling large amount of data from condition monitoring and extracting the right features while inferring part refers to memorizing the right features to make correct conclusions when it faces similar data next time. This is unsupervised learning as compared to supervised learning with machine learning. It helps asset managers in saving significant time on working with big data while achieving concise and reliable analysis results because it facilitates the use of more

complex set of features than the traditional machine learning techniques. Deep learning has not been widely applied to asset health management or prognostics health management field though some attempts have been made in the past. Reference [8] applied a deep learning classification method based on deep belief networks to diagnose electric power transformer health states.

Devising a three-steps methodology can help in leveraging IoT and analytics technologies for improving data collection and component maintenance performance:

1. Collect observational data to assess component condition, defining thresholds or rules to initiate actions or notification to enable condition-based maintenance.
2. Analyze historical data, as well as component failure and work order history, to uncover new patterns that can aid predictions of component failure.
3. Leverage component condition using analysis tools to assess the economic, safety, environmental and public relations effects of failures while also analyzing alternative strategies for handling assets (repair, replace, load shifting, run to fail and so forth).

Though getting to the third step may be the ultimate goal, utilities can derive significant benefit from taking even a first step toward leveraging IoT for condition-based maintenance.

3 Predictive Analytics and Big Data for Utilities

Predictive analytics in predictive maintenance is used to make predictions about unknown future events. In general, analytics refers to way of interpreting data elements and communicating their meanings to the appropriate users. The introduction of data analytics in the form of data mining, statistical modeling and machine learning techniques help in better future prediction. Of course, it is feasible only if high quality data is recorded/produced on its assets. For instance, depending on utilities, data quality can vary a lot, and consolidating the right data can be difficult. This is because utilities often divide responsibilities for assets among departments that have their own information and communication technology (ICT) systems. Separate ICT systems need to be integrated; so all the asset data they generate feed into a single warehouse. These data also need to be structured so they can be used seamlessly in asset-health indexes, asset-criticality assessments, and asset-management decision models. External databases can help fill gaps in an utilities' own data. The amount of data is going to be huge, as it is commonly called "big data". Big data refers to data sets that are so large and complex they are not easily manipulated using the commonly available database tools. Some of the already big data technologies that have been developed or under development are cloud computing, Internet of Things (IoT), granular computing, and quantum computing. IoT can be seen as a major breakthrough towards future improvements in asset management.

3.1 Role of Internet of Things (IoT) in Future Asset Management

Electric utilities have to be proactive in the future because (at least some) will have a tremendous job of dealing with an 'old' network while building the new one. In such cases, utilities have to find methods that will allow them to assess and monitor the condition of the assets in order to prolong their lifetimes (where possible) because they cannot afford (they do not have the resources and capacity) to refurbish/reinvest the whole network at the same time. Internet of Things (IoT) paves way for such a transition.

Talking of Internet of Things (IoT), it is not a single technology and rather a concept in which most new things are connected and enabled. It will pave way for big data analysis and performing operations that were impossible to find without IoT. Some of the advantages why IoT would be beneficial for asset managers and planners include:

- **Requires less human interaction and dependency**: IoT controls the data workflow and can work with less human interactions. It will require less humans and cost to manage the workflows as most of the system will be automated.
- **Better maintenance schedule**: IoT can be used to maximize uptime reliability to ensure that orders can be fulfilled quickly. Maintenance makes it easy to schedule downtime, and maintenance can be scheduled around by proper planning thereby avoiding missed deadlines.
- **Easy tracking of assets**: A program starts with a central database that keeps track of work orders, asset conditions and other data needed to make decisions regarding the management of assets. It is very much necessary for the business to know from where the asset is coming in order to cut the cost. With access to asset data where asset managers can track investment on asset purchase and maintenance, IoT will facilitate in easy handling of resources towards maintenance or investment on new assets.

As the IoT concept enables us to obtain a great amount of data to monitor physical assets, there is an increasing demand for determining asset health conditions in a variety of industries. Accurate asset health assessment is one of the key elements which enable predictive maintenance strategy to increase productivity, reduce maintenance costs and mitigate safety risks.

Smart metering devices (e.g., PMUs, WAMs, SCADA devices) are being installed widely by utilities to collect data as data becomes increasingly important in supporting organizational decisions together with the decision making process. In such case, they are now continually generating more data than at any other time. More data, however, does not necessarily mean better information, or more informed decisions. In fact, many are finding it difficult to use the data. Study reveals that more than 70% of generated data is never used [6], and also suggests that bad data is worse than no data at all. In other cases, at a needed time, the right data is unavailable or in a different format which makes it difficult for analysis

purpose. This situation can be termed as "more data and less information". In today's era, it is just not data anymore, rather big data which is fueling up utilities and it needs more research. There are many major obstacles to the development and implementation of big data analytics in electric power systems, namely, the lack of innovative use cases and application proposals that convert big data into valuable operational intelligence. This repeats the above situation of "more data and less information." More challenges are discussed in next section.

With IoT and big data comes data security and governance into picture. The risk of failing to adhere to data privacy and data protection standards or the advent of data security is a big concern. With big data and advent of new technology, data security and piracy needs to be taken care of. Data security has never been much of an issue for utilities because the infrastructure concerned was hard to access and the potential rewards for hackers were low. Situation has changed lately, proof is that the Ukrainian power grid was hacked twice in 2016 and 2017 through the data collecting devices that system operators need at control centers [9]. In the future, as smart homes get connected to transmission network via the distribution grid, it will be easier for hackers to access the power network. Hence, adequate cyber security measures are needed to counter attack cyber-attacks. Cyber security is a huge and evolving challenge for the bulk power system. Progress in terms of efficient methodologies is observed in the last decade and more is expected in future.

3.2 Challenges in Developing Big Data Applications

The challenges in developing big data applications for asset management can be divided into two stages. The first stage is to design a flexible system architecture that accommodates and optimizes big data analytic workloads (ICT side). The second one is to develop scalable mathematical tools capable of processing distributed data (statistical analysis side). The challenges can be further simplified into three types, namely:

1. **Data challenge**: Data collected in power systems suffer from three primary issues. (1) They are incomplete in nature. (2) They come from heterogeneous sources and therefore are difficult to merge. (3) Systems update or make their data available at different rates. Heterogeneity in power system data exists because often the data was intended for a specific application and not collected for a holistic purpose.
2. **Computational scalability**: Traditional mathematical methods are not adequate in handling the inherently large size, high-dimensional, distributed data sets in situ. To address high-dimensionality, machine learning, statistics and optimization algorithms such as classification, clustering, sampling, and linear/nonlinear optimization algorithms need to be easily scalable. Alternatively, scalable and flexible dimension reduction techniques are needed to extract latent

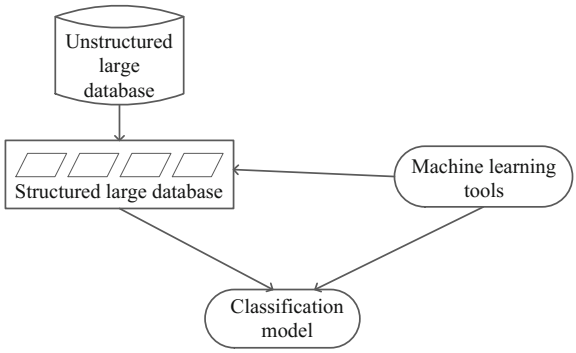


Fig. 1 Supervised learning methodology for big data

features and relevant subsets while balancing accuracy and degree of reduction according to user specification. Large scale data sets in power transmission systems tend to be inherently heterogeneous and distributed. Insufficient research on big data analytics system architecture design and advanced mathematics for large amounts (say, zettabytes of data) to find hidden patterns from the data is an existing hurdle for utilities. To efficiently analyze the distributed data, the algorithms need to come to the data rather than moving data set to the algorithm. Figure 1 shows how unstructured big data can be used to extract classification models for appropriate use cases. In general, such methodology models tend to have two separate steps such as feature extraction and prediction. The first step is to extract features that are indicative of failure or degradation from the data. The second step is to build a prediction model to assess or predict the health condition. This two-step approach involves two separate optimization procedures, which often requires the iteration of the two separate procedures until any acceptable result is achieved.

3. **Making it tractable**: After tackling the scalability issue, the last one is combating tractability. The major classes of data in the asset management decision making process are: (1) failure statistics data, (2) data on the measured conditions of the assets on the transmission network, (3) other data such as weather, load forecasts; (4) asset management data (when and how equipment are maintained and refurbished). Through various types of communication networks, the heterogeneous and complex data sets are transmitted and stored in traditional relational databases, data warehouses, web servers, application servers and file servers. For an efficient use of this data, models should be developed accordingly. Hadoop[1] or Spark[2] is a promising platform since it overcomes many of the constraints with traditional technologies, such as limitations of storage and capacities of computation of huge volume of data. The processing power needed for these platforms is very high, and the cloud

[1]Website: http://hadoop.apache.org/.
[2]Website: https://spark.apache.org/.

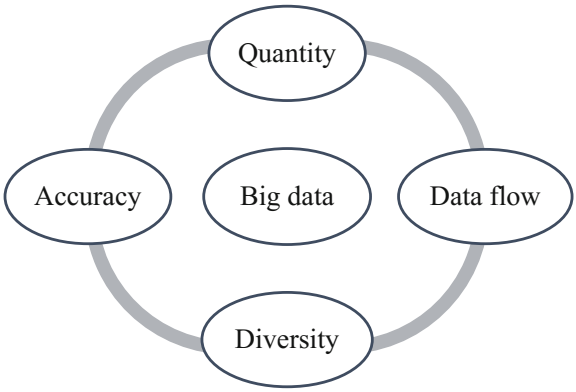


Fig. 2 Components of big data

computing and distributed technologies are useful tools to be able to analyze data at real-time domain also.

Figure 2 shows the constituents of "big data". The figure can be explained taking an example of the power transformer.

- The *quantity* of data collected and to be collected is going to increase in the future. For instance, an utility collecting data on physical condition of transformers, the quantity would be obviously huge. New data accumulating with the old historical data, and increasing the size of database. There are issues in data capture, transmission, processing, storage, searching, sharing, analysis and visualization. Extracting the relevant data out of gigabytes of data and then transform that relevant data into useful, actionable information is important. Data is increasing at exponential rates, but the improvement of information processing methods is relatively slow which takes us to the next point.
- The *data flow* needs not be constant for all transformers in different locations. As the data flow varies, modelling aspects for transformers will also vary. The condition monitoring is helpful when one can devise maintenance modelling according to the data collected in the database.
- *Accuracy* of data collected is still an open question. Uncertainty due to data inconsistency and incompleteness, latency and model approximations account for accuracy of collected data.
- At the end, *diversity* in the collected data where we recall Fig. 1, where it is explained about extracting classification models from structured large databases. In the case of transformers, the health status in terms of gas measurement, visual data like rust and other physical aspects contribute to data diversity. While the big data discussion is often centered around quantity, it can be suspected the real problem with data will be diversity. Multiple variables, often from disparate sources such as different sensors on a transformer, can now be used as inputs for inclusion in models which detect the existence of incipient failures. In some cases, the data does not necessarily need to be a numerical value as in rust on the

asset's surface. Technological advancements in the future will allow models to read non-numerical values as inputs to predictive maintenance models.

A number of questions arises when data collection for condition monitoring is linked with big data. Taking the above example of the transformer, questions on data collection will be like: which part of the transformer is to be monitored, which type of sensor is to be placed and what kind of data are expected from the sensor, how frequently the data is to be collected, whether collected data reached the database on time or not, how to deal with bad data (or missing information). From the huge chunk of data, it will be worth to explore pathways using data analytics to read monitoring data and also learn data correlation (e.g. demand vs weather factors) for efficient maintenance plans. After data collection, new questions on data analytics be like: which technique accurately models the failure rate or ageing, how to deal with missing data and how will it affect the modelling technique, how to deal with situations where there are multiple anomalies detected at the same time. Utilities will need to act soon starting with a small step. Figure 3 shows the pathway in adopting data analytics with four basic steps, namely, building base, learn and adapt, test (small scale), and improvise. Specifying a fixed time-scale for each step is not permissible since different individual steps vary within different utilities. Building a base to collect data while identifying the right tools or database management platforms is vital. Key performance indicators (KPIs) play an important role in data analytics as they determine the significance of each kind of asset data collected. Foundation is not strong without participation of analysts and proper training of existing workforce about data analytics. Once the base is strong, next step is learning and adapting to analytics while enhancing participation in other units within the utility and embracing data security at the same time. Before rolling out the practice in a full-fledged scale, kicking off at small scale is beneficial to

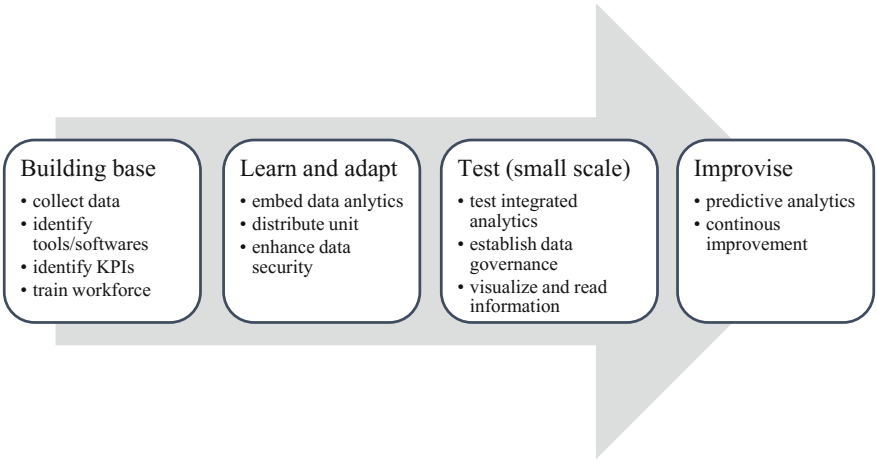


Fig. 3 Future directions on implementing data analytics (KPI refers to Key Performance Indicator)

utility. Establishing data governance at the application stage is critical. For any utility, data governance can be defined as a collection of practices related to data that ensures security, quality, usability and availability [3]. Following the small-scale application, adapting to predictive analytics becomes easier with room for continuous improvement

4 Conclusion

In conclusion, big data will have a large impact on the management of utilities in case of fast deployment of ICT and intelligent sensing within the transmission network. There are many challenges which would affect the success of big data applications in future for utilities. Combination of a significant increase in data, reduced cost of storage, the advancement of cloud based data analytic technology, combined with the emergence of data analyst and scientist roles who know data is the new oil. Presently, experience in integrating big data with current framework is limited. In particular, analytics must be supported by true optimization models to automate the maintenance planning and outage scheduling. It is intended to discover correlations or patterns to make holistic decisions and with the help of analytics utilities can consider all aspects of a decision—the financial side, the maintenance side, as well as the operations side. Also, real application of big data is the ability to understand what data sample is required, ways to analyze and interpret, and then use it. Without completed fields, or validated data, analysis is not possible. So good amount of effort is needed in future to be spent to develop more advanced and efficient algorithms for data analysis that can be easily accepted by utilities. In the end, effective maintenance will be a result of quality, timeliness, accuracy and completeness of information related to machine degradation state, based on which decisions are made.

Acknowledgements The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 608540 GARPUR project http://www.garpur-project.eu. The scientific responsibility rests with the authors.

The authors would like to thank Rémy Clement and Pascal Tournebise of RTE France, Maria Daniela Catrinu-Renstrøm of Statnett SF Norway, and the anonymous reviewers for constructive and insightful feedback.

References

1. CIGRE (2011) WG B3.12. Obtaining value from on-line substation condition monitoring., s. l.: CIGRE
2. GARPUR (2014) GARPUR: generally accepted reliability principle with uncertainty modelling and through probabilistic risk assessment. Available at: http://www.garpur-project.eu/

3. Khatri V, Brown CV (2010) Designing data governance. Commun ACM 53(1):148–152
4. Khuntia SR, Rueda JL, Bouwman S, van der Meijden M (2016) A literature survey on asset management in electrical power [transmission and distribution] system. Int Trans Electr Energy Syst 26(10):2123–2133
5. Khuntia SR, Tuinema B, Rueda JL, van der Meijden MA (2016) Time-horizons in the planning and operation of transmission networks: an overview. IET Gener Transm Distrib 10 (4):841–848
6. Koronis A (2006) Foreword of challenges of managing information quality in service organisations. Idea Group Inc., USA
7. Peng X et al (2013) Application of K-Means method to pattern recognition in on-line cable partial discharge monitoring. IEEE Trans Dielectr Electr Insul 20(3):754–761
8. Tamilselvan P, Wang P (2013) Failure diagnosis using deep belief learning based health state classification. Reliab Eng Syst Saf 115:124–135
9. Zetter K (2016) The Ukrainian Power Grid Was Hacked Again. Available at: https://motherboard.vice.com/en_us/article/bmvkn4/ukrainian-power-station-hacking-december-2016-report. Accessed 2 Feb 2018
10. Zhou C, Michel M, Hepburn D, Song X (2009) On-line partial discharge monitoring in medium voltage underground cables. IET Sci Meas Technol 3(5):354–363

Decision-Making in Asset Management Under Regulatory Constraints

Dragan Komljenovic, Georges Abdul-Nour and Jean-François Boudreau

Abstract Asset management (AsM) policies are influenced by various factors (strategic, technical/technological, economic, organizational, regulatory/legal, safety, markets, competition, etc.). Therefore, the AsM decision-making process should take into account relevant factors for balancing risks, performance, costs, and benefits. Modern organizations use various quantitative models to assess related risks as well as performance, costs and benefits. The results of these models are important input for the overall AsM decision-making process. Decision-makers occasionally give an overwhelming importance to numerical results. Ignoring the limitations of the quantitative risk models used, this practice may be misleading and result in potentially inappropriate decisions. A holistic Risk-Informed Decision-Making (RIDM) process developed for AsM is described in this paper. This process is adapted in the presented case study to analyze possible modification strategies for a nuclear power plant's emergency core cooling system (ECCS). During an operators' training preparation, a design weakness has been discovered in the ECCS, which called for major modifications to the system. The Probabilistic Risk Assessment (PRA) technique has been used to calculate the risk levels of various analyzed options. Nevertheless, the results of the PRA alone were insufficient to demonstrate regulatory compliance. The RIDM process has been used to provide the fundamental insights necessary to obtain regulatory approval for the proposed modification strategy.

D. Komljenovic (✉) · J.-F. Boudreau
Hydro-Quebec's Research Institute (IREQ), Varennes, Canada
e-mail: komljenovic.dragan@ireq.ca

J.-F. Boudreau
e-mail: boudreau.jean-francois@ireq.ca

G. Abdul-Nour
Université du Québec à Trois-Rivières (UQTR), Trois-Rivières, Canada
e-mail: georges.abdulnour@uqtr.ca

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_32

1 Introduction

Asset Management (AsM) has become widespread among modern companies as an effective approach allowing to deliver value from assets and to ensure the sustainability of the business and its operations [21, 26]. The advance of AsM experience and knowledge across various industries has been enacted into a new International Standard on AsM, the standard ISO 55000 [19].

Contemporary enterprises operate in a market, natural, technical, technological, organizational, regulatory, legal, political and financial environment (hereafter called business and operational environment), which is complex and characterized by significant intrinsic uncertainties.

In this context, the decision-making process related to AsM may reveal very challenging due to sometimes significant uncertainties related to the nature of often conflicting influence factors [10, 26, 30].

Modern organizations try to address these issues by using various models and tools that help decrease uncertainties and quantify risks in their asset management decision-making process. The results provided through these models and tools are an important input within the AsM decision-making process, but decision-makers occasionally give them an overwhelming importance. Ignoring the limitations of the quantitative risk models used is misleading and result in potentially mistaken decisions.

This paper presents a holistic Risk-Informed Decision-Making (RIDM) approach developed for AsM aiming at addressing the above-discussed challenges. In the case study presented, possible modification strategies for a nuclear power plant's emergency core cooling system (ECCS) are analyzed, and limits of the decision-making based only on the results of quantitative models are shown. This example tackles a particular application case where quantitative models are not sufficient to address the regulatory compliance.

2 Asset Management in Different Industries

The concept of Asset Management has generated significant interest across various industries and is still growing [13, 26].

The nuclear power industry has invested efforts in developing asset management approaches and methods tailored to its needs and particularities. It developed the Nuclear Asset Management (NAM) and the Risk-Informed Asset Management (RIAM) processes [15].

Some other specific AsM processes were also elaborated. The petrochemical industry has developed its own processes at the end of the 1980s [13, 26]. Power generation, transmission and distribution utilities worked on their specific AsM approaches [1, 5, 11, 14]. Actors in the field of infrastructure management have been using their specific AsM for many years [23, 24, 29]. The transportation industry also carried out works in this area [4, 12].

3 Risk-Informed Decision-Making Model in Asset Management

Asset management is composed of a set of interacting and interdependent activities and constituent elements within a multilevel structure. As per best practices, it should be closely linked to the strategic planning of an enterprise [19, 26]. New concepts and approaches in AsM should systematically take into account the overall complexity of the business and operating environment. In such circumstances, contemporary enterprises should be considered as complex adaptive systems (CAS) [25]. Consequently, the overall asset management should also be considered and analyzed as a CAS.

As far as the decision-making in AsM is concerned, it is necessary to develop a holistic model that is able to adequately integrate the overall complexity of the business and operational environment. In this regard, the Risk-Informed Decision-Making (RIDM) is deemed to be the best suited approach [16, 21], and it will be used in this study. Risk-Based Decision-Making as an alternative approach has shown limitations in tackling intangible and unquantifiable factors [2].

The RIDM is a concept elaborated by the U.S. nuclear power industry in the late 1990s. The initial idea was presented in the White Paper of the U.S. Nuclear Regulatory Commission [27]. Its application was further expanded and the framework defined through regulatory documents [28]. There is no unique definition of the RIDM, and several ones may be found across references [6, 22, 27]. The RIDM is a concept which involves considering, appropriately weighting, and integrating a range of often complex inputs and insights into decision-making [6]. It is deliberative and iterative. The RIDM is essentially carried out by decision makers.

The RIDM has been embraced by other industries at risk, such as aerospace [22] and dam safety [17]. This concept is opposed to a risk-based approach in which decision-making is solely based on the numerical results of quantitative risk assessments [2, 21, 27, 28].

The proposed RIDM approach in AsM is primarily suitable for the strategic and asset management decision-making process affecting both mid and long-term performance, as well as the sustainability of an enterprise. It is intended to be generic, applicable and adaptable to any size and any type of organization. Figure 1 depicts the proposed RIDM model in AsM, which is composed of seven sub-models [21]:

1. *Market*: Usually external to an organization/enterprise. It has a major impact on an organization's global performance;
2. *Reliability, availability and maintainability (RAM) factors*: Mainly internal to an enterprise. In principle, it may be efficiently controlled and influenced;
3. *Operations and operational constraints*: Generally internal to an enterprise;
4. *Revenue and costs*: Both internal and external to an enterprise;

Fig. 1 Risk-informed decision-making global model in asset management

5. *Organizational and business factors*: Primarily internal to an enterprise;
6. *Other factors and constraints*: Mainly external to an organization, but may have a major impact on its global performance;
7. *Strategic plan influence*.

The connections between the sub-models and their constituent elements are of different strengths and types such as physical, informational, cyber, geospatial, etc. [20]. Also, these sub-models and their internal constituent elements interact in a complex manner, and lead to the behavior of the whole AsM process that is not obvious from the individual behavior of each sub-model and its elements.

The RIDM model in AsM also enables identifying and mapping the type and strength of the connections between the sub-models and their constituent elements. It also allows identifying uncertainties within the sub-models. Thus, with the main elements defined, an enhanced decision-making framework becomes available, what will be illustrated next using a case study.

4 Case Study

Possible modification strategies for a Canadian CANDU nuclear power plant's emergency core cooling system (ECCS) are analyzed in the case study. The ECCS is one of four special safety systems in CANDU nuclear power plants. Its conception is fairly complex. It functions in three phases after its initiation consists of (a) high pressure, (b) mid-pressure and (c) low pressure phases (including associated components and equipment) with corresponding involvement of control logic, instrumentation, and various support systems (air, electricity, water) (Fig. 2). It is credited as a key mitigating system in numerous nuclear accident scenarios. This system is subject to strict regulatory performance requirements and scrutiny. Its minimum allowable performance standards shall be defined and referenced in the Safety Report.

In this study, a weakness in the design discovered in the ECCS during an operators' training preparation is described, as well as the RIDM process used in determining the best strategy to implement the necessary corrections. Key details are given below.

Throughout the phase of mid-pressure of the ECCS, water is drawn from the dousing tank (red arrow in Fig. 2). The two pneumatic valves (PV) should close at the end of this stage before initiating a long-term low-pressure phase (blue rectangle in dashed line in Fig. 2). Failure to close one of the two valves leads to the aspiration of air into the ECCS pumps (blue oval dashed line in Fig. 2), and consequently their failure. Thus, the function of the ECCS low pressure cannot be

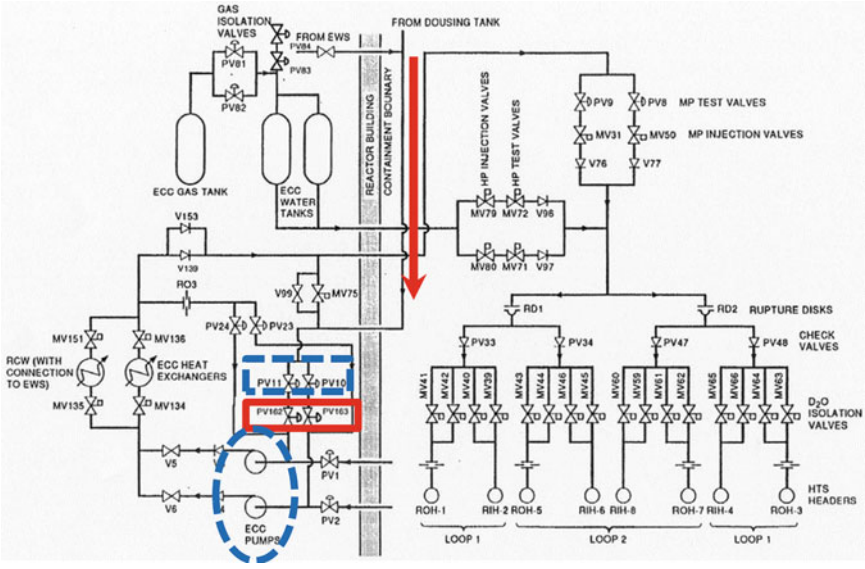


Fig. 2 Simplified ECCS schema with installed PV in series [9]

fulfilled, and the system fails its mission. Such a situation violates the basic safety principles [3, 18]. As such, the situation was rather complex, and cannot be tolerated on a permanent basis. Adequate corrective measures were required, although the failure of the ECCS during the Large Loss of Coolant Accident (LLOCA) as the worst case scenario is an analyzed situation in the Safety Report. The latter demonstrates that the plant is put in a safe state following such an accident even if it represents a significant challenge for operators.

Initially, operators have tested various maneuvering scenarios aiming at avoiding the loss of the system through operational procedures, but the time available to do it revealed insufficient. Afterward, an analysis of possible design changes in the ECCS was performed. Following deliberations led by the plant management and knowledgeable experts, the retained solution consisted in installing two other pneumatic valves in series with the existing ones (red rectangle in solid line, Fig. 2). Given the strict performance requirements for the ECCS, the engineering work, testing and procurement of required equipment represent a lengthy process (roughly one year in total). It should also be highlighted that the studied plant was going to undergo a major refurbishment project in three years, aiming at extending its useful life for another 30 years.

The acceptance of the solution also involved thorough discussions with the Regulatory Body. The next stage consisted in making a decision upon possible scenarios of the installation. Three options were proposed to the Regulatory Body:

- *Option 1*: the installation during a 6 week specific shutdown foreseen uniquely for installing the PV two years ahead of the refurbishment.
- *Option 2*: the PV installation during a planned shutdown one year before the refurbishment.
- *Option 3*: the installation of the PV during the refurbishment.

This situation represented a complex decision-making process in AsM. In the current study, it was conducted through the adaptation of the generic RIDM within the asset management model depicted in Fig. 1. Figure 3 presents the adapted model that restrains relevant influence factors and refers to the sub-models depicted in Fig. 1.

There are five groups of influence factors considered (Fig. 3). The complexity of the decision-making context increases knowing that these factors are not entirely independent. The interdependencies among them are of various types: physical, organizational, procedural/regulatory, spatial, informational and financial.

This paper focuses on the importance and limits of the probabilistic risk assessment (PRA) insights in the decision-making process. Its results provide the assessment of the risk levels required to establish an allowable "time-at-risk" (within the "Nuclear safety requirements" in Fig. 3). It should be stressed that the PRA input be one of the key influence factors in the final decision-making.

The probabilistic risk assessment (PRA) or probabilistic safety assessment (PSA) is a sophisticated risk evaluation technique. In the nuclear power industry, it consists in a comprehensive and integrated assessment of the safety (or risk) of a

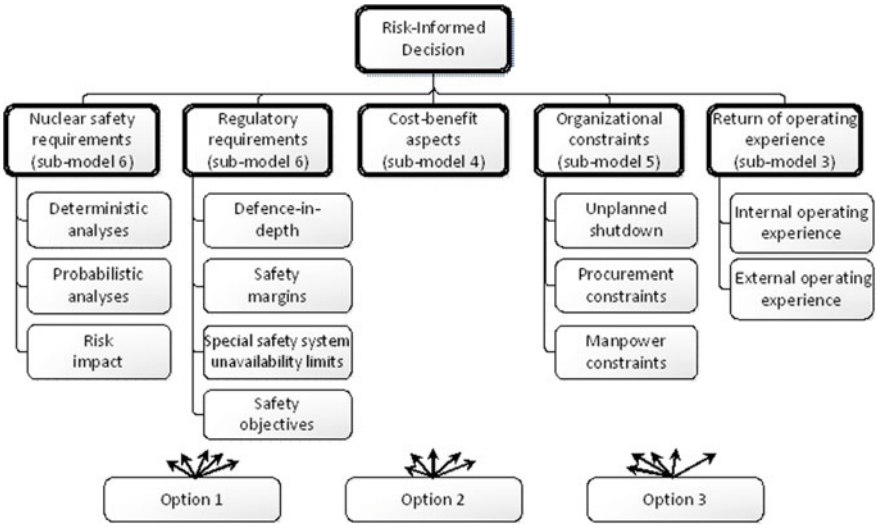


Fig. 3 Risk-informed decision-making in the case study

nuclear reactor facility [8]. The assessment involves the probability, progression and consequences of equipment failures or transient conditions to derive numerical estimates that provide a consistent measure of the risk of the reactor facility, as follows:

- In a level 1 PRA, the sequences of events that may lead to the loss of the reactor's core structural integrity and massive fuel failures are identified and their probabilities are quantified.
- In a level 2 PRA, the level 1 PRA results are used to analyze the containment behavior, evaluate the radionuclides released from the fuel failures, and quantify the releases into the environment.

In this study, the risk quantification is limited to Severe Core Damage Frequency (SCDF) and Large Release Frequency (LRF) assessment as a metric for the risk increase in the full power state. The results obtained are shown in Table 1. With and without PV, the values obtained are inferior to the corresponding quantitative safety limits of 1E-04 for SCDF and 1E-05 for LRF [7, 28]. These safety limits or safety objectives are integrated within "Regulatory requirements" in Fig. 3. Based on the quantitative PRA results only, the design modification may not seem justifiable given that risk levels are below the quantitative safety limits. Thus, the RIDM process (Fig. 3) was used to identify all other relevant fundamental insights necessary to make a final decision. The criteria such as "defense-in-depth" and "absence of a single point of vulnerability" [3] have emphasized the necessity to install the PV in order to meet regulatory requirements, and to comply with the fundamental safety principles.

Table 1 PRA quantification results with and without PV installed

Metric	With PV installed (y/y)	Without PV installed (y/y)	ΔSCDF ΔLRF (increase)
Severe core damage frequency (SCDF)	3.83E−05	4.01E−05	1.80E−06
Large release frequency (LRF)	9.47E−08	1.22E−07	2.73E−08

Hence, the results of the PRA showed the compliance with quantitative safety limits, but were unable to demonstrate regulatory compliance with regard to the "defense-in-depth" and "absence of a single point of vulnerability" requirements. Detailed analyses and deliberations were carried out by the plant management and knowledgeable experts in order to integrate insights of all the influence factors depicted in Fig. 3 in order to identify the most favorable option. The uncertainties were mostly related to the limits of the model, the assumptions made, as well as the quality of the data. Sensitivity analyses were performed to assess the impact of those uncertainties on the results. These analyses showed that the safety objectives are always met. A detailed cost/benefit analysis (CBA) was also completed. For example, the Options 1 and 2 showed a negative NPV (Net Present Value), and the Option 3 had a positive NPV.

Finally, the option 3 was retained as the solution (installation during the refurbishment of the plant). A sensitivity analysis was performed which showed the robustness of the solution. Since the "time-at-risk" was three years for that option, the above process was used to build a final safety and business case, and to obtain regulatory approval for the proposed strategy. The Regulator approved the proposed solution, and the PVs were installed three years later. It should be highlighted that the other CANDU stations have completed this installation through other design modification projects.

5 Conclusions

This paper introduces a holistic Risk-Informed Decision-Making (RIDM) process developed for AsM. It takes into account the complexity of both the operational environment and the internal structure of organizations. In such a decision-making process, insights from quantitative models are rather important. However, the study shows that they do not constitute a sufficient base of information to address complex issues in asset management. Those tools are usually unable to capture intangible influence factors (e.g. non-quantitative regulatory requirements) which may become dominant in a final decision-making process.

The case study carried out at a CANDU nuclear power plant illustrates that the insights from the Probabilistic Risk Assessment (PRA) tool were insufficient to demonstrate both the regulatory compliance and the compliance with fundamental

safety principles in the case of a major activity related to the installation of additional equipment within the Emergency Core Cooling System. It confirmed the compliance with the quantitative safety limits, but was not able to show the acquiescence with the requirements of defense-in-depth. Other factors and measures have been used to demonstrate it, although the input from PRA was of chief importance in the decision-making process. This example illustrates the need to cautiously consider quantitative input in a decision-making process in order to avoid wrong decisions.

This illustrative case from the nuclear power industry may serve as an example for other industries where an overwhelming reliance on quantitative models may sometimes be misleading in the decision-making process.

References

1. Adoghe AU, Awosope COA, Ekeh JC (2013) Asset maintenance planning in electric power distribution network using statistical analysis of outage data. Electr Power Energy Syst 47:424–435
2. Apostolakis GE (2004) How useful is quantitative risk assessment? Risk Anal 24(3):515–519
3. Atomic Energy Control Board (AECB) (1991) R-9, requirements for emergency core cooling systems for CANDU nuclear power plants, regulatory policy, Ottawa, Canada. http://nuclearsafety.gc.ca/pubs_catalogue/uploads/R-9E.pdf. Accessed on 21 Jan 2017
4. Ballis A, Dimitriou L (2010) Issues on railway wagon asset management using advanced information systems. Transp Res Part C 18:807–882
5. Bollinger LA, Dijkema GPJ (2016) Evaluating infrastructure resilience to extreme weather—the case of the Dutch electricity transmission network. EJTIR 16(1):214–239
6. Bujor A, Gheorghe R, Lavrisa T, Ishack G (2010) Risk-informed approach for the CNSC power reactor regulatory program, basis document with worksheets and examples of actual and potential applications, Revision 7. Canadian Nuclear Safety Commission CNSC, Ottawa, (internal document)
7. Canadian Nuclear Safety Commission (CNSC) (2009) RD-152 Guidance on the use of deterministic and probabilistic criteria in decision-making for Class I nuclear facilities. Draft Regulatory Document, Ottawa, Canada. http://nuclearsafety.gc.ca/pubs_catalogue/uploads/rd-152d-e.pdf. Accessed on 21 Jan 2017
8. Canadian Nuclear Safety Commission (CNSC) (2014) REGDOC-2.4.2, probabilistic safety assessment (PSA) for nuclear power plants, regulatory document, Ottawa, Canada. http://nuclearsafety.gc.ca/eng/acts-and-regulations/regulatory-documents/published/html/regdoc2-4-2/index.cfm. Accessed on 21 Jan 2017
9. Canteach, Reference Library on CANDU Technology, https://canteach.candu.org/Pages/Welcome.aspx, Toronto, Canada. Accessed on 14 Jan 2017
10. Chopra SS, Khanna V (2015) Interconnectedness and interdependencies of critical infrastructures in the US economy: implications for resilience. Phys A 436:865–877
11. Dashti R, Yousefi S (2013) Reliability based asset management in electrical distribution systems. Reliab Eng Syst Saf 112:12–136
12. Dornan DL (2002) Asset management: remedy for addressing the fiscal challenges facing highway infrastructure. Int J Transp Manag 1:41–54
13. El-Akruti D, Dwight R, Zhang T (2013) The strategic role of engineering asset management. Int J Prod Econ 146:227–239

14. Electrical Power Research Institute (2007a) Program on technology innovation: enterprise asset management—executive primer. EPRI, 1015385Palo Alto, CA
15. Electric Power Research Institute (2007b) Nuclear asset management (NAM) process model. EPRI, 1015091, Palo Alto, CA
16. Electric Power Research Institute (2005) Risk-informed asset management (RIAM), method, process, and business requirements. EPRI, 1009632, Palo Alto, CA
17. Federal Energy Regulatory Commission (FERC) (2015) Risk-informed decision-making (RIDM), dams safety and inspections, initiatives. https://www.ferc.gov/industries/hydropower/safety/guidelines/ridm.asp. Accessed 15 Jan 2017
18. International Atomic Energy Agency (IAEA) (1996) Defence in depth in nuclear safety, INSAG-10. A report by the International Nuclear Safety Group, International Atomic Energy Agency, Vienna
19. ISO 55000 (2014) Asset management—overview, principles and terminology. International Standard
20. Katina PF, Ariel Pinto C, Bradley JM, Hester PT (2014) Interdependency-induced risk with applications to healthcare. Int J Crit Infrastruct Prot 7:12–26
21. Komljenovic D, Gaha M, Abdul-Nour G, Langheit C, Bourgeois M (2016) Risks of extreme and rare events in asset management. Saf Sci 88:129–145
22. NASA (2010) NASA risk-informed decision-making handbook. NASA/SP-2010-576, Version 1.0, Office of Safety and Mission Assurances, NASA Headquarters
23. Park S, Park SI, Lee S-H (2016) Strategy on sustainable infrastructure asset management: focus on Korea's future policy directivity. Renew Sustain Energy Rev 62:710–722
24. Shah R, McMann O, Borthhwick F (2017) Challenges and prospects of applying asset management principles to highway maintenance: a case study of the UK. Transp Res Part A 97:231–243
25. Stacey RD, Mowles C (2016) Strategic management and organisational dynamic; the challenge of complexity to ways of thinking about organisations, 7th edn. Pearson, London
26. The Institute of Asset Management (2015) Asset management—an anatomy V3. https://theiam.org/what-is-asset-management/anatomy-asset-management. Accessed on 15 Jan 2017
27. Travers WD (1999) Staff requirements—SECY-98-14—white paper on risk-informed and performance-based regulation, U.S. nuclear regulatory commission (NRC). http://pbadupws.nrc.gov/docs/ML0037/ML003753601.pdf. Accessed 15 Jan 2017
28. U.S Nuclear Regulatory Commission (NRC) (2011) An approach for using probabilistic risk assessment in risk-informed decisions on plant-specific changes to the licensing basis (R.G. 1.174), Washington, DC 20555-0001
29. Younis R, Knight MA (2014) Development and implementation of an asset management framework for wastewater collection networks. Tunn Undergr Space Technol 39:130–143
30. Zio E (2016) Challenges in the vulnerability and risk analysis of critical infrastructures. Reliab Eng Syst Saf 152:137–150

From Asset Provider to Knowledge Company—Transformation in the Digital Era

Helena Kortelainen, Ari Happonen and Jyri Hanski

Abstract Digitalization and the industrial internet is expected to bring major changes in manufacturing and service delivery—following the transformation in business to consumer markets. The industrial internet enables new business by connecting intelligent devices and people using them into cloud based analytics and decision-making systems. The objective of this research was to understand the transformation from an asset provider to a knowledge service company by analysing companies' service offerings. 'Data to decision'—framework is used to categorize service concepts according to their knowledge intensity. The framework suggests that the service provider could offer new value to its customers by being able to provide knowledge as a service. In this paper, we discuss shortly the steps that characterize the transformation and the possible business impacts this transformation generates. Additionally, we connect the steps into the framework to help further research in this field. For further research, we suggest that the framework could be used for similar analysis with comparable data and to help to understand the ongoing transition in business ecosystems.

1 Introduction

Digitalization and the industrial internet is predicted to transform many sectors, including industry automation, manufacturing and mining [13, 17]. The industrial internet enables new business by connecting intelligent devices and people using them into cloud based analytics and decision-making systems. Along with the digitalization trend, formerly product-centric companies have shifted their focus from product delivery to value adding asset life cycle services [2–4]. Rapid

H. Kortelainen (✉) · J. Hanski
VTT Technical Research Centre of Finland Ltd., Tampere, Finland
e-mail: helena.kortelainen@vtt.fi

A. Happonen
Lappeenranta University of Technology, Lappeenranta, Finland
e-mail: ari.happonen@lut.fi

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_33

development of industrial internet enhances the collection and analysis of information from their products in customer sites in a very profound way (e.g. [15]).

One of the commonly used definitions of digitalization in business context is using "digital technologies to change a business model and provide new revenue and value-producing opportunities" [8]. For example, the industrial internet combines intelligent machines, analytics, and the people who are using them [7], which generates new business opportunities. Product manufacturers—i.e. asset providers—are in a middle of a transformation process that the rapid information technology development drives. Because of speed of the change digitalization prongs into the table, organisations have to be strategically agile [5]. However, digital strategies at forerunner organisations go beyond the technologies. They target to improvements in innovation, decision-making and transforming how the business work [9]. Data, information, knowledge and analytics, and the use of data in decision-making and processes is in the core of this transformation process.

The objective of this research was to understand the transformation process from an asset provider to a knowledge service company by analysing companies' service offerings. This paper is a part of ongoing research, which aims to find new ways to gather, analyze and understand, and then utilize information in fleet level operations in a value network and create innovative technological solutions and systems around information. In this paper, we characterized and discuss shortly the steps that are needed in the transformation and its business impacts. Additionally, we connect the steps into the Data to decision framework to help further research in this field.

2 Data to Decisions Framework

Knowledge-intensive services are designed to provide support for decision-making. Thus, the understanding of decision-making situations and relevant alternatives is the basis for data collection and analysis. Our 'Data to decision'—framework [12] aims to describe the content and the role of data in the context of knowledge intensive services. The framework also illustrates the data refinement and increase in value process in which the data is systematically converted in to information, into knowledge or even to wisdom (see DIKW-hierarchy e.g. in [14]). DIKW hierarchy is widely accepted as a basic model describing the levels of understanding and as a path to make data useful for decision-making. The 'Data to decision'—framework suggest that the service provider could offer new and increasing value to its customers by being able to provide knowledge and consequent actions or action plans as a service instead of mere data collecting and sharing. In this case, service provider is capable in the transformation of the data and the information into such a knowledge that supports the decision making of the customers (asset owners) to improve their business. Such decisions could deal with daily operations, developing the assets and operations, or with long-term strategic investments, and a wide variety of other technical issues and business situations. Figure 1 illustrates the phases of the 'Data to decision' process.

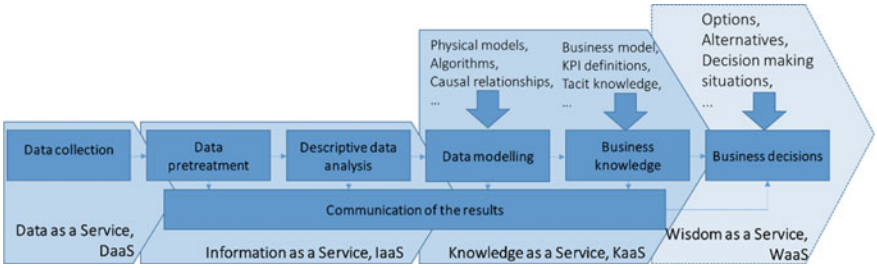


Fig. 1 Schematic presentation of the data to decision framework

The data, information, knowledge and wisdom (DIKW) hierarchy [1] or the 'knowledge pyramid' [14] offers a theoretical background for our work. Table 1 summarizes the descriptions for DIKW-hierarchy that we apply in the context asset-related decision-making and knowledge-intensive services.

Traditionally asset owners collect data, create information and knowledge, and gain wisdom that is applicable and useful for their business. Manufacturers and service providers can support asset owners by different services based on their experience and knowledge about installed base. However, the Data to decision—framework (Fig. 1) suggests that the service and/or product provider could offer

Table 1 DIKW hierarchy descriptions and examples

DIKW hierarchy	Description	Example
Data	Saved facts, i.e. values and observations about selected variables	Measured value by a technical sensor
Information	Data transferred to a form that is meaningful and useful to a human	Trends of failure rate or vibration amplitude. More generally, analyzed and processed data that has a easier to interpreted meaning over time
Knowledge	Ability to interpret trends or other signs and to recognize when actions are needed	A professional skill developed over a long time to understand the information and take actions with good results (e.g. ability to understand trends and a need for actions)
Wisdom	Ability to recognize relevant options in the current situation and to compare those as pros and cons with a skill set to utilize tools to make optimal decisions	Ability to recognize that the available vibration measurement based options are to change the component right away or wait until the next planned stoppage and to evaluate risks and benefits related to these options and also skills to use e.g. simulation tools to make educated hypothesis what are the risks versus rewards in different scenarios

new value to its customers by being able also to provide knowledge as a service. This would require that the service providers develop new competencies, analytical capacities and deep understanding of the customer's value creation. On the other hand, service providers might have more knowledge (big data) to build these competencies than asset owners do as they are in position to collect and analyse data from their installed base and from a plenitude of use environments.

3 Application of Data to Decision—Framework to Service Model Evaluation

3.1 Methods and Materials

The empirical study was carried out in an on-going large-scale Finnish research project S4Fleet—Service Solutions for Fleet management[1] that is characterized by strong industrial involvement. The S4Fleet case companies represented machinery manufactures, technology and knowledge and service suppliers, and IT service and infrastructure providers. In a structured workshop, the companies were asked to present their development goals and plans related to data-based industrial services. Eleven companies participated the workshop (see Table 2). During the workshop, the researchers recorded the presentations and analysed afterwards the presented descriptions and prospects of the asset management related services according to the stages in the Data to decision framework (Fig. 1) and DIKW model (Table 1).

3.2 Analysis of Companies Capabilities with the Framework

The planned future asset service concepts presented by the companies were analysed according to the DIKW-hierarchy and Data to decision - framework (DaaS, IaaS, KaaS, and WaaS). Table 2 summarises both the service descriptions (more details in [10]) and the evaluation of the service knowledge intensity. The researcher team evaluated the service model of the described offering.

3.3 Findings Related to the Transformation and Value Creation

The company offering in Table 2 consists from two broad categories; service concepts that enable fleet services (ICT) and the fleet service concepts that are based

[1]See details in http://www.dimecc.com/dimecc-services/s4fleet/.

Table 2 Analysis of the fleet service concepts identified by S4Fleet case companies using DIKW hierarchy and data to decision framework

Company identifier[a]	Company size	Planned service concept (short description)	Service model
ICT1	Small	Platform and technical means for IoT solutions	DaaS
ICT2	Large	Analytics and platform for predictive maintenance and asset optimization solutions for fleet management	DaaS
ICT3	Medium	Offering collaboration platforms which combine data from different sources	DaaS
ICT4	Small	Helps customers to focus development efforts to improve OEE. Uses software and working practice to manage reliability and RAMS requirements	IaaS
ICT5	Medium	Platform/infrastructure service for extending the existing service from measuring the asset level to the fleet-level	DaaS
SER1	Large	Service manager view to integrated fleet level and customer profiling information	DaaS
SER2	Medium	Logistics services and material stream analyses for improved logistic processes.	IaaS/ KaaS
OEM1	Large	Services for predicting equipment failures, predictions for unit's performance, maintenance and defining best practices	IaaS/ KaaS/ WaaS
OEM2	Large	Service for optimizing operations and maintenance activities, and estimating total life cycle costs	KaaS/ WaaS
OEM3	Large	Services for predictable and guaranteed failure free run to the customer based on machinery usage and data from equipment	IaaS/ KaaS
OEM4	Large	Service for analysis of maintenance, failures and lifetimes	IaaS/ KaaS

[a]ICT refers to companies offering ICT systems and infrastructure, SER refers to the companies offering solely technical services and OEM to the companies offering products and services

on systematic collection of asset data (OEM, SER). ICT companies seem to focus in developing platforms and analytical tools but do not indicate plans to incorporate substance knowledge in their offering. Possibly the ICT companies are still not getting indications from their customers, that the customers would be ready to pay enough for the service for it to cover the costs of the substance knowledge integration in reasonable time frame. OEM companies have more advanced plans that is reasonable when thinking their close relationships with end-users. All OEMs are large companies with a global installed base and they look for upgrading the service model by advanced service concepts (KaaS, WaaS).

From the perspective of the Data to decision—framework, the short descriptions for future service concepts offer most information on the communication of results. This is understandable as identification of the benefits for different user groups are crucial at an early stage of service concept level. In some cases (OEM1, OEM2) described service concepts enable a wide variety of service models depending on

the availability of data, the nature of customer relationship and business model choices of the asset owner and service provider. The appropriate service model (IaaS, KaaS, WaaS) may need to be selected case-by-case depending on the customer needs and limitations on the available data. The choice for the service model depends of course also from the competitive environment, as the high-end knowledge intensive business may be lucrative also to other service providers. In those cases, the company needs to develop competencies and acquire tools for the most demanding service offering possible, to be able to keep and grow up the market share.

4 Changing Roles in the Industrial Ecosystems

In a constantly evolving market, the manufacturing companies have to evaluate and change their role in the ecosystem in order to stay competitive. Rapid development of digitalisation drives the transformation process but the transformation also applies to the service content. According to Davies [6] the competitive advantage is not simply about providing services, but how services are combined with products to provide high value "integrated solutions" that address a customer's business or operational needs. As indicated in Table 2, manufacturing companies are actively developing service offering that would improve their customer's performance. What needs to addressed too is, that 3rd party players have also noticed the possibilities these data rich markets could offer. For example, Travis Kalanick (CEO of Uber) said that "If we can get you a car in five minutes, we can get you anything in five minutes" [5], meaning that they are really prepared to fight this digitalization "game", creating new and interesting information and data rich platforms. This statement could be interpreted as a remainder to all other business sectors, that digitalisation may disrupt "traditional" field of industry in a crucial and unforeseen way.

'Data to decision'—framework and DIKW hierarchy suggest that the service provider could offer new value to its customers by being able to provide knowledge (or even wisdom) as a service. The acknowledged challenges include e.g. fragmented fleet data and the challenge of data sharing between various players in industrial networks. Also it should be noted that the data collection environments are not equal, meaning that the data-analysis will have it's challenges too. The service concepts presented by the machinery suppliers (OEMs) indicate that data collection or sharing information is no longer sufficient but the development goals encompass significant increase of knowledge intensity. Even technical elements for fleet services are available the methods to extract useful information from the data, and models the value generation from the asset and fleet management perspectives require validated approaches [11]. The extending service offering covers also issues that the asset owners have traditionally had a strong knowhow. This means that the service provider has to develop the excellence in refining data in a way that delivers more value to the asset owner.

An interesting finding in the role of the ICT companies is from the ecosystems. ICT companies in our study (see Table 2) seem to be satisfied with role as an ICT infrastructure and software tools provider. This role would be also supplementary to the expanding role of the OEM and offer opportunities for growth along with the rapidly growing market of IoT-solutions. This finding is well in line with recently published collaboration schemes in global companies (e.g. KONE and IBM; ABB and Microsoft). The strategy in the digital transformation may be that not all companies aspire the same role as a "knowledge company" covering the whole range in the Data to decision process. The companies may update and upgrade their role in the ecosystem as an enabler, information deliverer, asset provider or asset service provider with a limited or extended offering.

Digital strategies at forerunner target to improvements in innovation, decision-making and transforming how the business work. Ecosystems are highly valuated in the B2C products [16]. Refining the data for business knowledge in ecosystems, offering data to, using data from 3rd parties and taking the full advantage of the fleet data and the application of open interfaces and APIs might be elements for the "intelligence layer" also in the B2B ecosystems.

5 Conclusions and Future Research

In this paper, we discussed the steps that characterize the transformation of the machinery suppliers to knowledge service providers, changing roles in the ecosystems and the possible business impacts this transformation generates. As an actual impact, the utilization of fleet life-cycle data can work as enabler for numerous possibilities to upgrade data into knowledge with business value. The DIKW hierarchy and Data to decision—framework emphasise both the development of data refining capabilities when planning the creation of new fleet asset services. As a conclusion, we summarize our findings of the transformation, the changing roles of actors in the manufacturing ecosystem and the changes in the knowledge intensity of the asset services provided by the companies in a following way:

- ICT (& Knowledge) companies offer data pipelines, platforms and ICT infrastructure for OEMs and service providers
- ICT (& Knowledge) companies focus in developing platforms and analytical tools.
- ICT (& Knowledge) companies are key players in the future ecosystems as platform operators that are essential for all knowledge based services of OEMs and service providers
- OEM companies develop their information or knowledge based offering mainly to support the customer's daily operations.
- OEMs and service providers have to develop the excellence in refining data in a way that delivers more value to the asset owner.

– OEM companies aim at becoming "knowledge companies" that do not own production assets but have the control over the platform used in their business ecosystem.

For further research, we suggest that the framework should be tested for similar analysis with comparable data and to help to understand the ongoing transition in business ecosystems. The digital transformation of the business models in B2C sector is already widely researched and there are multiple examples of large-scale business transformations (e.g. Uber, Airbnb and Etsy). It is more than likely that same phenomena will occur in B2B sector, too. Those OEM and service companies that have developed capabilities to handle and analyse the data and connect the demand and supply of the knowledge, can win in this game.

Acknowledgements The research leading to these results has received funding from the Finnish Funding Agency for Innovation (Tekes). The authors gratefully acknowledge Tekes for the financial support that made this study possible.

References

1. Ackoff R (1989) From data to wisdom. J Appl Syst Anal 16(1):3–9
2. Ahonen T, Reunanen M, Pajari O, Ojanen V, Lanne M (2010) Maintenance communities—a new model for the networked delivery of maintenance services. Int J Bus Innov Res 4(6):560–583
3. Baines T, Lightfoot HW (2013) Servitization of the manufacturing firm. Int J Oper Prod Manag 34(1):2–35
4. Brady T, Davies A, Gann D (2005) Creating value by delivering integrated solutions. Int J Proj Manag 23:360–365
5. Choudary S, van Alstyne M, Parker G (2016) Platform revolution: how networked markets are transforming the economy and how to make them work for you, 1st edn. W. W. Norton & Company, 256 p
6. Davis A (2004) Moving base into high-value integrated solutions: a value stream approach. Ind Corp Change 13(5):727–756
7. Evans PC, Annunziata M (2012) Industrial internet: pushing the boundaries of minds and machines. GE, 26 Nov 2012. http://www.ge.com/docs/chapters/Industrial_Internet.pdf. Accessed 14 Jan 2016
8. Gartner (2016) IT Glossary. http://www.gartner.com/it-glossary/digitalization/. Accessed 21 Apr 2016
9. Kane G, Palmer D, Phillips A, Kiron D, Buckley N (2015) Strategy, not technology, drives digital transformation. MIT Sloan Management Review and Deloitte University Press, July 2015
10. Kortelainen H, Hanski J, Valkokari P, Ahonen T (2016) Tapping the value potential of extended asset services—experiences from Finnish companies. Maintenance Performance Measurement and Management Conference, MPMM 2016, 28 Nov 2016, Luleå, Sweden. Proceedings, Luleå University of Technology
11. Kortelainen H, Happonen A, Kinnunen S-K (2016) Fleet service generation—challenges in corporate asset management. In: Proceedings of the 10th world congress on engineering asset management (WCEAM 2015) (Lecture notes in mechanical engineering). Springer, Berlin, pp 373–380

12. Kunttu S, Ahonen T, Kortelainen H, Jantunen E (2016) Data to decision—knowledge-intensive services for asset owners. Euro maintenance 2016, 30 May–1 June 2016, Athens, Greece. Proceedings. EFMNS, European Federation of National Maintenance Societies, pp 75–83
13. OECD (2015) OECD digital economy outlook 2015. OECD Publishing, Paris
14. Rowley J (2006) The wisdom hierarchy: representation of the DIKW hierarchy. J Inf Sci 33 (2):163–180
15. Thomas RJ, Kass A, Davarzani L (2014) From looking digital to being digital: the impact of technology on the future of work, report. https://www.accenture.com/. Accessed 14 Jan 2016
16. Van Alstyne M, Parker G, Choudary S (2016) Pipelines, platforms, and the new rules of strategy. Harv Bus Rev
17. WEF (World Economic Forum) (2014) Industrial internet of things: unleashing the potential of connected products and services, report. http://reports.weforum.org/industrial-internet-of-things. Accessed 23 Feb 2016

Method to Determine Internal Leakage of Aircraft's Hydraulic Servo

Jouko Laitinen and Kari Koskinen

Abstract This paper addresses a method to determine possible internal hydraulic leak of airplane control servo. Internal hydraulic leakage is difficult to perceive due the leak doesn't have any visible evidence outside the control servo. Leaking happens past the hydraulic piston or through faulty control valves. Such leaking can have serious consequences due servo's decreased capacity and force especially situations where all the aerodynamic force of control surface is needed to maintain the ability to control the airplane. This kind of hidden failure can advance unobserved because decreased capacity of the servo can be enough to keep the airplane in control during normal flight conditions. However, in heavy flight manoeuvres such as aerobatics, the decreased rate of control surface's angular velocity can lead to in-flight loss of control. One relevant problem is how to find out when inner leakage is so severe, that it endangers safety of the airplane. To solve this kind of problem, a detailed quantitative model has been built. This model utilizes recorded in-flight process data and aircraft's aerodynamical characteristics to determine condition of the control servo.

1 Introduction

The aircraft must be controllable and manoeuvrable under all anticipated operating conditions and it shall be possible to make smooth transitions from one flight maneuver to another (e.g. turns, sideslips, changes of engine power, changes of aeroplane configurations) even in the event of failure of engine.

The aim of this study is find out if it is possible to determine oncoming failure of aircraft's stabilator, which can lead to loss of control.

J. Laitinen (✉) · K. Koskinen
Tampere University of Technology, Tampere, Finland
e-mail: jouko.laitinen@tut.fi

K. Koskinen
e-mail: kari.t.koskinen@tut.fi

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_34

Fighter aircraft Hornet F/A-18 is equipped with a solid-state recorder (USSR = universal solid-state recorder) which makes possible to record and analyse aircraft's bus messages i.e. process data, after the flight mission. F-18's command signal of the stabilizing surface, stabilator (a combined stabilizer and elevator at the tail of an aircraft), called 'Pitch command' and the position signal of both stabilators called 'Left stab position' and 'Right stab position' are registered on the frequency of 20 Hz i.e. 20 times in a second. The position information of the stabilators' and their actuating devices, control servos, was studied by calculating how much the positions of the stabilators deviate from the value of the command signal. When the command-position-difference of the stabilators (the difference between the command signal and the stabilators' acquired position) was calculated, it was noticed that there are big differences of the values at the various phases of the flight, Fig. 1.

In the ideal case the value of the command-position-difference is zero, i.e. the position of stabilator corresponds to the command that is given to it. In reality, the command-position-difference deviates from zero more or less and is either negative or positive. The differing of the value from zero means error, in other words the fact that the stabilator is not able to carry out a control command totally.

At first it was assumed that the big errors meant the fact that the servo of the stabilator is failing. When the data collected from several flights was compared, it was noticed that on flights where many big and quick control deflections were used (for example the aerobatics), the command-position-difference also was big, Fig. 1.

Because the condition of the stabilator's servo cannot be directly concluded from the command-position-difference, the issue was studied by comparing the command-position-difference with the amount of the force required for the rotating of the stabilator surface during flight. Force required for rotating the stabilator, or the hinge moment (HM), to be precise, can be calculated from the dimensions of the

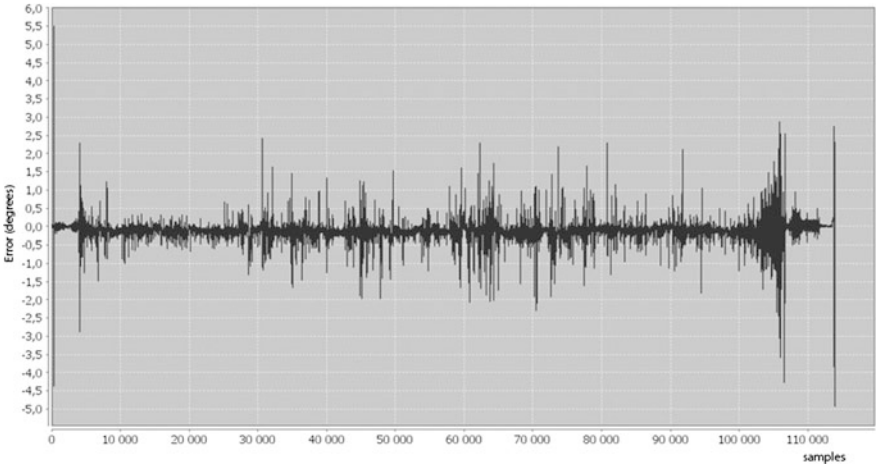


Fig. 1 Typical stabilator command-position-difference (error) in flight with heavy manouvers

stabilator surface and from the aircraft's speed, altitude and from the hinge moment coefficient of the control surface. It is not possible to calculate the hinge moment coefficient directly but it had to be defined on the basis of the graphs based on the wind-tunnel tests. The hinge moment coefficient is calculated by using values of the angle of attack of the airplane and the stabilator and the angle of sideslip which was supposed to be zero in this study.

In this study, the idea is: the bigger force is needed for rotating the stabilator, the bigger error can be accepted. On the other hand, if the need of the force is small but the error is big, it may mean a failure of the stabilator servo or of the hydraulic system.

The case of this study was a flight in which the warning system of the aircraft gave warning of the failure of the right stabilator servo. This made sure that it was a question of a functional failure of the stabilator servo.

The study case was a test flight and during the flight the right engine of the aircraft was turned off and was started again. Functioning of the control servos was examined in two sections of a flight: the first section was before the failure, in which both engines were running and second just before the failure in which the right engine had been turned off. The reason why the second examination section was just before the failure but not during it, was due to the fact that one wanted to see if an oncoming failure can be seen in advance.

2 Calculating the Hinge Moment

For calculating the stabilator's hinge moment, values presented in Figs. 1, 2 and 3 are needed. The hinge moment is calculated on Eqs. (1) and (2). It is possible to simplify the calculation by replacing the physical dimensions of the stabilator with factor 1 in the equation, because the relative value of the hinge moment is 1 (the size of the control surface will not change when the values of the hinge moment are calculated in different points of the flight) [1].

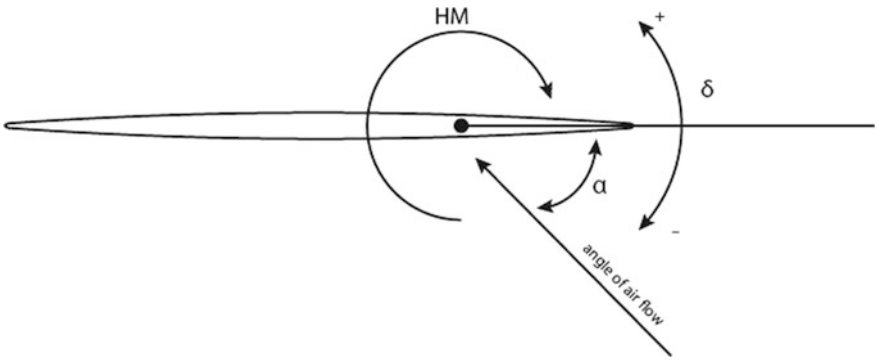


Fig. 2 Angles of the stabilator α = angle of attack with respect to air flow, δ = deflection angle of the stabilator. HM = direction of the hinge moment in regard to hinge axis of the stabilator

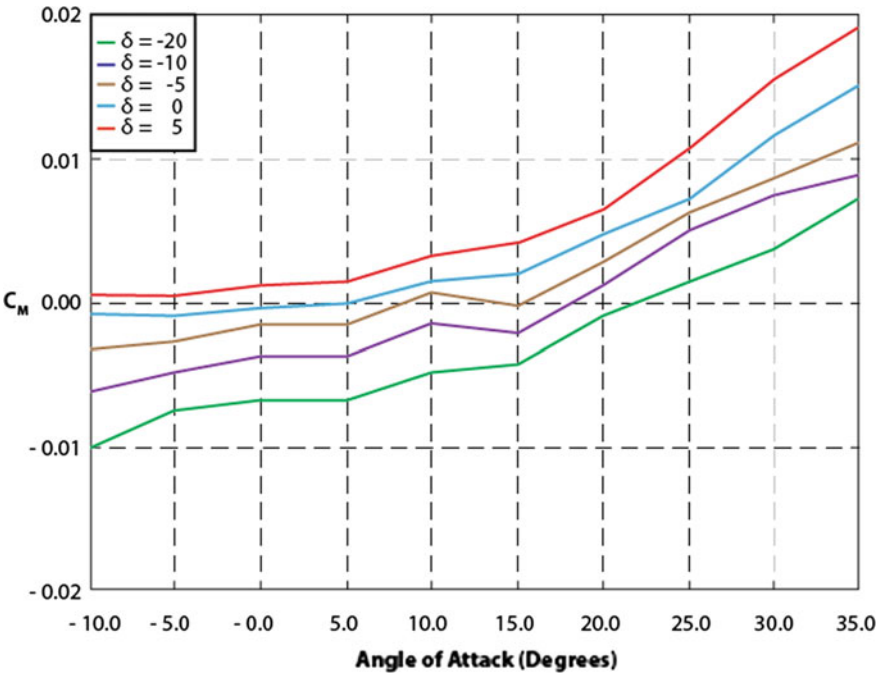


Fig. 3 Values of the hinge moment coefficient in graph that has been drawn up on the basis of the wind tunnel tests. On the vertical axis is the value of the hinge moment coefficient, on the horizontal axis is the angle of attack of the aircraft. The different curves describe the change in the hinge moment coefficient at different angle of attack of the control surface. Angle of Sideslip (β) = 0

Stabilator's hinge moment is calculated with the equation

$$HM = 0.5\rho V^2 C_M S_e c_e \tag{1}$$

Coefficients of the term $0.5\rho V^2$ in Eq. (1). ρ = density of air and V = velocity of the aircraft are obtained from the records of the solid state recorder. Coefficients S_e = area of the stabilator surface and c_e = the chord of the stabilator are set to 1, because all stabilators in the same side of F-18's are similar.

Hinge moment coefficient is calculated with the equation

$$C_M = C_{h_0} + C_{h_\alpha}\alpha_h + C_{h_{\delta_e}}\delta_e \tag{2}$$

where α_h = stabilator's angle of attack and δ_e = stabilator's deflection angle. The hinge moment coefficient which appears in Eq. (1) is calculated from Eq. (2). The coefficients $C_{h_0}, C_{h_\alpha}, C_{h_{\delta_e}}$, which describe the properties of the stabilator under different flight conditions are not within reach of but the value of the hinge moment coefficient was determined on the basis of wind tunnel tests [2, 3].

3 Comparing of the Command-Position-Difference with the Hinge Moment of the Stabilator

The amount of the command-position-difference, i.e. the difference between wanted and acquired stabilator position, was compared with the amount of the hinge moment needed in the different flight states. The comparison was made by dividing the error with the value of the hinge moment calculated and by normalizing the result. The results of the calculation were divided into 50 classes so that it was possible to draw a frequency distribution from the results, Figs. 4, 5, 6 and 7.

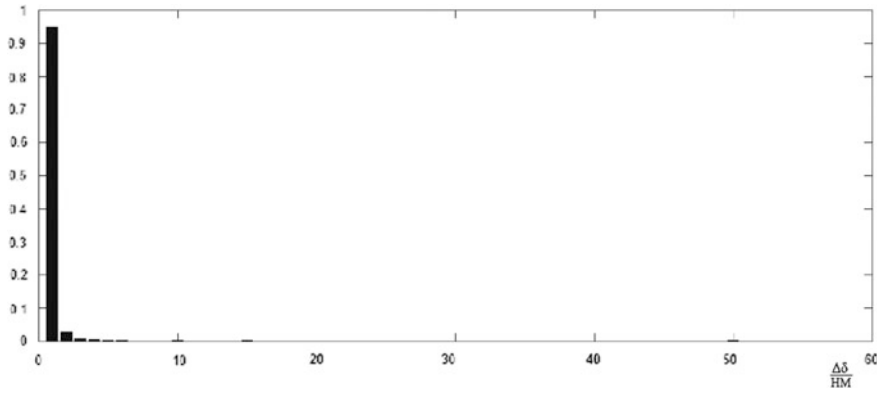


Fig. 4 Distribution of the proportion of command-position-difference to stabilator's hinge moment (Δδ/HM) of the **left** stabilator before failure. Most of the values, about 95%, is near zero

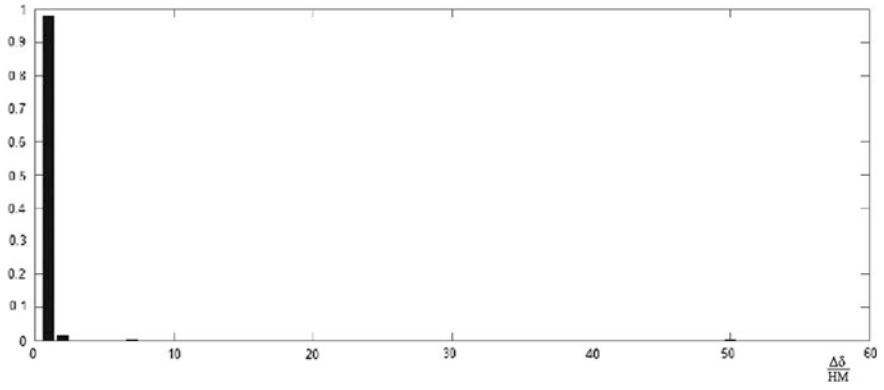


Fig. 5 Distribution of the proportion of command-position-difference to stabilator's hinge moment (Δδ/HM) of the **right** stabilator before failure. Almost all the values are near zero

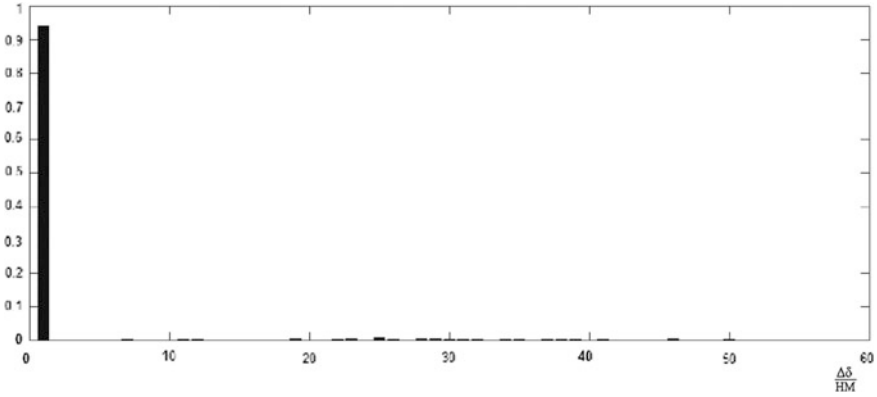


Fig. 6 Distribution of the proportion of command-position-difference to stabilator's hinge moment (Δδ/HM) of the **left** stabilator just before failure. Most of the values, about 94%, is near zero even though some of the values are deviating from zero

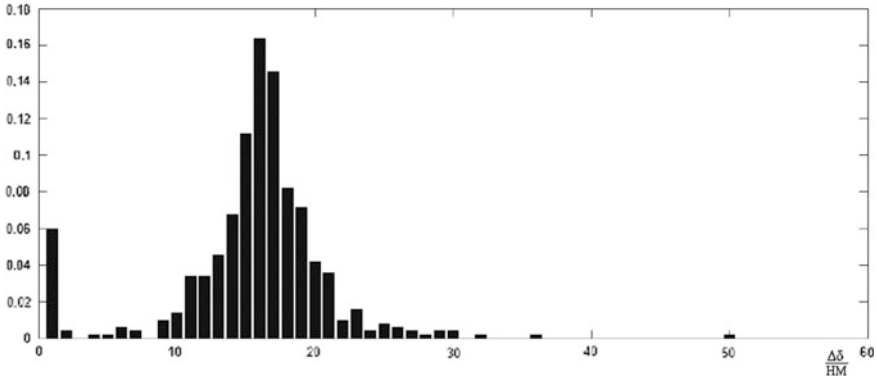


Fig. 7 Distribution of the proportion of command-position-difference to stabilator's hinge moment (Δδ/HM) of the **right** stabilator just before failure. The distribution is located on the right which means that the capacity of the servo is inadequate

Figures 4 and 6 describes the distribution of the proportion of command-position-difference to stabilator's hinge moment of the left stabilator and in Figs. 5 and 7 right stabilator. When the proportion is small (the distribution is centred on near zero), it means that even the required hinge moment is big, the error is yet small. In an opposite case (the proportion is big) the error is big even though the required hinge moment is small. In the graphs, the proportion of the error and hinge moment $\Delta\delta/HM$ increases with the horizontal axis to the right and on the vertical axis is the relative portion of each class.

According to Figs. 5 and 6, the operation of the left and right control servo was normal when both aircraft engines were running. Even though the aircraft makes heavy manouvers, the proportion of command-position-difference to stabilator's

hinge moment was near zero. This means that even though control servo didn't reach the desired position, it can be accepted if the flight maneuver is of such nature that the moving of the stabilator requires high force. In this kind of a situation the force of the control servo is not enough even if the servo is in good condition.

From Figs. 6 and 7 one can see that there is a failure in the right control servo. The distribution of the proportion of command-position-difference to stabilator's hinge moment ($\Delta\delta/HM$) of the right stabilator has been centred on distinctly on the right, which means that even though the required hinge moment is small, the position error of the stabilator is significant. In this kind of a situation the capacity of the control servo is not necessarily enough even in a light flight manouvers.

4 Summary

It is possible to conclude from the results of this study, that it is possible detect the failure of the stabilator control servo (for example caused by internal hydraulic leakage) by studying the force requirement of the stabilator by examining the recorded air data. This is made possible by the research results of the aerodynamics of the aircraft and by the data, which are obtained by analyzing the air data and by the command and position signals of control surfaces. The method is possible quite easily to automate. This is important, because there is no another available techniques for in-flight internal servo leak detection in F-18 aircraft.

The findings of this study needs further examination, because they can be affected by deflection of the hydraulic system because one of the engines had been turned off and thus the output of the hydraulic pressure was possibly smaller than both engines running. Was the reason in the hydraulic system or in the stabilator servo, the results show that the ability to function of the stabilator servo had been restricted. If in the cases of failure of the control surface's servo valve the fault of the hydraulic system can be excluded by analyzing data, the failure can be more surely directed to the operation of the servo valve.

References

1. Mulder M (2009) Aerodynamic hinge moment coefficient estimation using automatic fly-by-wire control inputs. In: AIAA modelling and simulation technologies conference, 10–13 Aug 2009, Chicago, Illinois
2. Candida D (2000) An evaluation technique for an F/A-18 aircraft loads model using F/A-18 systems research aircraft flight data. National Aeronautics and Space Administration, Washington DC, 2000
3. MacLaren LD (1993) Low-speed pressure distribution measurements over the aft-fuselage, fins and stabilators of a 1/9 scale F/A-18 wind tunnel model. Research report 9, Aeronautical research laboratory, Melbourne, Victoria, Australia. http://www.dtic.mil/dtic/tr/fulltext/u2/a274870.pdf

Study on the Vibration Reduction Performance of Smart Springs

M. M. Li, D. Ni, W. M. Wu, R. P. Zhu and S. M. Li

Abstract Based on the smart spring's active vibration method, the micro-displacement of piezoelectric ceramic actuators is generated by applying controlled voltage and alternation of the impedance characteristics—no complex displacement amplifying device and high driven voltage are required. A device with active vibration reduction was designed according to the concept of a smart spring. This device is compact in structure, easily assembled by friction blocks and with convenient adjustment of clearance. An excitation test using a single-stage smart spring vibration reduction was carried out, which verified the feasibility and validity of the smart spring's vibration reduction performance. The experiment and the theoretical analysis described in the paper presents the value of the design of the smart spring system.

1 Introduction

Active damping technology relies on external energy to adjust system parameters, which can be adapted to different bandwidths and low-frequency vibrations [6, 13]. Smart spring damping systems used in active damping technologies, apply a control

M. M. Li (✉) · W. M. Wu · R. P. Zhu · S. M. Li
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
e-mail: limiaomiao@nuaa.edu.cn

W. M. Wu
e-mail: 273962355@qq.com

R. P. Zhu
e-mail: rpzhu@nuaa.edu.cn

S. M. Li
e-mail: smli@nuaa.edu.cn

D. Ni
AECC Hunan Aviation Powerplant Research Institute, Zhuzhou 412002, China
e-mail: 532958264@qq.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_35

voltage to the piezoelectric ceramic actuator (PZTA) resulting in micro-displacements; thereby indirectly altering the impedance of the vibration system, which is reliant on damping, stiffness and mass to reduce the transmission of vibrations [4]. Compared to direct methods of suppressing the excitation force [5], active damping technologies do not require a complex displacement amplification device high drive voltage and are therefore easier to design and build [1].

Chen et al. conducted research on performance, testing and control strategies of the smart spring system [1–3, 11, 12]. Results of these studies demonstrate that significant damping effect of smart spring systems may have important future prospects in active damping technologies. Moreover Ni et al. [7–10] studied the influence of parameter analysis, system dynamic response analysis, and design methods on the vibration performance of smart spring systems.

In practical applications, the range of the control voltage with PZTA is limited. Studies related to the design of the PZTA structure are merited, to obtain optimal damping effects within a limited voltage range. To date, no literature has been published on the topic.

2 Damping Principle of Smart Spring

The smart spring damping system consists of a basic spring-mass system and active spring-mass system, as shown in Fig. 1.

The basic spring and active spring are arranged in parallel, with spring stiffness K1 and K2 respectively, and damping coefficients C1 and C2 respectively. Mass m1 represents the equivalent mass of the vibration structure and other components connected to the basic spring, and m2 is the equivalent mass of the PZTA and other components connected to the active spring. For convenience, we will refer to the basic spring—mass system as the main system, and the active spring—mass system as the auxiliary system.

The upper end of the main system withstands an external stimulus F(t) by passing the disturbance energy to a target object via the smart spring device, which is located in the transmission path of the vibration as it moves from the vibration source toward the target.

The spring is an active smart control device, wherein the PZTA adjusts the sliding friction Fd to control the degree of binding between the main system and auxiliary system. Furthermore, actively adjusting the dynamic impedance characteristics (stiffness, damping and mass properties) of the vibration control system can suppress the transmission of the vibration energy.

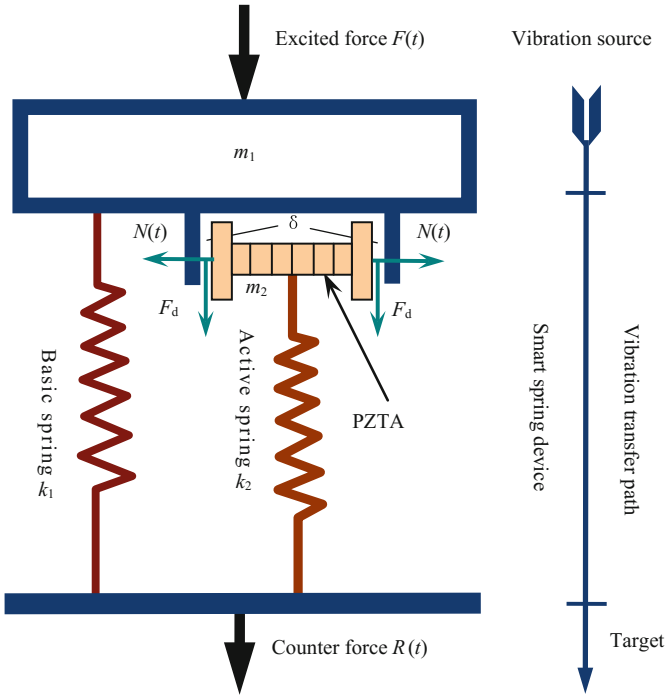


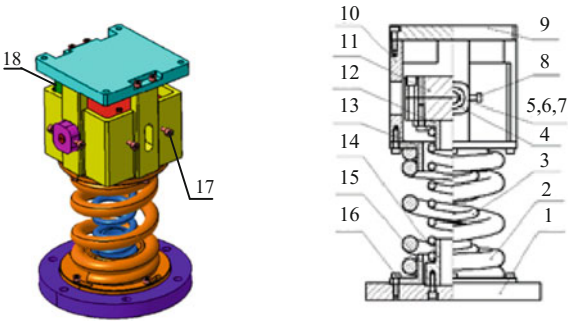
Fig. 1 Composition and damping principle of smart spring vibration system

3 Design of Test-Bed for Smart Spring Vibration Reduction System

A test prototype was designed and built, as shown in Fig. 2. The device is an independent modular component, connected to the controller system via the base and output connecting plate. The system provides continuous optimization and is expected to be relevant to a variety of applications.

The structure acts as a base for the active spring in the basic spring component, and therefore not only saves on installation space but also ensures symmetry of the structure and uniformity of the spring restoring force. The basic spring is connected to the output connection plate via the piezoelectric ceramic shield, which is placed in the protecting mask. The outer friction surface of the cover protects the piezoelectric ceramic. The shield is designed as a symmetrical structure with four friction surfaces: two working friction surfaces and two friction surfaces on standby used only if the primary working friction surfaces become worn out or torn.

Piezoelectric ceramic components with bidirectional outputs are proposed to ensure friction in the outer shield. The piezoelectric ceramic components can be subdivided into the piezoelectric ceramic actuator, output shaft, and friction block, making installing and disassembling the piezoelectric ceramic component more



1.Base; 2.Basic spring; 3.Active spring; 4.Friction block; 5.Output shaft; 6. Increase washer; 7.The screw for adjusting clearance; 8.Locking screw; 9. Output connecting plate; 10.The piezoelectric ceramic shield; 11.Bidirectional Output piezoelectric ceramic actuator (PZTA); 12. First flange; 13.Second flange; 14.Third flange; 15.Fourth flange; 16.Screw connection; 17.Set screws; 18.Standard clearance plate.

Fig. 2 Schematic diagram of test prototype

convenient. The output shaft of the piezoelectric ceramic passes through a hole in the friction surface of the shield and the friction block sits on top of the piezoelectric ceramic output shaft on the outer side of the friction surface. When the control voltage is switched on, the ceramic is compressed onto the friction surface.

To improve hardness and wear resistance, the friction surfaces are treated by carburization, and the friction block is anodized. Two planes of the friction block can be used, and one side can be used to replace the other once it is worn out, thus extending the lifespan of the device. Piezoelectric ceramics are brittle and their shearing resistance is poor. To prevent damage, an anti-shear design is adopted. Output displacement of the piezoelectric ceramic is directly related to the control voltage. Due to the limitations of the electromechanical coupling characteristic, both the control voltage and output displacement are limited, and maximum allowable free deformation is usually only in the region of a few tens of microns, as shown in Fig. 3. Moreover, the gap between the auxiliary system and main system must be easy to adjust. To this end, the thickness of the friction block is designed to be slightly larger than the length of the friction surface of the piezoelectric ceramic output shaft, and the gap adjusting screw is used to achieve a desired gap width. The friction block is then fixed by locking screws.

To ensure optimal performance of the damping device, the outer side surface of the piezoelectric ceramic and inner side of the shield must be in parallel, thus ensuring a constant width between them. Otherwise, the shield of the ceramic block may be locked into the friction. There is no relative motion between the basic spring and the cylindrical, therefore, adjustment of the damping device cannot be achieved. Considering manufacturing and installation errors, the lower end of the base spring member is fixed at one end. The installation hole of the fourth flange is

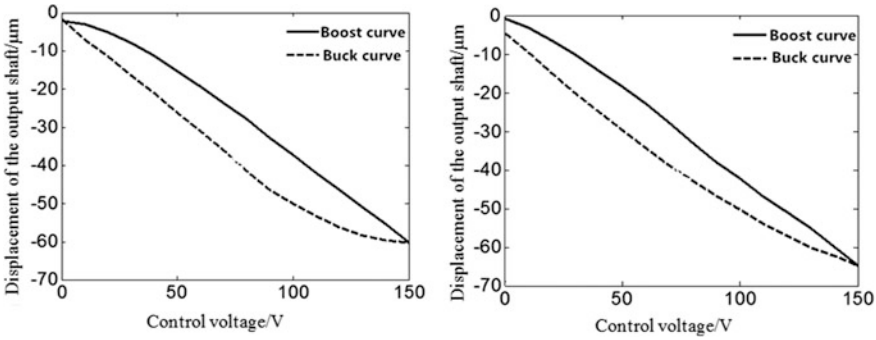


Fig. 3 Relationship between control voltage and output displacement of two way output piezoelectric ceramics

designed as a long circular hole. The entire assembly which is connected to the basic spring, can be rotated about its center within certain range. As shown in Fig. 2, the standard gap plate is inserted between the piezoelectric ceramic and protective cover of the piezoelectric ceramic on the side of the opening of the output connecting plate, using two threaded holes on opposite sides of the piezoelectric ceramic protection cover and two fastening screws. Once its position is fixed, the lower end of the basic spring component is fixed onto the base. Then, by loosening the two fastening screws, after the standard clearance plate is drawn out, the outer side of the piezoelectric ceramic actuator is guaranteed to be parallel to the inner side of the piezoelectric ceramic protective cover, creating a constant gap width between the two sides.

Furthermore, to improve the vibration reduction effect, the basic spring should be stiffer, i.e., the spring stiffness of the basic spring is larger than the spring stiffness of the active spring. The size of the basic spring components and structure of the active spring are limited and relatively small when the compression spring is used directly. The active spring and connection piece (first and third flange) are replaced with a cylindrical structure. The vibration reduction performance using both types of active spring structures were tested.

The basic parameters of the test system are: $k1 = 1.95 \times 105$ N/m, $m1 = 3.5$ kg, $m2 = 0.4$ kg, When compression structure is adopted by active spring $k_2 = 4.45 \times 104$ N/m. When a tubular structure is adopted for the active spring: k_2 is approximated as infinite.

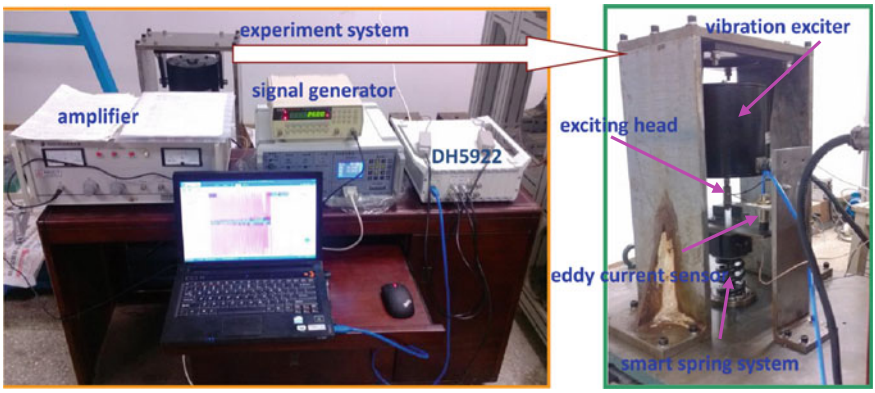


Fig. 4 Test on vibration reduction performance of smart spring vibration system

4 Testing Vibration Reduction Performance of Smart Spring System

The experimental setup for testing the damping performance of the smart spring is shown in Fig. 4. The excitation source of the test system is a vibration exciter. Control signals were generated using the signal generator, then passed to the vibrator through an amplifier. The exciting head measured the vibrating force and acceleration generated by the exciter. The maximum excitation force was 200 N. The vibration displacements were measured using an eddy current sensor within the test system, and a dynamic signal acquisition and analysis system (DH5922, Donghua, China) was used for measurements and further analysis. The control voltage of the PZTA was provided by a dedicated ceramic controller.

The maximum control voltage was set to 150 V, vertical stiffness of the ceramic was approximately 10 N/m, output force was approximately 1000 N, and response frequency was 500 Hz.

4.1 Aims and Methods

In this paper, no force sensor or output displacement feedback detection device in the two-way output type piezoelectric actuator exists and the piezoelectric ceramic exhibits hysteretic creep behavior and cannot obtain the control force. Therefore, the aim of this experiment was to test the actual damping effect of the smart spring device, as well as variation of the vibration reduction system under different control voltage and the influence of active spring parameters on vibration reduction performance of the system. The results were quantitatively compared with those of the

theoretical analysis. The specific process and methods of the experiment are as follows:

(1) For an active spring with tubular structure, frequency sweep tests were performed at an initial frequency of 10 Hz and termination frequency of 90 Hz. The frequency interval was 1 Hz with 0.5 s time interval. The control voltage was between 0 and 140 V and the amplitude was 20 V.
(2) Three excitation frequencies in the sub-critical, critical, and super-critical regions were selected for testing and the control voltage was applied in the same range as test 1.
(3) For an active spring with compression structure, tests 1 and 2 were repeated. The control force of the smart spring vibration damper can be controlled by inhibiting displacement of the piezoelectric actuator, thereby reducing the gap between the friction contact surface in order to increase the maximum value of the effective control force. Therefore, the gap between the contact surface was adjusted to zero.

4.2 Experimental Results and Analysis

Results of sweep tests under different control voltages are shown in Fig. 5. Vibration amplitudes of the system were reduced for all excitation frequencies, moreover, the maximum vibration absorption rate was close to 60%. The natural frequency of the test system was measured, as shown in Fig. 5a, approximately 37 Hz, and based on our theoretical calculation was 37.57 Hz. Therefore, the experimental and theoretical data are in agreement. After changing the control voltage during the test, the results produce a certain zero drift, therefore, only peak values were used to evaluate the damping properties of the system.

As shown in Table 1, an increase in control voltage enhances the vibration performance of the test system followed by a reduction in performance. The optimal displacement value is obtained between the 100–120 V, and the natural frequency of the system increases, consistent with theoretical predictions.

Vibration reduction performance of the smart spring vibration system was tested at a constant frequency of 25 Hz (subcritical region), 35 Hz (critical) and 45 Hz (supercritical) and three excitation frequencies were investigated. The amplitude and damping efficiency of the test system with tubular structure under constant excitation frequency are presented in Table 2. At an excitation frequency of 35 Hz or 25 Hz, damping performance of the system can be enhanced by increasing the control voltage. When the control voltage is greater than 100 V, the amplitude of the vibration decreases. At 45 Hz, the system undergoes a trans-critical process. As the control voltage increases, vibration performance of the system first increases, followed by decrease in performances. For an excitation frequency of 35 Hz, damping effects of the system are optimal (reaching a maximum of 82.29%).

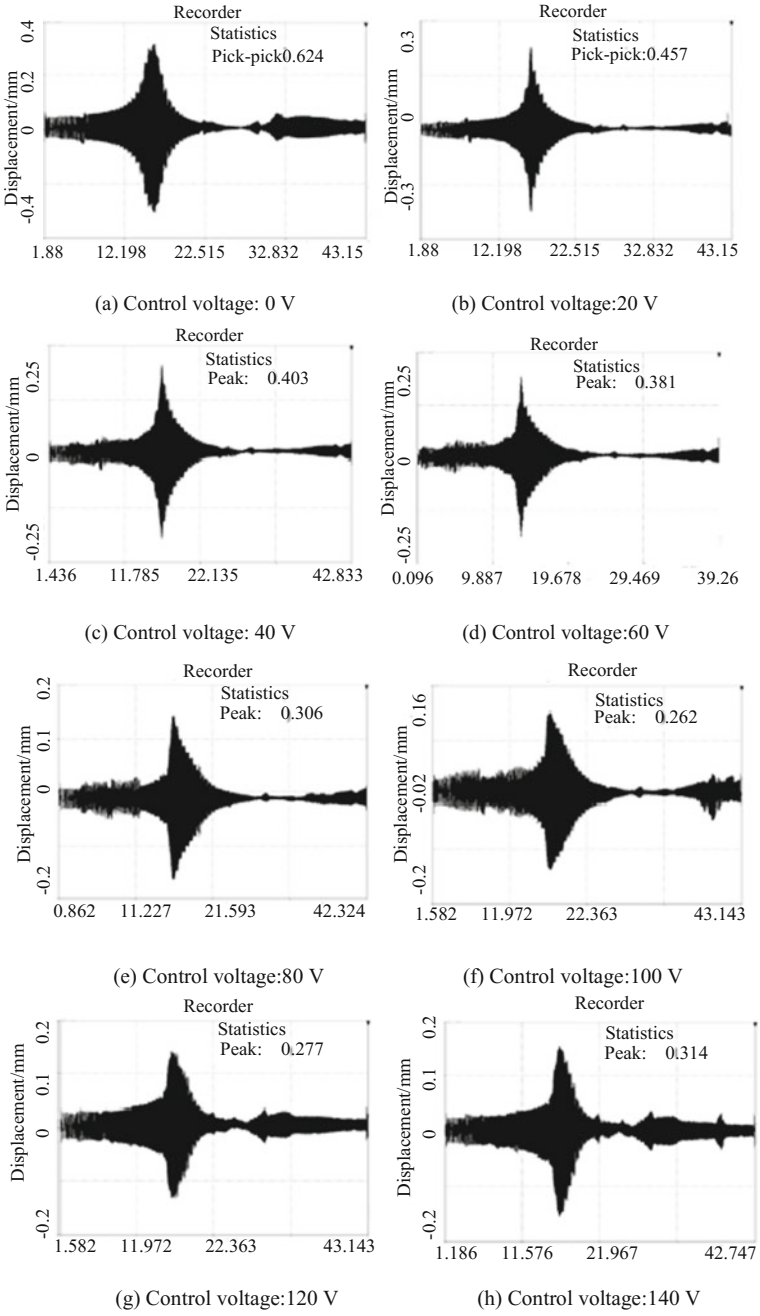


Fig. 5 Test results under different control voltages across critical process (time is on the horizontal axis)

Table 1 Amplitude and vibration reduction rate of test system with tubular structure for transcritical process

Excitation frequency	Control voltage (V)	0	20	40	60	80	100	120	140
10–90 Hz	Peak-peak value (mm)	0.624	0.457	0.403	0.381	0.306	0.262	0.277	0.314
	Damping efficiency (%)	0	26.76	35.42	38.94	50.96	58.01	55.61	49.68
	Natural frequency (Hz)	37	38.5	40	41.5	42.75	44	44.75	45.5

Table 2 Amplitude and damping efficiency of test system with tubular structure under different excitation frequencies

Excitation frequency	Control voltage (V)	0	20	40	60	80	100	120	140
35 Hz	Peak-peak value (mm)	0.367	0.299	0.291	0.258	0.131	0.09	0.069	0.065
	Damping efficiency (%)	0	18.53	20.71	29.7	64.31	75.48	81.2	82.29
25 Hz	Peak-peak value (mm)	0.097	0.092	0.082	0.066	0.052	0.044	0.043	0.041
	Damping efficiency (%)	0	5.15	15.46	31.96	46.39	54.64	55.67	57.73
45 Hz	Peak-peak value (mm)	0.138	0.129	0.115	0.105	0.09	0.08	0.102	0.114
	Damping efficiency (%)	0	6.52	16.67	23.91	34.78	42.03	26.09	17.39

Then, 25 Hz (57.73) and 45 Hz (42.03) are the worst. This suggests damping performance of the system is closely related to its operating frequency. The results demonstrate the feasibility of using the smart spring device proposed in this paper, and damping effects are clearly observed.

Considering the active spring with a compression spring structure, the natural frequency of the non-coupled system and coupled system are very close, and based on the theoretical analysis, damping performance of the system is worse. System damping effects are extremely weak and binding force is reduced when the basic spring and active spring were fully blinded. As shown in Table 3, damping performance of the test system with compression spring structure is greatly reduced compared to the tube structure. Sweep frequency test results show vibration reduction performance is optimal at a control voltage of 80 V and excitation frequency of 35 Hz. However, the vibration reduction performance is improved with a control voltage of 60 V. Both values are lower than those of the tubular structure.

Table 3 Amplitude and damping efficiency of test system with compressed structure under different excitation frequencies

Excitation frequency	Control voltage (V)	0	20	40	60	80	100	120	140
10–90 Hz	Peak-peak value (mm)	0.696	0.654	0.593	0.449	0.438	0.45	0.467	0.485
	Damping efficiency (%)	0	6.03	14.8	35.49	37.07	35.34	32.9	30.32
35 Hz	Peak-peak value (mm)	0.584	0.544	0.414	0.396	0.452	0.461	0.534	0.535
	Damping efficiency (%)	0	6.85	29.11	32.2	22.6	21.06	8.56	8.39
25 Hz	Peak-peak value (mm)	0.231	0.224	0.218	0.202	0.188	0.187	0.186	0.186
	Damping efficiency (%)	0	3.03	5.63	12.55	18.61	19.05	19.48	19.48

When the excitation frequency is 35 Hz and the control voltage is greater than 120 V, the system is close to the boundary state and natural frequency is close to the operating frequency, therefore, the vibration reduction effect is very weak. When the excitation frequency is 45 Hz, the system has almost no damping effect, and the results are not presented in the table. In summary, the test results are consistent with our theoretical analysis.

5 Conclusions

This paper presents the design of an active vibration damping device with a compact structure and relatively straightforward installation requirements. Adjustment of the initial gap width is convenient. Excitation testing of a single-level system of the smart spring damping system was also conducted. The frequency sweep test results under different control voltages show that the vibration of the system can be reduced and the maximum vibration absorption rate is nearly 60%. The excitation frequency testing for 25 Hz (subcritical), 35 Hz (critical) and 45 Hz (supercritical) demonstrated the damping properties of the smart spring system under constant frequency. The vibration performance of the active spring system with tubular spring is superior to the system with compression spring structure.

Acknowledgements Authors are pleased to acknowledge the financial support provided by National Natural Science Foundation of China (Grants No. 51505215)

References

1. Chen Y, Wickramasinghe VK, Zimcik DG (2005) Smart spring impedance control algorithm for helicopter blade harmonic vibration suppression. J Vib Control 11(4):543–560
2. Chen Y, Zimcik DG, Wickramasinghe VK et al (2004) Development of the smart spring for active vibration control of helicopter blades. J Intell Mater Syst Struct 15(1):37–47
3. Chen Y, Zimcik DG, Wickramasinghe VK et al (2003) Research of an active tunable vibration absorber for helicopter vibration control. Chin J Aeronaut 16(4):203–211
4. Gennaretti M, Poloni L, Nitzsche F (2003) 'Smart spring' identification for hovering rotor aeroelastic-stability augmentation. Aeronaut J 107(1071):233–240
5. Gu HC, Song GB (2007) Active vibration suppression of a flexible beam with piezoceramic patches using robust model reference control. Smart Mater Struct 16(4):1453–1459
6. Li X-P, Liao F-H, Chen X-D et al (2013) Optimization design for structure parameters of active vibration isolator. J Vib Eng 26(1):61–67
7. Ni D, Zhu R-P (2012) Influencing factors of vibration suppression performance for a smart spring device. J Vib Shock, 31(23):87–91, 98
8. Ni D, Zhu R-P, Bao H-Y, Hu Z-G (2013) Response analysis of smart spring vibration isolation system. J Vib Shock 32(15):161–167
9. Ni D, Zhu R-P, Bao H-Y, Lu F-X, Fu Q-J (2014) Parameters design of smart spring vibration suppression system based point of amplitude frequency characteristic curve cluster on common. J Aerosp Power 29(9):2121–2128
10. Ni D, Zhu R-P, Lu F-X et al (2014) Vibration transmissibility characteristics of smart spring isolation system. J Cent South Univ 21(12):4489–4496
11. Nitzsche F (2007) Direct-active and semi-active smart structure control systems for aeroelastic application. In: Varoto PS, Trindade MA (eds) Proceedings of the XII international symposium on dynamic problems of mechanics, Ilhabela: DINAME 2007, pp 1–13
12. Wickramasinghe VK, Chen Y, Zimzik DG (2008) Experimental evaluation of the smart spring impedance control approach for adaptive vibration suppression. J Intell Mater Syst Struct 19(2):171–179
13. Zhang T, Hongbiao HUANG, Bin WANG et al (2010) Stiffness amplification of actuator and governing error analysis in active vibration isolation. J Vib Shock 29(5):50–53 (in Chinese)

Single-Sensor Identification of Multi-source Vibration Faults Based on Power Spectrum Estimation with Application to Aircraft Engines

Shunming Li, Yu Xin and Xianglian Li

Abstract Identifying the vibration fault of aircraft engines in lack of sufficient sensor information is a big challenge to researchers due to the restriction of sensor installation location and the nature of complicated multi-source vibrations on turbine rotors. In this paper a new method of single-sensor identification for multi-source vibration faults is proposed based on the blind source separation (BSS), the empirical mode decomposition (EMD), and power spectrum estimation. First the observed single-channel multi-source vibration signal is decomposed by EMD decomposition method, so that the intrinsic mode function (IMF) component can be obtained and new observation signal can be reconstructed. Second the source number is to estimate using the ratio matrix of power spectral density function, then the observation signals are reconstructed based on the estimated source number. By blindly separating the mixed signals matrix composed of the reconstructed observation signals and the original measurement signal, the original multi-source vibration components are obtained. The final step is to analyze the ratio of power spectral density function of signals to identify the characteristics of multi-source vibrations. Simulation and experiment results for applications in the turbine rotor system validated the proposed method on decomposing vibration signals and identifying the characteristics of vibration faults.

S. Li (✉) · Y. Xin
College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
e-mail: smli@nuaa.edu.cn

Y. Xin
e-mail: xy_apte@163.com

X. Li
College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
e-mail: lixle@nuaa.edu.cn

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_36

1 Introduction

Separating and identifying the characteristics of multi-source vibration faults from single sensor observation is a promising signal processing technology [1]. For the vibration fault detection of turbine motors, due to the restriction of sensor installation location, the complicated characteristics of vibration fault signal can only be obtained from a few measurements [2]. In recent years, Blind Source Separation (BSS) method has been widely used for vibration signal processing [3, 4]. Some researchers utilized BSS method for the separation of rotor vibration signals [5, 6]. The BSS method has also been used in the area of mechanical vibration fault detection [7]. Restricted by the BSS algorithm itself, it is always assumed that the number of observation signals is not less than the number of vibration sources. However, this hypothesis doesn't stand for many real engineering practices. Blind source separation method with fewer measurement signals than vibration sources is known as "under determined blind source separation". The research of under determined blind source separation mainly includes three categories, i.e., source number estimation and separation algorithm [8], singular value decomposition based method [9], and potential function based method [10]. However, these existing methods are all essentially based on the sparse representation of source signal. The blind separation results will be affected by poor sparse representation. Although a few literatures presented blind separation methods with single-channel mechanical signal, no method has been proposed for the critical source number estimation with single-channel information [11]. Therefore, it still cannot be applied for practical complicated vibration source identification like in aircraft engines.

For the under determined vibration problem of multi-source rotor with the difficulty of installing multiple sensors, we proposed a new method with the combination of BSS algorithm and empirical mode decomposition (EMD) algorithm, which includes the following procedures: collecting single-sensor vibration signal, decomposing the vibration signal based on EMD algorithm, estimating the number of vibration sources based on power spectral density function, reconstructing the signal, blindly separating the signals matrix composed of reconstructed observation signals and the original signal. The feasibility of the proposed method are verified by simulation and real vibration fault signals of turbine rotor.

2 Basics of BSS and EMD Algorithms

An affiliation is required for all authors. Please create a footnote for the affiliation and use the "affiliation–footnote" style in it. Please style affiliations as in the footnoted example. The envelope in brackets is used to mark the corresponding author.

Blind source separation (BSS) is a type of processing method to estimate source signals only through sensor signals, with all the source signals and the channel

parameters unknown. The sensor signal is comprised of source signal and noise, expressed as N

$$\boldsymbol{X}(t) = \boldsymbol{A}\boldsymbol{S}(t) + \boldsymbol{N}(t) \tag{1}$$

where $\boldsymbol{X}(t) = [x_1(t), x_2(t), \ldots, x_M(t)]^{\mathrm{T}}$ are the M sensor signals, $\boldsymbol{S}(t) = [s_1(t), s_2(t), \ldots, s_N(t)]^{\mathrm{T}}$ are the N unknown source signals, t is discrete time, $\boldsymbol{A}$ is a mixing matrix of dimension $M \times N$, $\boldsymbol{N}(t) = [n_1(t), n_2(t), \ldots, n_M(t)]^{\mathrm{T}}$ are Gauss noise signals in the M sensors. For M dimension mixed signal vectors $\boldsymbol{X}(t)$, after founding a separation matrix, we can through

$$\boldsymbol{Y} = \boldsymbol{W}\boldsymbol{X} \tag{2}$$

to estimate the N independent source signals $\boldsymbol{S}(t)$. where $\boldsymbol{W}$ is a matrix of $M \times N$, which is called a separation matrix or a solution matrix.

The method of empirical mode decomposition (EMD) decomposes the original signal into multiple intrinsic mode function (IMF) components. In every moment, there only one frequency component of IMF signals.

3 Number Estimation and Separation of Source Signals from Single-Sensor Signal

3.1 Number Estimation of Source Signals

Blind source separation method requires that the number of observation signals is not less than the number of source signals. We need to estimate the number of source signals using the method based on power spectrum density function [12] when the source signals is unknown. The processing procedures for single sensor based identification include: First, we decompose the single sensor signal $x_1(t)$ by using EMD method to obtain its IMF, $x_{1\mathrm{imf}} = [c_1, c_2, \ldots, c_n, r_{1n}]^{\mathrm{T}}$. Then, we clear away the trend terms and false components in the IMF. Suppose the number of remaining IMF components of x1imf is M. Constructing an arbitrary $(N-1) \times M$ matrix, and multiply with $x_{1\mathrm{imf}}$, we can get the $N-1$ new observation signals, denoted as $x_1, x_2, x_3, \ldots\ldots, x_N$, $N \geq 3$. Combining these $N-1$ new observation signals with the original sensor signal $x_1(t)$ to form a new vector of observation signals $\boldsymbol{X}(t)$, with $\boldsymbol{X}(t) = [x_1, x_2, \ldots, x_{\mathrm{N}}]^{\mathrm{T}}$. The vector $\boldsymbol{X}(t)$ can be used to estimate the number of source signals.

The cross-power spectrum of any sensor signal $x_k(t)$ with $x_1(t)$ can be described as $P_{k1}^{x}(\omega_q)$, and cross-power spectrum of any observation signal $x_k(t)$ with $x_2(t)$ can be described as $P_{k2}^{x}(\omega_q)$, $k = 1,2, \ldots, N$. The ratio of them at frequency ω_q can be expressed as

$$\lambda_{k,2,1}(\omega_q) = \frac{P_{k2}^{x}(\omega_q)}{P_{k1}^{x}(\omega_q)} \tag{3}$$

For all the N observation signals, the average value of $\lambda_{k,2,1}(\omega_q)$, k = 1,2, …, N, *is* $\lambda_{2,1}(\omega_q)$ [13].

It is same way for the other IMF components and frequencies. Therefore, the matrix of power spectral density ratios at K frequency points with non same-frequency from any observation signal $x_k(t)$ with $x_1(t)$ and $x_i(t)$ can be described as

$$\boldsymbol{P}_1 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_{2,1}(\omega_1) & \lambda_{2,1}(\omega_2) & \cdots & \lambda_{2,1}(\omega_K) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N,1}(\omega_1) & \lambda_{N,1}(\omega_2) & \cdots & \lambda_{N,1}(\omega_K) \end{bmatrix} \tag{4}$$

The ratio matrix of power spectral density at V points with same-frequency can be described as

$$\boldsymbol{Q}_1 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_{2,1}(\Omega_1) & \lambda_{2,1}(\Omega_2) & \cdots & \lambda_{2,1}(\Omega_V) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N,1}(\Omega_1) & \lambda_{N,1}(\Omega_2) & \cdots & \lambda_{N,1}(\Omega_V) \end{bmatrix} \tag{5}$$

Similarly, the ratio matrix of power spectral density $\boldsymbol{P}_2$, $\boldsymbol{Q}_2$ can be obtained from observation signal $x_k(t)$ and $x_2(t)$, at K points with none same-frequency and at V points with same-frequency. Similarly, $\boldsymbol{P}_3$, $\boldsymbol{Q}_3$, $\boldsymbol{P}_4$, $\boldsymbol{Q}_4$, …, $\boldsymbol{P}_N$, $\boldsymbol{Q}_N$ can be obtained. The overall ratio matrix of power spectral density can be expressed as

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{Q}_1 \\ \boldsymbol{P}_2 & \boldsymbol{Q}_2 \\ \vdots & \vdots \\ \boldsymbol{P}_N & \boldsymbol{Q}_N \end{bmatrix} \tag{6}$$

If two columns or more than two columns are equal in the ratio matrix R of power spectral density, the corresponding frequencies must be non same-frequency that comes from the same source signal, which can determine the lower limit number of source signals. For those columns not equal or close to each other, the corresponding frequencies may be the same, and may also be frequency components corresponding to the single frequency but non same-frequency source signals, which can estimate the upper limit number of the source signals.

3.2 Identification of Multiple Vibration Sources from Single-Sensor Signal

After the number of source signals being estimated, the number of constructed observation signals required by the blind source separation method can be determined. The observation signals are reconstructed by using IMF components, and then the blind source separation can be realized only using a single sensor signal. Through a lot of simulation and experiment, we summarize the single channel signal blind identification method as following steps:

(1) Decompose the single sensor signal $x_1(t)$ by EMD to obtain the intrinsic mode function $x_{1\text{imf}}$.
(2) Use the IMF components to reconstruct $N - 1$ ($N \geq 3$ and N is a positive integer) observation signals, together with the original sensor signal to form a new vector of observation signals $\boldsymbol{X}(t) = [x_1, x_2, \ldots, x_N]^{\mathrm{T}}$.
(3) Estimate the number M of source signals using the power spectral density ratio matrix.
(4) With the estimated number M of source signals, reconstruct $M - 1$ new observation signals using IMF components.
(5) Using BBS to separate the new vector of signal $\boldsymbol{X}(t)$ composed of reconstructed $M - 1$ observation signals together with the original signal, to obtain the independent components.
(6) Analyze the independent components and identify the signal characteristics.

4 Simulation

Consider three mechanical vibration source signals, i.e.,

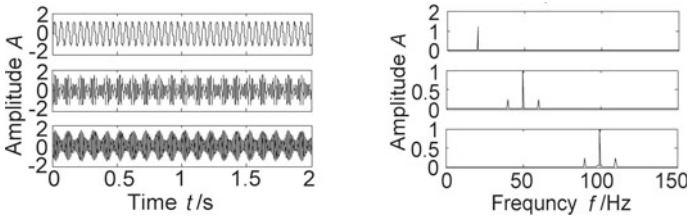
$$\begin{aligned} s_1 &= 1.2\cos(2\pi f_1 t + 10) \\ s_2 &= (1 + \beta \sin(2\pi f_r t))\sin(2\pi f_2 t) \\ s_3 &= (1 + \beta \sin(2\pi f_r t))\sin(2\pi f_3 t) \end{aligned} \tag{7}$$

where $f_1 = 20$ Hz, $f_2 = 50$ Hz, $f_3 = 100$ Hz, $f_r = 10$ Hz and $\beta = 0.5$. The sampling frequency is 1024 Hz. The three source signals in time domain and frequency domain are shown in Fig. 1, respectively.

Adding a Gaussian white noise signal to the source signal described in Eq. 7.

$$s_4 = \text{White Gaussian Noise} \tag{8}$$

with the same mixing matrix $\boldsymbol{A} = [0.37, 1.23, 0.66, 0.60]$, We can construct a primitive mixed signal by this mixing matrix.



(a) time domain (b) frequency domain

Fig. 1 Source signals

From the frequency domain in Fig. 2 we can see that the characteristics of source signals have been recovered from the separation signals. The figure with no obvious peak corresponds to Gaussian noise signal. All the separation signals contain noise, but it does not affect the identification and the frequency characteristic extraction for the source signals.

5 Multi-source Vibration Fault Identification of Turbine Rotor

We applied the proposed method to the fault diagnosis of an aircraft engine. The aircraft engine only has one turboshaft. The compressor shaft runs through low-pressure compressor, high pressure, and the combustion turbine. And the operation speed is 32,100 r/min. A vibration sensor is installed in the vertical direction of turbine. Sampling frequency is 12,800 Hz. The single-sensor vibration signal is shown in Fig. 3a, and its spectrum is shown in Fig. 3b. It can be seen from the spectrum in Fig. 3b that the signal has many primary vibration frequencies, which means the signal is a mixed multi-source vibration signal. The signal itself is

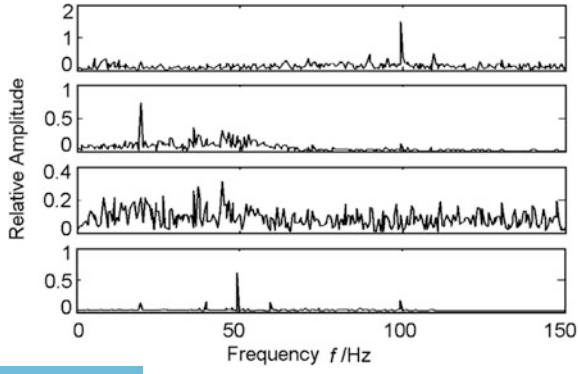
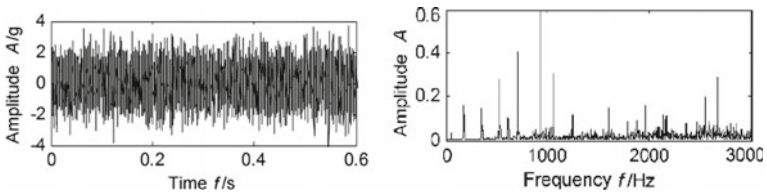


Fig. 2 Separation signals of signal with noise in frequency domain



(a) time domain (b) frequency-domain

Fig. 3 Vibration signal in the vertical direction of turbine

a mixture of multiple vibration source signals, don't need a 'hybrid' and it just need to separated.

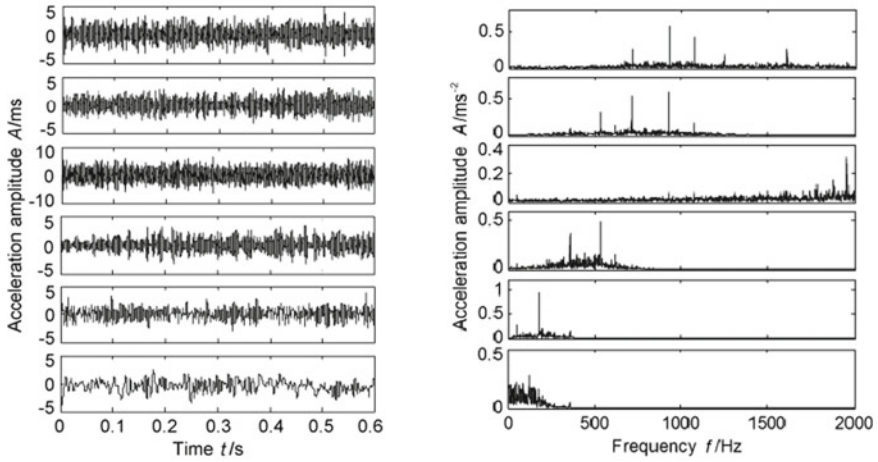
Decomposing the vibration signal by EMD method, 13 IMF components can be obtained. The last 4 almost have no waveform information. We select the first 9 order IMF components for the observation signal reconstruction. Using the reconstruction method and source number estimation method in Sect. 3, the ratios of power spectral density function are shown in Table 1.

Considering the high-frequency components are mainly from the noise signal, we focus on the frequency components below 2000 Hz. From Table 1 we can see that the ratios of power spectral density function corresponding to 177.9, 355.8, 535, and 1959 Hz are far from each other. Thus, these four frequency components may be from four single frequency signals. The ratios of power spectrum density function on 712.9, 926.6, and 1069 Hz are almost equal to each other, and the ratios of power spectrum density function on 1247 and 1604 Hz are approximately the same, which can be concluded that they are from the same sources. Therefore, we can estimate source signals are 6. With arbitrarily chosen mixing matrix $\boldsymbol{B}$= [0.4020 0.9983 0.5508 0.8392 1.1940 0.0983; 1.7734 1.7166 0.1701 0.3548 0.0268 1.1800], we can get five reconstruction signals. Using the second-order blind identification (SOBI) algorithm of blind source separation, we can further separate into six signals. The separation signals are shown in Fig. 4.

From Fig. 4b we can see that the fourth separate signal has two obvious peaks at 355.8 and 535 Hz, respectively. The frequency of 535 Hz is the rotor frequency

Table 1 Ratio of power spectral density function of the vibration signal of turbine rotor

Frequency (Hz)	$\lambda_{2,1}$	$\lambda_{3,1}$	$\lambda_{3,2}$
177.9	1.96	1.48	0.76
355.8	1.45	0.70	0.48
535.0	1.07	0.85	0.79
712.9	0.65	0.61	0.94
926.6	0.61	0.61	0.99
1069.0	0.64	0.64	0.99
1247.0	0.71	1.04	1.45
1604.0	0.73	1.06	1.45
1959.0	1.23	3.86	3.13



(a) time domain (b) frequency domain

Fig. 4 Separation signals in **a** time domain and **b** frequency domain

corresponding to the engine rotating speed of 32,100 r/min. The frequency of 355.8 Hz is the 2/3 times of rotor frequency, and the energy in this frequency component is larger. It can be inferred that there is the rotor imbalance component. The fifth separation signal contains frequency component of 177.9 Hz, it is the 1/3 times of rotor frequency. That is typical character of rubbing fault in the rotor system when 1/3 and 2/3 times of rotor frequency appear at the same time.

The main frequency components of the first separation signal and the second separation signal are 535, 712.9, 926.6, and 1069 Hz. These frequency components correspond to the rotor frequency, 4/3 times, 5/3 times, and 2 times of rotor frequency, respectively, which are further emphasis that there is some rubbing fault in the rotor system. The third and the sixth separation signals are mainly the noise signals.

After disassembling the turbine rotor engine, we found that there is rubbing problem in the turbine rotor due to rotor imbalance. This verifies the effectiveness of the proposed analysis method.

6 Conclusions

By combining the empirical mode decomposition (EMD), power spectrum estimation, and blind source separation (BSS), a new multi-source vibration fault identification method from single-sensor signal is proposed in this paper. Based on the ratio of power spectral density function and signal reconstruction using intrinsic mode function (IMF) components, the number of source signals can be estimated accurately, which provides crucial information for blind source separation method.

From the simulation results of separation signals from original noise-free signal and with-noise signal, we can conclude that the proposed method using single-sensor signal can extract the characteristics of the multiple-source signals with good anti-interference performance. Application to the fault diagnosis of an aircraft engine turbine rotor successfully verified demonstrated the effectiveness of the proposed method.

Acknowledgements The research was supported by National Natural Science Foundation of China (51675262) and also supported by the Project of National Key Research and Development Plan of China "New energy-saving environmental protection agricultural engine development" (2016YFD0700800) and the Open Project of State Key Laboratory for Strength and Vibration of Mechanical Structures (SV2015-KF-01). Special thanks should be expressed to the editors and anonymous reviewers for their valuable suggestions.

References

1. Li SM (2011) Aviation Industry Press of China. Blind source separation technology and its application in vibration signals
2. Li SM (2005) Blind source separation of rotor vibration Faults. J Aerosp Power 20(5):751–756
3. Liu HN, Liu CL, Huang YX (2011) Adaptive feature ex traction using sparse coding for machinery fault diagnosis. Mech Syst Signal Process 25:558–574
4. Liu RB, Randal S (2005) Blind source separation of internal combustion engine piston-slaps from other vibration signals. Signal Process 19:1196–1208
5. Li SM, Lei YB (2008) Blind separation of rotor vibration signals by second order non-steady arithmetic. J Propul Technol 29(6):747–752
6. Song Y, Liu CK, Li QH (2002) Denoise of rotor vibration signal based on higher-order cumulant. Chin J Aerosp Power 17(3):363–366
7. Popescu TD (2010) Blind separation of vibration signals and source change detection—application to machine monitoring. Appl Math Model 34:3408–3421
8. Tan BH, Yang ZY (2009) Source signals' number estimation and separation arithmetic in underdetermined blind separation. Sci China F 39(3):34–356
9. Shen YJ, Yang SP, Kong DS (2009) New method of blind source separation in under-determined mixtures based on singular value decomposition and application. Chin J Mech Eng 45(8):64–70
10. Zhang Y, Li BW, Wang YH (2010) Fault diagnosis based on underdetermined blind source separation using potential energy function. J Aerosp Power 25(1):218–223
11. Wu WF, Chen XH, Su XJ (2011) Blind source separation of single-channel mechanical signal based on empirical mode decomposition. Chin J Mech Eng 47(11):12–16
12. Li N, Shi T (2008) Estimation of blind source number based on power spectral density. Journal of Data Acquisition & Processing 23(1):1–7
13. Belouchrani A, Abed Meraim K, Cardoso JF, Moulines E (1997) Ablind source separation technique using second order statistics. IEEE Trans Signal Process 45(2):434–444

Centrifugal Compressor Diagnosis Using Kernel PCA and Fuzzy Clustering

X. Liang, F. Duan, D. Mba and B. Ian

Abstract Centrifugal compressors are critical components in most oil and gas industries. Performance degradation or unplanned shutdowns of the centrifugal compressors can significantly increase financial cost for a company. Through the years, the early fault diagnosis of centrifugal compressors based on monitoring performance parameters remains a challenge because such data can sometimes be large and complicated. Given machine learning techniques are suitable for high dimension and non-linear data, this paper uses seventy-four (74) performance parameters from an operational centrifugal compressor to identify 'fouling'. By employing Kernel PCA, the data dimension is reduced, then, after applying fuzzy c-means clustering, the fouling fault of the centrifugal compressor is detected. Compared to k-means and other hard clustering approaches, fuzzy clustering better solves the problem of which cluster the data points belong to, by providing corresponding membership values of each data point belong to several groups. This paper features the use of the Kernel PCA and fuzzy c-means clustering approach which hitherto has not been used for centrifugal compressor diagnosis.

Keywords Centrifugal compressor · PCA · Kernel PCA · Fuzzy clustering

X. Liang (✉) · F. Duan
School of Engineering, London South Bank University, London SE1 0AA, UK
e-mail: liangx3@lsbu.ac.uk

F. Duan
e-mail: duanf@lsbu.ac.uk

D. Mba
Faculty of Technology, De Montfort University, Leicester LE1 9BH, UK
e-mail: mbad@lsbu.ac.uk

B. Ian
Team Lead Technology–Rotating Equipment, Shell Global Solutions International, B.V., 2288 GS Rijswijk, Netherlands
e-mail: Ian.Bennett@shell.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_37

1 Introduction

In most oil and gas industries, centrifugal compressors are indispensable utilities for the main production processes [8]. One or more components degrade or fail, may cause efficiency loss or emergency shutdown, which will threaten the production and safety, probably resulting in economic losses. Therefore, the condition monitoring and fault diagnosis of centrifugal compressors are essential.

There are several papers being published on rotating machine diagnosis with relatively few focused on centrifugal compressor diagnosis; of those published, several have employed vibration analysis [1, 9, 12–14, 16, 18]. The diagnosis systems using performance parameters are limited, and they have generally utilized low dimensional data sets.

Compared with measured vibration indicators, the performance parameters are also sensitive, especially to the inside parts of the compressor. As the performance parameters change slightly or fluctuate abnormally before the occurrence of a failure, the diagnostic system using performance parameters can detect early faults and prevent the occurrence of unplanned shutdowns. In addition, when the acquired parameters are non-linear and have very high dimensions, it is usually difficult for an operator to estimate the health condition of the system or to discover an incipient fault. This challenge can be overcome with the use of principle component analysis, which is an efficient feature extraction method, and fuzzy clustering approaches [2, 15].

This paper presents a fault detection methodology using a high dimensional performance data set as inputs and employing Kernel PCA and fuzzy c-means clustering approach, which have seldom been used in centrifugal compressor diagnosis. The structure of the paper is organized as follows. Performance indicators of centrifugal compressors and diagnosis process are demonstrated in Sect. 2. Two data dimension reduction methods: principle component analysis (PCA) and Kernel PCA, and fuzzy c-means clustering approach are introduced in Sects. 3 and 4, respectively. In Sect. 5, a data set obtained from an industry maintenance log with 74 performance parameters is used as an input, and the diagnosis system realized fault detection of the compressor. Finally, the conclusions are drawn in Sect. 6.

2 Performance Monitoring of Centrifugal Compressor

2.1 Centrifugal Compressor Performance Indicators

Performance monitoring is typically used for identifying abnormal compressor problems as these performance parameters are usually already available on the controlling software of the machine [11]. Generally, when a fault occurs in the gas flow path, for example fouling, stalling and/or surge, the performance parameters

will change more rapidly than measured vibration parameters [10]. In addition, the measured performance parameters of the centrifugal compressor contain useful information that can reflect the health condition of the compressor. By detecting and analysing the abnormal variations of performance parameters, it can be used for predicting the occurrence of centrifugal compressor faults, identifying fault types, predicting fault locations, and informing predictive maintenance. Therefore, it has become one of the most used techniques in compressor diagnosis, especially suited for detection of fouling and surge.

The performance parameters are generally easy to measure. For centrifugal compressors, the typical performance parameters include outlet pressure, outlet temperature, bearing temperature, mass flow, actual power, polytropic efficiency etc.

2.2 *Diagnosis Process*

As many failure modes have measurable responses and develop over periods of time, appropriate condition monitoring parameters can provide valuable information about the condition of system. Early warning of potential failures can be given by properly analysing the condition monitoring data. Figure 1 depicts the process of a general fault diagnosis framework.

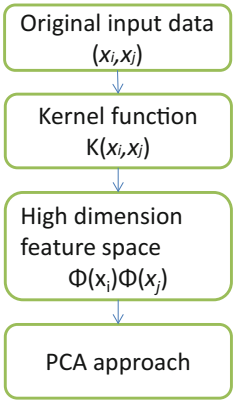


Fig. 1 Depicts the process of a general fault diagnosis framework

3 Data Dimension Reduction

3.1 PCA

Principal component analysis (PCA) is an effective method for data dimension reduction and feature extraction. By calculating the eigenvectors of the covariance matrix of the original input data, PCA linearly transforms a high-dimensional input vector into a low-dimensional with components uncorrelated [4].

Assume there are i numbers of samples, and each sample has j parameters, such as temperature, pressure etc. In order to eliminate the errors caused by different variable dimensions, the first step is data normalization:

$$\overline{X}_{ij} = \left(X_{ij} - E\left(X_{ij}\right)\right) / \sqrt{var\left(X_{ij}\right)} \tag{1}$$

where $\sqrt{var(X_{ij})}$ is the standard deviation of the original input data.

The next step is to create a covariance matrix and calculate the eigenvectors and eigenvalues of the covariance matrix. Then a new data set with uncorrelated variables and low dimension is obtained.

$$X = t_1 p_1^T + t_2 p_2^T + \cdots + t_j p_j^T + E = X_p + E \tag{2}$$

where p_j is the eigenvector of the covariance matrix X. t_j is the score matrix of the principal components. E represents a residual error matrix.

In the PCA approach, the first principal component p_1 corresponds to the projection having the largest variance. p_1 can reflect the most percentage of the original input data. The second component p_2 is orthogonal to the p_1 and again maximizes the variance of the data points projected on it.

The variance contribution rate of the kth principal component to the model can be explained by the percentage of the summary of total variances divided by each principal component variance.

$$C_k = \lambda_k / \left[\sum_{i=1}^{j} \lambda_i\right] \times 100\% \tag{3}$$

where λ_k is variance of kth principal component, $\sum_{i=1}^{j} \lambda_i$ is the summary of total variances.

The cumulative variance ratio explains the validation of the data dimension reduction. The cumulative variance ratio can be obtained by the equation below:

$$C_{cum} = \sum_{i=1}^{k} \lambda_i / \sum_{i=1}^{j} \lambda_i \times 100\% \tag{4}$$

The original data set is i × j matrix, and after PCA approach, the new data set will be an i × k matrix (k < j). Generally, the first k principal components will satisfy the usage if $\sum_{i=1}^{k} \lambda_i / \left[\sum_{i=1}^{j} \lambda_i\right] \geq 85\%$, which means that with first k principal components, the new low dimension data set can reflect more than 85% features of the original data set.

3.2 Kernel PCA

Kernel PCA can be regarded as a nonlinear PCA approach using the kernel method (H. [6]. The idea of kernel PCA is to firstly map the original input vectors (x_i, x_j) into a high-dimensional feature space $\Phi(x_i)\Phi(x_j)$ via a nonlinear function called kernel function and then calculate the linear PCA in the high-dimensional feature space [4].

By transforming the original inputs into a high-dimensional feature space, the Kernel PCA is able to extract more numbers of principal components than PCA. As Kernel PCA is doing PCA in high dimension space, it can also obtain a low dimension data set if only considering the first several eigenvectors.

4 Fuzzy c-Means Clustering

Objective function-based clustering approach is composed of two categories, hard clustering analysis and fuzzy clustering analysis [17]. The significant difference between them is that each data point only belongs to one group in the hard clustering analysis, while in fuzzy clustering analysis, each data point can belong to several groups with corresponding membership degrees.

In fuzzy clustering analysis, fuzzy c-means (FCM) clustering is an unsupervised machine learning technique that has been successfully applied to feature analysis, clustering, and classifier designs. Bezdek et al. [3] gave a general description of the fuzzy clustering. The objective of FCM clustering algorithm is to minimize the objective function:

$$\min J_m(U, v) = \sum_{k=1}^{N} \sum_{i=1}^{c} (u_{ki})^m d_{ki}^2 \tag{5}$$

Assume Y is the data, $Y = \{y_1, y_2, \ldots, y_N\} \subset R^n$. c is the number of clusters in Y; u_{ki} is each sample point's membership degree to the center of clusters k; m is the weighting coefficient, $1 \leq m < \infty$; d_{ki} is Euclidean distance from data sample point to the center of clusters; U is a $c \times N$ matrix, $U = [u_{ki}]$; v is vectors of centers, $v = (v_1, v_2, \ldots, v_c)$.

By minimizing Eq. 5, the input data can be divided into different clusters. At the same time, the centers of clusters can be calculated.

5 Case Study

To validate the diagnosis model, we used data from an operational centrifugal compressor suffering from a fouling fault. Compressor fouling is the most common problem causing compressor deterioration in oil and gas applications [3, 5]. Fouling is the deposition of particles in the fluid and sticks to the impellers and diffuser areas [7]. With comparison to the non-fouled compressor, the result of the fouling is increasing the turbulence levels, lifting the outlet temperatures of the compressor, lowering compressor outlet pressures and can cause up to 9% efficiency loss for the large gas compressor. Therefore, it is significant to detect and diagnose fouling problems to improve the performance of a compressor, and thus reducing financial costs.

The original data set has 74 performance parameters, including outlet pressure, outlet temperature, bearing temperature, mass flow, actual power, polytropic efficiency etc. The data set is from 27/04/2013 12:00:00 to 12/05/2013 04:00:00. By checking maintenance log, the first 275 h of compressor operation was healthy, but at 08-05-2013 22:47 (about 275 h after the beginning), some of their measurements passed the threshold. The machine did not shut down, no maintenance action was taken, and the maintenance log was recorded "Watch Keeping".

As the data dimension is very high, and some parameters are strongly correlated with each other, the first step for data analysis was to reduce data dimension. By applying kernel PCA, the original data which contained 74 parameters are reduced to only 3 dimensions whilst retaining 99% information of the original data. The obtained low dimension data are shown in Fig. 2a, b respectively.

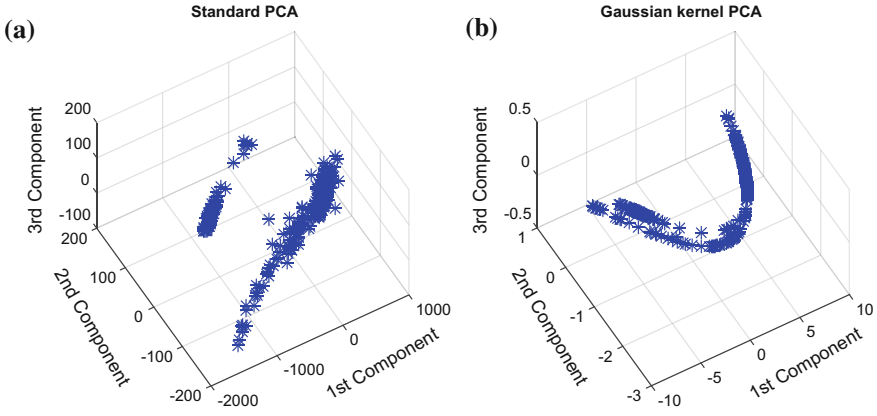


Fig. 2 **a** Data dimension reduction by standard PCA, **b** data dimension reduction by Kernel PCA

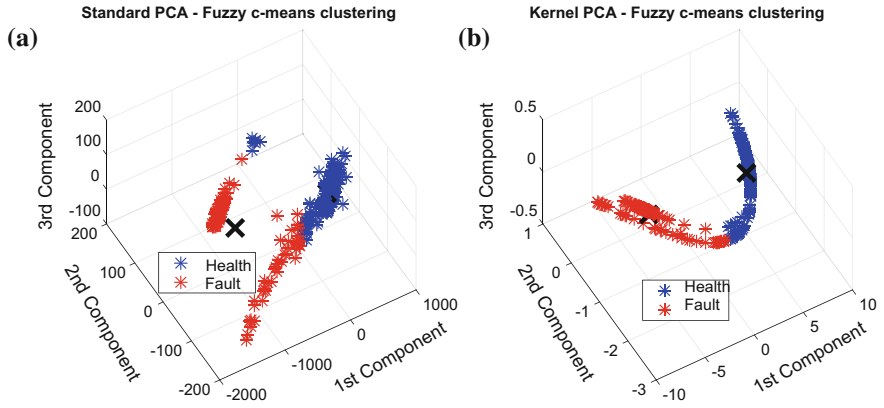


Fig. 3 **a** Fuzzy c-means clustering using data obtained by standard PCA, **b** fuzzy c-means clustering using data obtained by Kernel PCA

From Fig. 2a, b, it can be observed that the data dimension is reduced from 74 to 3. However, the fault data are mixed with healthy data, which need another step to distinguish the fault and healthy data.

The next step of fault diagnosis is data clustering after obtaining the new low dimension data set. The Fuzzy c-means clustering approach is employed, since it provides each point's membership degree to different cluster centres compared with the k-means algorithm and other hard clustering methods. It allows for more useful information and enhances decision making. The fuzzy c-means clustering method can automatically divide the data into two clusters, by using a few historical data training the fuzzy logic system, the healthy data and fault data can be identified. The result is shown in Fig. 3.

By employing the fuzzy clustering approach, the values 0 and 1 indicate no membership and full membership of fault, respectively. A membership value 0.5 was used as a threshold. From Fig. 4, it can be clearly observed that, at 110 h of operation, the membership value crosses the threshold, but then remain below the threshold. At 210 h of operation, the membership value crosses the threshold more frequently, which can be regarded as an incipient fault. These results validate actual condition from operation, i.e., first 275 h of compressor operation is known to be healthy though from 275 h the system shows degradation in the monitored data due to fouling.

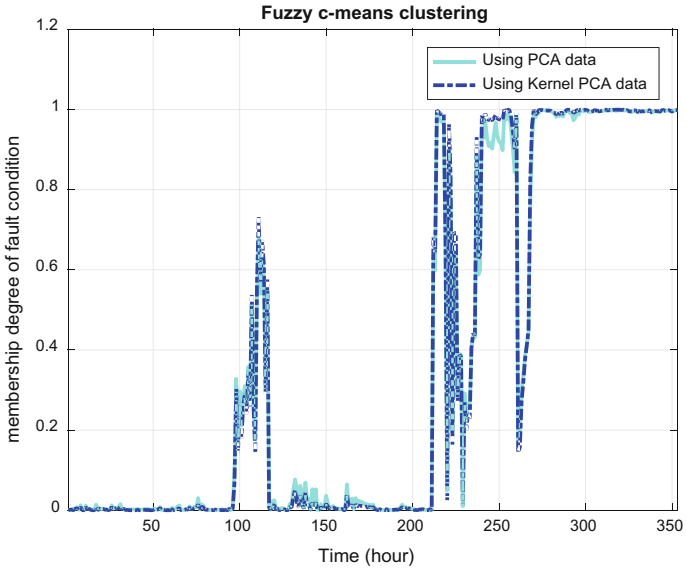


Fig. 4 Membership degree of fault condition

6 Conclusion

This paper presents a methodology for 'fouling' diagnosis of a centrifugal compressor. To validate the diagnosis model, operational data from a compressor with known with fouling are employed. The data contains seventy-four (74) performance parameters, including outlet pressure, outlet temperature, mass flow, and polytropic efficiency etc. As the data dimension is quite high, standard PCA and Kernel PCA are applied for data dimension reduction. The results show that the use of Kernel PCA can effectively reduce data dimension which is then categorized using fuzzy c-means clustering. The results show incipient fouling can be detected using fuzzy c-means clustering and Kernel PCA.

References

1. Al Yahyai M, Mba D (2014) Rotor dynamic response of a centrifugal compressor due to liquid carry over: a case study. Eng Fail Anal 45:436–448
2. Alpaydin E (2014) Introduction to machine learning. MIT Press
3. Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy c-means clustering algorithm. Comput Geosci 10(2–3):191–203
4. Cao LJ, Chua KS, Chong WK, Lee HP, Gu QM (2003) A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. Neurocomputing 55(1–2):321–336

5. Fouflias D et al (2010) Experimental investigation of the influence of fouling on compressor cascade characteristics and implications for gas turbine engine performance. Proc Inst Mech Eng Part J Power Energy 224(7):1007–1018
6. Hoffmann H (2007) Kernel PCA for novelty detection. Pattern Recognit 40(3):863–874
7. Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685
8. Kopanos GM, Xenos DP, Cicciotti M, Pistikopoulos EN, Thornhill NF (2015) Optimization of a network of compressors in parallel: Operational and maintenance planning—the air separation plant case. Appl Energy 146:453–470
9. Li B, Chow MY, Tipsuwan Y, Hung JC (2000) Neural-network-based motor rolling bearing fault diagnosis. IEEE Trans Ind Electron 47(5):1060–1069
10. Lu Y, Wang F, Jia M, Qi Y (2016) Centrifugal compressor fault diagnosis based on qualitative simulation and thermal parameters. Mech Syst Signal Process 81:259–273
11. Norton MP, Karczub DG (2003) Fundamentals of noise and vibration analysis for engineers. Cambridge University Press
12. Ogaji SOT, Singh R (2003) Gas path fault diagnosis framework for a three-shaft gas turbine. Proc Inst Mech Eng Part J Power Energy 217(2):149–157
13. Rai VK, Mohanty AR (2007) Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert-Huang transform. Mech Syst Signal Process 21(6):2607–2615
14. Ruiz-Cárcel C, Jaramillo VH, Mba D, Ottewill JR, Cao Y (2016) Combination of process and vibration data for improved condition monitoring of industrial systems working under variable operating conditions. Mech Syst Signal Process 66–67:699–714
15. Witten IH, Frank E, Hall MA, Pal CJ (2016) Data mining: practical machine learning tools and techniques. Morgan Kaufmann
16. Xu S, Jiang X, Huang J, Yang S, Wang X (2016) Bayesian wavelet PCA methodology for turbomachinery damage diagnosis under uncertainty. Mech Syst Signal Process 80:1–18
17. Yin S, Huang Z (2015) Performance monitoring for vehicle suspension system via fuzzy positivistic c-means clustering based on accelerometer measurements. IEEEASME Trans Mechatron 20(5):2613–2620
18. Yuan S-F, Chu F-L (2006) Support vector machines-based fault diagnosis for turbo-pump rotor. Mech Syst Signal Process 20(4):939–952

A Study of the Torsional Vibration of a 4-Cylinder Diesel Engine Crankshaft

T. R. Lin and X. W. Zhang

Abstract This paper presents a study on the torsional vibration of a multi-sectional diesel engine crankshaft using both discrete lumped-mass spring model and finite element model. A dynamical torsional stiffness matrix is established from the model and is used to calculate the torsional response of the crankshaft due to an external torsional excitation. The result is then compared to that of finite element analysis. It is found that although the result calculated using the discrete model agrees well with that obtained using finite element analysis in general, large discrepancies can also be observed between the two results. Results obtained from this study prompts for the need to establish a more accurate continuous 3-dimensional model for the multi-section crankshaft.

1 Introduction

Diesel engines play a critical role in the modern society. They are widely used in various industries such as mining, power plants, offshore platforms and military vehicles. Failures of diesel engines can result in significant economic losses and severe accidents such as human casualty. A continuous monitoring of engine performance, in a reliable and timely manner, is imperative for an early detection of a malfunction development in an engine so that correction measures can be deployed to prevent it from deteriorating further to become a functional failure.

Several condition monitoring (CM) techniques such as oil and wear analysis, in-cylinder pressure analysis, vibration analysis, acoustic emission analysis and instantaneous angular speed analysis have been developed for diesel engines in the last few decades. Each of these techniques can only be employed to deal with some

T. R. Lin (✉) · X. W. Zhang
School of Mechanical and Automotive Engineering, Qingdao University of Technology, 777 Jialingjiang Road, Huangdao District, Qingdao 266525, People's Republic of China
e-mail: trlin888@163.com

X. W. Zhang
e-mail: zhxw2016@sina.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_38

specific aspects in diesel engine monitoring. For instance, oil and wear analysis [24, 26] is typically employed to analyse faults or degrading issues of an engine caused by wear. In-cylinder pressure analysis [2] can provide a direct indication of the state of engine combustion process and the healthy condition change in a cylinder. However, its application is limited by the intrusive nature of the technique as well as the cost of high temperature pressure sensors. Vibration analysis [1, 17], on the other hand, is frequently employed to monitor the overall performing of a diesel engine and its auxiliary devices. Applications of vibration technique are often affected by the harsh working conditions (e.g., high dust and temperature) of a diesel engine and the clarity of vibration signal is hampered by the strong background noise during an engine operation.

Acoustic emission (AE) based CM technique [6, 7, 12, 14, 18, 19, 21, 22] has recently been successfully deployed to overcome the shortcomings of vibration monitoring technique in diesel engine applications. It is found that the signal acquired by the AE based technique would not be affected by the background noise and vibration generated by the reciprocating components of an engine attributing to the high frequency nature of the AE technique. Furthermore, AE signals generated by neighbouring cylinders would be largely attenuated when reach the location of an AE sensor installed to monitor a particular cylinder of a large multi-cylinder engine. Wu et al. [22] employed a semi-blind source separation technique to successfully separate the AE signals of a small diesel engine. Nevertheless, AE based technique also comes with inherent problems such as sensor non-linearity, costly data acquisition system and unwieldy large data. Instantaneous angular speed (IAS) analysis, as an alternative non-intrusive technique, is gaining wide applications for condition monitoring and fault detection of diesel engines [3, 4, 8, 11, 12, 23]. The variation of instantaneous angular speed links directly to the gas pressure torque produced by a diesel engine during the combustion process [3]. Combustion-related faults and other faults affecting the gas pressure can thus be diagnosed by the analysis of the IAS of the engine crankshaft. Nevertheless the performance of IAS analysis varies from one engine to another and depends largely on the cylinder number of an engine under monitoring. For instance, it has been pointed out by Charles et al. [4] that IAS technique is mainly effective when the cylinder number of an engine is less than 8. This restrictive condition can be relaxed a bit in detecting misfiring cylinders of a diesel engine with relatively large cylinder number when presenting the IAS waveform in a polar coordinate system [4]. Most recently, Lin et al. [15] proposed an Instantaneous Power Output Measurement (iPOM) technique in a preliminary study by utilizing the IAS technique and the instantaneous torque output of an engine determined by a crankshaft model. The experimental measurement on a mining engine as presented in the paper found that the technique is sensitive to the external loading conditions of an engine under monitoring.

Traditionally, the engine crankshaft is modelled as lumped spring-mass model [9, 10, 16, 20, 25] where each section of the crankshaft is considered as a discrete lumped inertia and the crankshaft is connected by torsional elastic springs. For instance, Kallenberger et al. [10] estimated the torque output of a 4-cylinder diesel

engine by utilizing a discrete lumped mass-spring model in conjunction with the in-cylinder pressure and engine speed measurement. Desbazeille et al. [5] presented a discrete model in analyzing the angular speed variation of an engine crankshaft which was then verified by a comparison with the measured angular speed variation. The model was successfully employed to identify the faulty cylinder of a 20-cylinder diesel engine.

Simplifications were made in determining the stiffness, damping and inertia of a discrete crankshaft model, though simplifications can vary from cases to cases and may not be accurate enough to capture the precise change of IAS or torque output signals of an engine in practical applications. To overcome this problem, a 3-dimensional torsional vibration model of a multi-cylinder engine crankshaft is developed in this paper so that a more accurate dynamic stiffness matrix of the engine crankshaft can be obtained in deriving the instantaneous output torque. The instantaneous torque and the measured IAS signal then form the basis of the iPOM technique for the real-time engine cylinder health monitoring [13].

The rest of this paper is organized as follows: Sect. 2 describes the 4-cylinder diesel engine and the multi-sectional crankshaft under study. Section 3 presents a lumped spring-mass model of the crankshaft and compares the predicted tangential displacement of the crankshaft with that obtained from finite element analysis. The main findings of this study are summarized in Sect. 4.

2 A Description of the 4-Cylinder Diesel Engine and the Engine Crankshaft

Figure 1 shows the 4-cylinder diesel engine under study. A Kistler high temperature pressure sensor is installed into the 4th cylinder of the diesel engine to monitor the pressure change of the cylinder. Other types of CM sensors such as AE sensors, accelerometers and angular displacement sensor have also been installed to monitor the condition of the engine though signals from these sensors are not used in this paper and thus will not be elaborated further.

Figure 2a shows the graphical presentation of the 4-cylinder engine crankshaft which is made up of crank pins, crank arms and four crankshaft sections and an

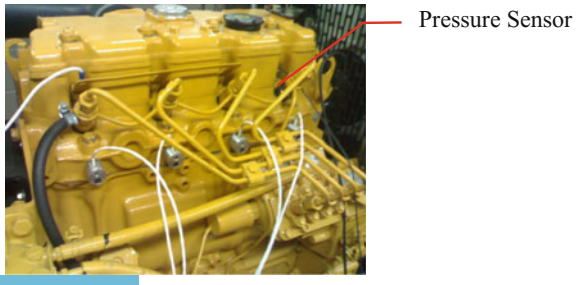


Fig. 1 A graphical illustration of the 4-cylinder diesel engine

output shaft. A schematic sketch of the projection of the crankshaft in 2-dimensional plane is also presented in the figure for a better illustration in the development of the mathematical model described in the following text.

3 A Traditional Lumped Mass-Spring Model of the Crankshaft

Traditionally, the torsional vibration of an engine crankshaft is modelled by discrete lump mass-spring models (Espadafor et al. 2014; Östman et al. 2008) [10, 16, 20]. Figure 3 shows the partition of the crankshaft torsional inertia in the discrete model and the corresponding lumped-mass torsional vibration model of the crankshaft. The 4-cylinder engine crankshaft is considered to be comprised by four lumped torsional inertia which are inter-connected by torsional elastic springs as shown in the figure.

Based on the discrete model described by Fig. 3b, a governing differential equation for the torsional vibration of the crankshaft can be written as:

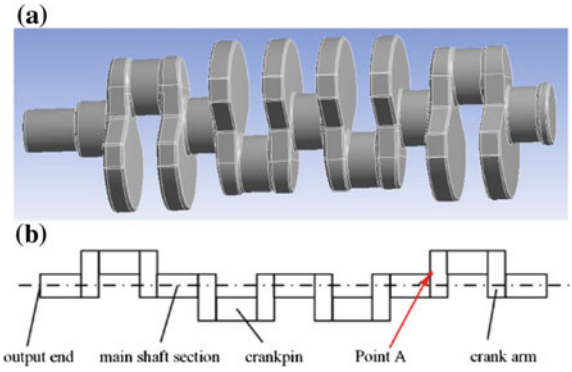


Fig. 2 **a** 3-dimensional model of a 4-cylinder crankshaft, **b** a schematic sketch of the crankshaft in a 2-dimensional plane

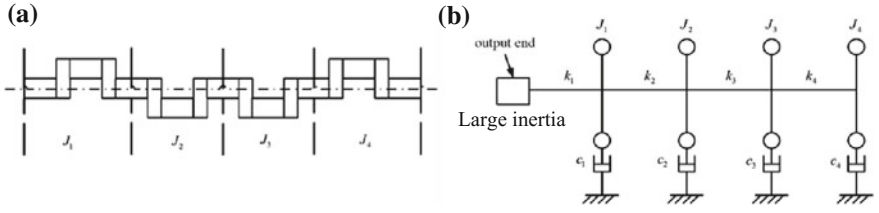


Fig. 3 **a** A graphical illustration of the partition of the crankshaft torsional inertia, **b** the lumped spring-mass model

$$[J]\left\{\ddot{\theta}\right\} + [C]\left\{\dot{\theta}\right\} + [K]\{\theta\} = \{T\} \tag{1}$$

where

$$[J] = \begin{bmatrix} J_1 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ 0 & 0 & J_3 & 0 \\ 0 & 0 & 0 & J_4 \end{bmatrix}$$ is the torsional inertia matrix,

$$[C] = \begin{bmatrix} c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & c_3 & 0 \\ 0 & 0 & 0 & c_4 \end{bmatrix}$$ is the damping matrix,

$$[K] = \begin{bmatrix} k_1 + k_2 & -k_2 & 0 & 0 \\ -k_2 & k_2 + k_3 & -k_3 & 0 \\ 0 & -k_3 & k_3 + k_4 & -k_4 \\ 0 & 0 & -k_4 & k_4 \end{bmatrix}$$ is the torsional stiffness matrix,

$$\{T\} = \begin{Bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{Bmatrix}$$ is the external torque applied to each section of the crankshaft, and

$$\{\theta\} = \begin{Bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{Bmatrix}$$ is the angular displacement vector of the discrete crankshaft model.

Table 1 lists the numerical values of the discrete crankshaft torsional inertia and the stiffness for each part of the 4-cylinder diesel engine which are estimated based on the finite element model of the crankshaft.

Table 1 The numerical values of the torsional inertia and stiffness

Inertia (kg m^2)	Stiffness (N m/rad)
$J_1 = 0.0615$	$k_1 = 6.612 \times 10^6$
$J_2 = 0.0607$	$k_2 = 9.667 \times 10^6$
$J_3 = 0.0607$	$k_3 = 9.667 \times 10^6$
$J_4 = 0.0619$	$k_4 = 6.365 \times 10^6$

Neglecting the damping and assuming harmonic, linear motions of the crankshaft, Eq. 1 can be further simplified to take the form:

$$\begin{bmatrix} k_1+k_2-J_1\omega^2 & -k_2 & 0 & 0 \\ -k_2 & k_2+k_3-J_2\omega^2 & -k_3 & 0 \\ 0 & -k_3 & k_3+k_4-J_3\omega^2 & -k_4 \\ 0 & 0 & -k_4 & k_4-J_4\omega^2 \end{bmatrix} \begin{Bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{Bmatrix} = \begin{Bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{Bmatrix} \tag{2}$$

Assuming the diesel engine is operated at a constant speed of 1500 rpm to give an angular speed of $\omega = 50\pi$ rad/s. Substituting this value and the parameters listed in Table 1 into Eq. 2 to have:

$$\begin{Bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{Bmatrix} = \begin{bmatrix} 1.628\times10^7 & -9.667\times10^6 & 0 & 0 \\ -9.667\times10^6 & 1.933\times10^7 & -9.667\times10^6 & 0 \\ 0 & -9.667\times10^6 & 1.603\times10^7 & -6.365\times10^6 \\ 0 & 0 & -6.365\times10^6 & 6.364\times10^6 \end{bmatrix}^{-1} \begin{Bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{Bmatrix} \tag{3}$$

According to Eq. 3, once the torque output curve produced by each cylinder is known, the angular displacements of the 4 crankshaft sections can be determined. To determine the pressure torque produced by each cylinder of the engine, the pressure signal of each cylinder needs to be measured. The pressure in Cylinder 4 is measured by the pre-installed Kistler pressure sensor as shown in Fig. 1. The other three cylinders are assumed to have the same pressure profiles as that of Cylinder but with a phase delay according to the firing sequences of the cylinders. The

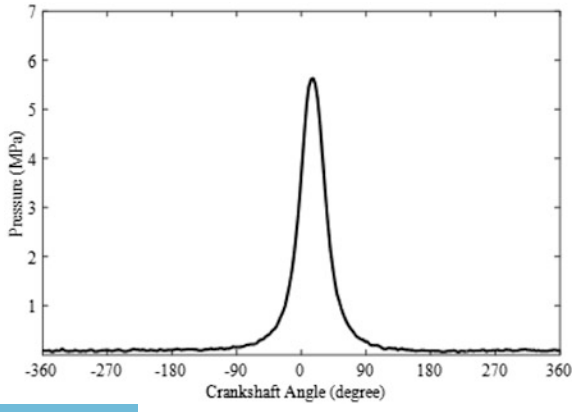


Fig. 4 The measured pressure profile of cylinder 4

pressure signal is converted into the true physical unit based on the information provided in the sensor calibration chart and is shown in Fig. 4.

Once the pressure profile of the cylinder is known, the gas pressure torque produced by the cylinder can be calculated by [20]

$$T_{gas} = P(\alpha) \cdot \frac{dV}{d\alpha} = -P(\alpha) \cdot A \cdot \left(r \cdot \sin(\alpha) + 0.5 \frac{r^2 \sin(2\alpha)}{\sqrt{l^2 - r^2 \sin^2(2\alpha)}} \right) \quad (4)$$

where P is the in-cylinder pressure as shown in Fig. 4, α is crankshaft angle, $l = 0.263\,\text{m}$ is the length of the connecting rod, $r = 0.05\,\text{m}$ is the radius of the crankshaft and $A = 0.00554\,\text{m}^2$ is the area of the piston.

Furthermore, there is inertia torque produced by the reciprocating components of the cylinder which can be calculated by [20]

$$T_{rec} \approx m_{rec}\omega^2 r^2 \left\{ \left[0.5\frac{r}{l} + 0.125\frac{r}{l} \right] \sin(\alpha) - \left[1 + 0.0625\left(\frac{r}{l}\right)^4 \right] \sin(2\alpha) - \left[1.5\frac{r}{l} + 0.56\left(\frac{r}{l}\right)^3 \right] \sin(3\alpha) - \left[0.5\left(\frac{r}{l}\right)^4 \right] \sin(4\alpha) \right\} \quad (5)$$

where $m_{rec} = 4\,\text{kg}$ is the total mass of the reciprocating components.

Figure 5 shows the calculated gas pressure torque, inertia torque and the resultant torque (which is the sum of the two) of the cylinder. Figure 5 shows that the resultant torque of each cylinder applied on the crankshaft varies according the crank position within each engine cycle. Assuming all cylinders have the same torque output profile and according to the cylinder firing sequence, i.e., 1–3–4–2 [14, 18, 22], the torque out of the other cylinders can be determined by shifting the pressure profile in Fig. 5 by 180°, 360° and 540° respectively. Summing up all the

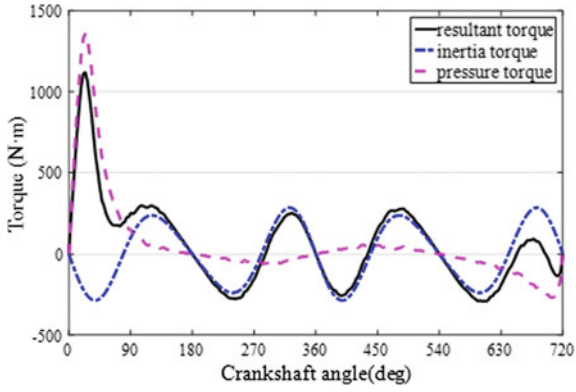


Fig. 5 The calculated gas pressure torque, inertia torque and resultant torque of cylinder 4

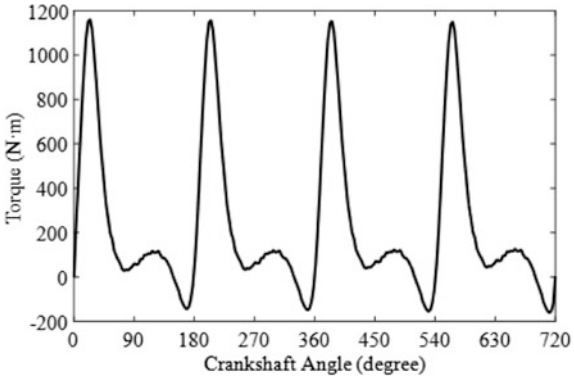


Fig. 6 The total torque outputs of the 4-cylinder diesel engine

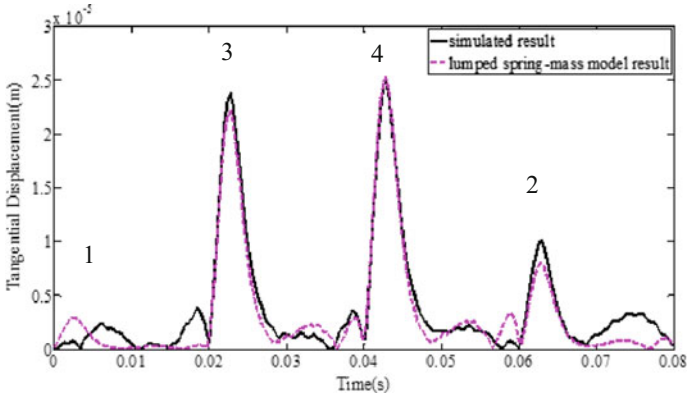


Fig. 7 A comparison of the tangential displacement at Point A calculated using the discrete model and finite element model. *Note* The numbers (1–4) shown in the figure indicates the combustion sequence of the engine

torque output of the 4 cylinders to have the total output torque profile produced by the engine which is shown in Fig. 6.

The torque output profiles of the four cylinders are used in Eq. 3 to calculate the angular displacements of the four main shaft sections. The tangential displacement at Point A of the shaft section next to the crank arm of Cylinder 4 as shown in Fig. 2 is calculated and shown in Fig. 7. Also shown in the figure is that obtained using finite element analysis for comparison.

It is showed in Fig. 7 that the tangential displacement at Point A calculated by both models agree well in general, particularly at the largest peak (note the peak is induced by the firing of Cylinder 4 which has the shortest transmission path to the calculation location, Point A) though the discrepancy induced by the lumped spring-mass model increases as the firing cylinders are away from the calculation

location. For instance, the error of the calculated tangential displacement by the lumped parameter model increases following the order Cylinder 3, Cylinder 2 and Cylinder 1 as shown in Fig. 7. The discrepancy can lead to inaccuracies if the discrete torsional vibration model of the crankshaft as described by Eq. 3 is used to determine the output torque using the measured angular displacements in practical situations. To overcome this problem, a 3-dimensional torsional vibration model of a multi-cylinder engine crankshaft needs to be developed so that a more accurate dynamic stiffness matrix can be obtained for the crankshaft to be used in the derivation of the instantaneous output torque of a diesel engine. The development of a 3-dimensional torsional model will be presented in a separate paper due to the limit scope of this paper.

4 Conclusion

A discrete lumped-mass spring model is used to describe the torsional vibration of a 4-cylinder diesel engine crankshaft as a preliminary study in the development of a novel non-intrusive diesel engine monitoring technique—Instantaneous Power Output Measurement (iPOM) technique. A finite element model of the crankshaft has also been established to provide a separate result to verify the torsional displacement of the crankshaft calculated using the discrete torsional vibration model. It was found that although the torsional response of the crankshaft calculated by the discrete model agrees well with that obtained from finite element analysis in general, large discrepancies can also be observed in the results. This then limits the usefulness of the discrete model in our attempt of establishing a novel instantaneous engine output measurement tool and prompts for the development of a continuous 3-dimensional crankshaft model, which will be described in a coming work.

Acknowledgements The financial support from Shandong provincial government of the People's Republic of China through the privileged "Taishan Scholar" program for this work is gratefully acknowledged.

References

1. Barelli L, Bidini G, Burrati C, Mariani R (2009) Diagnosis of internal combustion engine through vibration and acoustic pressure non-intrusive measurements. Appl Therm Eng 29(8–9):1707–1713
2. Carlucci AP, Chiara FF, Laforgia D (2006) Analysis of the relation between injection parameter variation and block vibration of an internal combustion diesel engine. J Sound Vib 295(1):141–164
3. Charchalis A, Dereszewski M (2013) Processing of instantaneous angular speed signal for detection of a diesel engine failure. Math Probl Eng 1:112–128
4. Charles P, Sinha JK, Gu F, Lidstone L, Ball AD (2009) Detecting the crankshaft torsional vibration of diesel engines for combustion related diagnosis. J Sound Vib 321(3):1171–1185

5. Desbazeille M, Randall RB, Guillet F, Badaoui ME, Hoisnard C (2010) Model-based diagnosis of large diesel engines based on angular speed variations of the crankshaft. Mech Syst Signal Process 24(5):1529–1541
6. Dykas B, Harris J (2017) Acoustic emission characteristics of a single cylinder diesel generator at various loads and with a failing injector. Mech Syst Signal Process 93:397–414
7. Elamin F, Gu F, Ball A (2010) Diesel engine injector faults detection using acoustic emissions technique. Modern Appl Sci 4(9):3–13
8. Gu F, Yesilyurt I, Li Y, Harris G, Ball A (2006) An investigation of the effects of measurement noise in the use of instantaneous angular speed for machine diagnosis. Mech Syst Signal Process 20(6):1444–1460
9. Guo YB, Li WY, Yu SW et al (2017) Diesel engine torsional vibration control coupling with speed control system. Mech Syst Signal Process 94:1–13
10. Kallenberger C, Hanedovic H, Zoubir AM (2008). Evaluation of torque estimation using gray-box and physical crankshaft modeling. In: IEEE international conference on acoustics, pp 1529–1532
11. Li Y, Gu F, Harris G, Ball A, Bennett N, Travis K (2005) The measurement of instantaneous angular speed. Mech Syst Signal Process 19(4):786–805
12. Lin TR, Tan ACC (2011a) Charactering the signal pattern of a four-cylinder diesel engine using acoustic emission and vibration analysis. In: Proceedings of world conference on acoustic emission-2011 Beijing, Chinese Society for Non-destructive Testing, Beijing, P. R. China, pp 506–515
13. Lin TR, Tan ACC, Mathew J (2013) The development of the instantaneous power output measurement technique for condition monitoring of diesel engines. In: Proceedings of the 8th world congress on engineering asset management, 30 Oct–1 Nov 2013, Hong Kong, P. R. China, pp 1–14
14. Lin TR, Tan ACC, Mathew J (2011) Condition monitoring and diagnosis of injector faults in a diesel engine using in-cylinder pressure and acoustic emission techniques. Dyn Sustain Eng 1:454–463
15. Lin TR, Tan ACC, Ma L, Mathew J (2011c) Estimating the loading condition of a diesel engine using instantaneous angular speed analysis. In: Engineering asset management 2011: Proceedings of the sixth annual world congress on engineering asset management, Duke Energy Center, Cincinatti, Ohio. Springer, Berlin, pp 259–272
16. Lin TR, Tan ACC, Ma L, Mathew J (2015) Condition monitoring and fault diagnosis of diesel engines using instantaneous angular speed analysis. Proc Inst Mech Eng Part C: J Mech Eng Sci 229(2):304–315
17. Liu S, Gu F, Ball A (2006) Detection of engine valve faults by vibration signals measures on the cylinder head. Proc Inst Mech Eng Part D: J Automob Eng 220(3):379–386
18. Lowe DP, Lin TR, Wu W, Tan ACC (2011) Diesel knock combustion and its detection using acoustic emission. J Acoust Emiss 29:78–88
19. Mba D (2006) Development of acoustic emission technology for condition monitoring and diagnosis of rotating machines: bearings, pumps, gearboxes, engines and rotating structures. Shock Vib Dig 38(1):3–16
20. Mendes AS, Meirelles PS, Zampieri DE (2008) Analysis of torsional vibration in internal combustion engines: modeling and experimental validation. Proc Inst Mech Eng Part K: J Multi-body Dyn 222(2):155–178
21. Nivesrangsan P, Steel JA, Reuben RL (2007) Source location of acoustic emission in diesel engines. Mech Syst Signal Process 21(2):1103–1114
22. Wu W, Lin TR, Tan ACC (2015) Normalization and source separation of acoustic emission signals for condition monitoring and fault detection of multi-cylinder diesel engines. Mech Syst Signal Process 64–65:479–497
23. Yang JG, Pu LJ, Wang ZH, Zhou YC, Yan XP (2001) Fault detection in a diesel engine by analyzing the instantaneous angular speed. Mech Syst Signal Process 15(3):549–564
24. Yin YH, Wang WH, Yan XP, Xiao HL, Wang CT (2003) An integrated on-line oil analysis method for condition monitoring. Meas Sci Technol 14(11):1973–1977
25. Yu YH, Yang JG, Zhou PL (2011) Fault diagnosis of a diesel engine by using the analysis of instantaneous angular speed with a flexible model. Intern J Veh Noise Vib 7(4):365–385
26. Zhang YL, Mao JH, Xie YB (2010) Engine wear monitoring with OLVF. Tribol Trans 54 (2):201–207

Engineering Asset Management for Various Power Sources: Common Concepts and Specificities

Jérôme Lonchampt, Karine Aubert, Emilie Dautrême and Roman Sueur

Abstract This paper describes a generic framework that can be used for managing physical assets in the field of electricity generation. This framework is based on a model describing both the technical and economic context of the assets. This paper presents how this generic method has been adapted to fit some specificities of various types of power generation such as nuclear, fossil, hydropower or wind-turbines.

1 Introduction

Asset management processes, focused on realizing value from physical assets, have been developed for years, but for the last one or two decades, these methods have been going from qualitative or semi-qualitative ones to quantitative management methods in order to support decision making. Asset management is a very important issue for the electric utility company EDF as it operates hundreds of production units for a total of 130 GW. The purpose of this paper is to describe the generic framework developed at EDF R&D to analyse and quantify physical assets management strategy and how it can be used to support decision making. While this generic framework can be applied to any type of power generation (nuclear, fossil, hydropower, wind-turbine…), or even any kind of physical asset, each power source has its own specifities. These characteristics will described in this document, as well as how the generic methods and tools can be adapted to address associated issues.

J. Lonchampt (✉) · K. Aubert · E. Dautrême · R. Sueur
Industrial Risk Management Department, EDF R&D, Chatou, France
e-mail: jerome.lonchampt@edf.fr

K. Aubert
e-mail: karine.aubert@edf.fr

E. Dautrême
e-mail: emilie.dautreme@edf.fr

R. Sueur
e-mail: roman.sueur@edf.fr

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_39

2 Engineering Asset Management for Power Generation

2.1 EAM System

Amadi-Echendu et al. [1] gives a very detailed and thoughtful definition of Engineering Asset Management as being an integrative management embracing all dimensions related to physical assets. It should then cover all areas such as data management, resources management, life cycle analysis… The research area covered by the EAM team at EDF R&D is focused on life cycle with an integrated view covering both the technical side of the assets and its economic dimension in order to develop methods and tools to support risk-informed decision making. The focus is then set here on the top lines of the EAM challenges table given by Amadi-Echendu et al. [1], that is to say "Tactics: Asset maintenance", "Strategy: Asset life cycle analysis" and "Decision and models".

The generic framework that has been developed by EDF R&D is then to build a model that describes both:

- Technical system: model describing the relationship between all physical components. This covers the reliability models of all physical assets, the global asset architecture describing the relationships between subsystems, the maintenance programs, the supply-chain… The model complexity will depend on the scope of the analysis, from a model representing a single elementary asset with a single failure mode up to a complete plant with system dynamic, redundancies…
- Economic system: model describing how the technical assets impact the economic dimension. Each technical event may generate cash-flows (positive or negative) and these must be valued in order to evaluate a life-cycle strategy. For this model the complexity will also depend on the scope of the study, from a single cost generated by a failure up to complex economic behaviour such as insurance covering part of generation losses or tax-shield issues related to spare parts.

2.2 Net Present Value

In finance, the Net Present Value (NPV) is an indicator used to quantify the profitability of an investment or a set of investments. It is defined as the sum of all cash flows associated with the investment, whether it is a positive cash flow (income) or a negative one (outcome). The cash flows are present values, that is to say their value is discounted according to a discount rate representing the time value of money. For a single investment I generating annual cash flows CF_i over N periods with a discount rate α, the NPV is calculated through Eq. 1:

$$NPV = -I + \sum_{i=1}^{N} \frac{CF_i}{(1+\alpha)^i} \quad (1)$$

Decision making, using this indicator is then straightforward: if the NPV is positive the investment is profitable. This definition of the NPV is a very simplistic one, but may be more complex in some cases (multiple investments, tax shield impact…) even though the philosophy remains the same (for a more detailed discussion on NPV, one could read [3]). One of the main drawback of using such a simple expression is to consider the cash flows to be deterministic ones; as a matter of fact, future cash-flows, especially long-term ones, are usually uncertain and can't be given as a single value. This is the reason why it is important to calculate not only the average value of the NPV, but its probabilistic distribution in order to support efficiently the decision process.

2.3 Life Cycle Cost

The value of an engineering asset may not be summarized with the sole purchase cost of this asset and the issue of quantifying the real cost of an asset has been studied for several years before the first definition of Life Cycle Cost has been given by the American Department of Defense (DoD) in the late seventies in order to improve the budget control for military assets [8]. The idea of Life Cycle Cost is to consider all costs associated to a given asset, from the very first development to its decommissioning through all operational costs such as maintenance costs, unavailability costs… The LCC is often referred as a cradle to grave cost or global owning cost. What decision makers would like to do is to minimize this LCC; the difficulty is that the various costs making the LCC are related and sometimes antagonists:

- Preventive maintenance tasks are expensive but may reduce the number of failures,
- Purchasing spare parts will improve the efficiency of maintenance task and reduce the outage durations,
- Upgrading an asset may improve its reliability…

To evaluate this LCC, an engineering asset manager needs to build the model of relations between all these costs and to provide input data (reliability, unavailability consequence, spare part lead time….) to this global cost model.

The LCC indicator is different from the NPV, even if the costs of the LCC may be discounted to take into account the time value of money and then be homogenous with the NPV. In engineering asset management the NPV of a strategy should be defined as the difference in the LCC induced by this strategy compared with a reference (or current) one. Figure 1 illustrates this definition of the NPV with investments (preventive replacement in orange and spare purchase in grey) reducing the possible

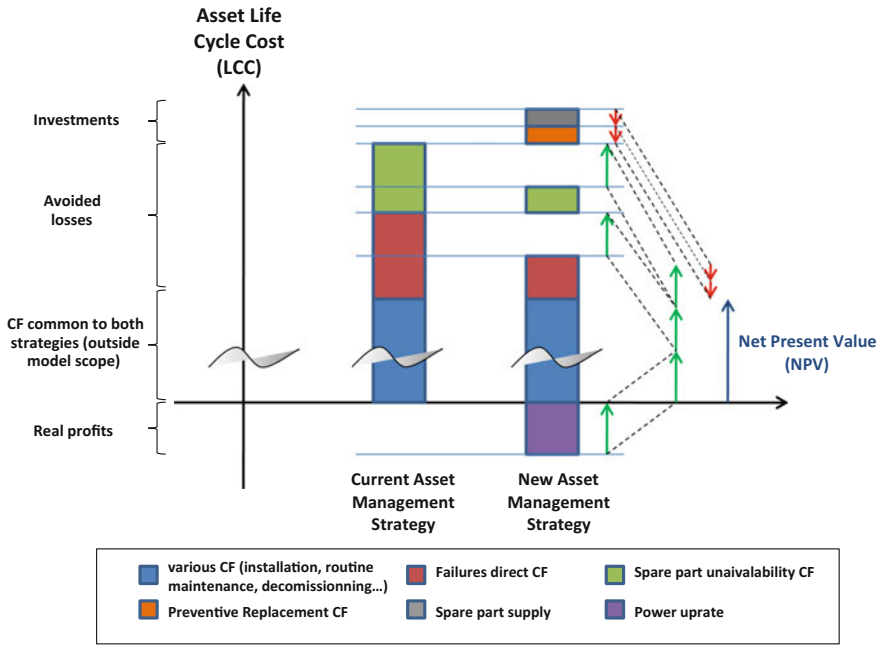


Fig. 1 NPV as the difference between LCCs

corrective costs (failures in red and spares shortage in green) and improving the production of the asset (power uprate in purple). When comparing these cash flows with the ones of the current strategy we can sum relative outcomes (red arrows) and incomes (green arrows) to evaluate the Net Present Value of these investments.

In the field of maintenance or spares investments, the profits generated by an investment are then more likely to be avoided losses. The above definition is presented as a deterministic one, but as the failures dates of assets are usually probabilistic ones, the LCC and thus the NPV are also probabilistic variables.

Once the global system is created we need to be able to compute values for a given life-cycle strategy or to optimize this strategy. There are two main types of methods:

1. **Numerical computation**: This type of methods, that has been described in detail in [4] relies on a model representation using stochastic process, usually Markov related such as Markov process, Semi-Markov process, Piecewise Deterministic Markov Process… Once the system has been modelled with these processes, numerical methods are used to calculate approximations of the strategy values (finite elements method, finite difference method…). These methods are usually used to compute expected values [5] or some simple statistics (single quantile or moments) but are not efficient to evaluate high level risk representation, like the Cumulative Distribution Function, because of computing time and/or memory issues.

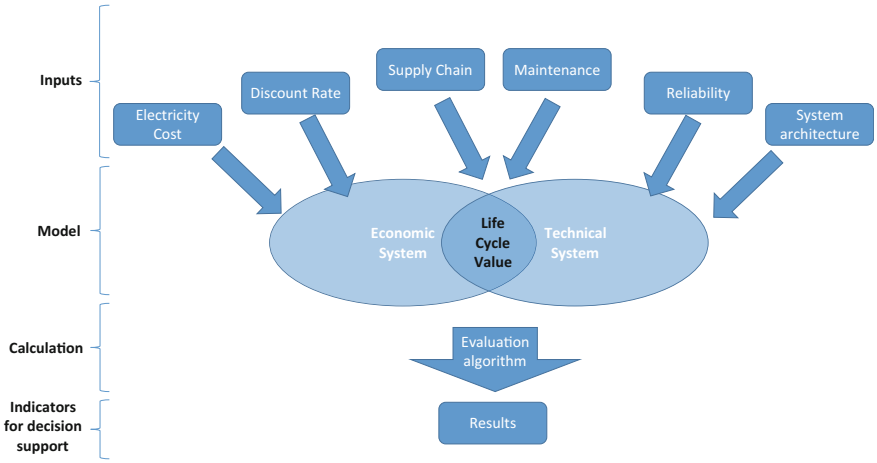


Fig. 2 Generic EAM framework

2. **Monte-Carlo simulation**: An alternative method is to represent the global system as an event model that is to say the occurrence dates of events, which can be deterministic or probabilistic, as well as a description of the interface between all potential events defining how the model updates after each event. Once the system is modelled, Monte-Carlo simulation is used to replicate the random aspect of the model and calculate risk-informed evaluations tending to their real values with the number of replications going up [9].

The generic framework used by EDF R&D in all its developments related to EAM is synthetized in Fig. 2. The goal of this framework is to define how to build and evaluate models for complex technical and economic systems of assets. The complexity of the systems is not only related to technical aspects (redundancies or complex supply chain) but also on the economic part of it (costs mutualization, insurance…). Interleaving economic and technical data is the key to have a realistic model of the assets and therefore to support efficiently decision making.

3 Specificities for Different Power Sources

The generic framework described above has been implemented and adapted to various power sources. As it will be explained, each generation type will have its own specificities that necessitates the adaptation of the generic methods, these characteristics may involve the type of fuel (renewable or not), the size of the fleets, the complexity of studied systems or the complexity of the economic context that can generate differences between a base-load large fleet of plants and a single plant piloted in real-time to cover the power demand.

3.1 Nuclear

Nuclear power in France is characterized by class effects, that is to say large number of identical plants. Currently EDF operates 58 plants belonging to 3 classes and 6 subclasses. This class effect enables to develop large synergies between the plants that are a source of significant savings. The down-side of this is that investments related to the ages of component will tend to be necessary on all plants in a very narrow time window with budget issues. This is the reason why the need is here to prioritize and optimize investments plans in order to smooth the budgets within safety and technical constraints.

This is the reason why the tools developed for the nuclear generation are based on simple and fast to compute models (semi-Markov graphs computed through finite differences methods) linked to an optimization algorithm (Genetic Algorithm) that enables decision makers to help them find the best investments planning that will fulfil all constraints. The risk aspect of life-cycle strategies is not used as an optimization goal but is checked a posteriori with a simplified Monte-Carlo simulation. The tool, named IPOP for Investments Portfolio Optimal Planning [7], has been used for several real cases and is part of EPRI Integrated Life-Cycle Management suite dedicated to long term operation. One of its applications was described in showing how the tool was used for real decision making regarding the investments program for the transformers of an American utility (see Fig. 3).

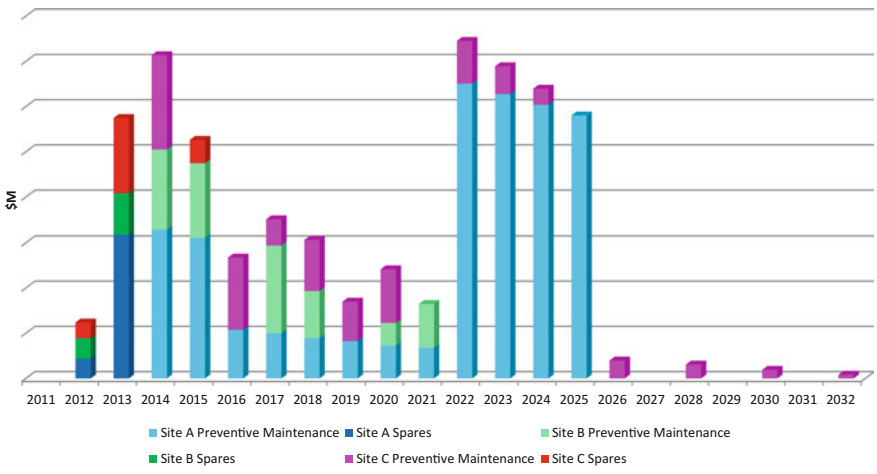


Fig. 3 Example of optimized investments planning for a family of transformers

3.2 Hydropower

The main specificity of hydropower generation comes from the power fuel itself, the water, which is probabilistic and depends on the weather, but also creates a dependency between plants. A failure of an asset in one plant may have consequences in downstream plants, if for example it stops the water and therefore not only generates a loss of production on the plant it is installed on but on all downstream plants. Another type of dependency between plants would be the opposite situation in which a failure would prevent the water regulation, potentially impacting the behaviour of downstream installations. The main difficulty is then to create a model to capture these dependencies, one solution could be to model an entire valley describing the interactions between plants as a dynamic system in which probabilistic events will be the failures of components as well as the downfall. An approximated model could focus on a single plant over-estimating the generation losses to take into account the potential downstream consequences. Research is on-going on this subject with a new project that started in 2017 with the final ambition to transfer operational tools to the engineering division. Another characteristic of the hydraulic area is the wide variety of installations (reservoir generating station, run-of-river generating station…) which involves different water management. This characteristic has to be taking account during the tool development.

3.3 Fossil

On the opposite of nuclear, fossil generation is made of very few plants with very different technologies, this is the reason why the risk-informed evaluation of investments is extremely important as risks won't be smoothed into a Gaussian-like profile but will rather show a low probability/high consequence residual risk. This is why EDF R&D has developed a tool, named VME (for valuation of exceptional maintenance), that enables the modelling of both complex technical system and complex economic system. If these kind of models are not reserved to fossil generations and were used for other power sources they are particularly suited to single investments regarding very few assets. An example of such an analysis was given in [6] describing how two different companies could optimize the spare parts global owning cost by sharing some. In this case both the technical and the economic systems were complex with a complete modelling of the supply chain taking into account the distance of the warehouse from each plant, the emergency supply with a shorter delay or the insurance of one company that partially covers generation losses. The results of this study (Figs. 4 and 5) were used for operational decision making with some spares currently shared by the two companies.

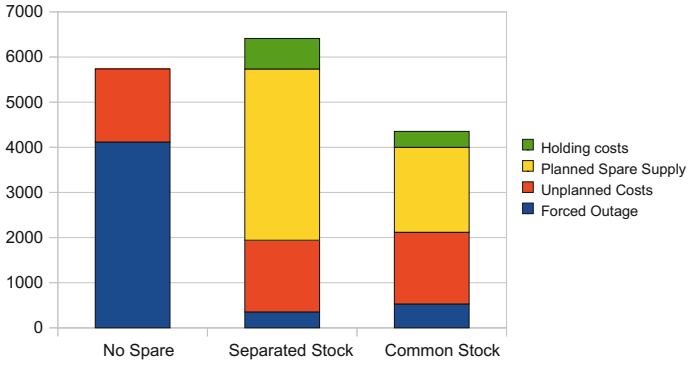


Fig. 4 Final cumulated cash flows for Company 1

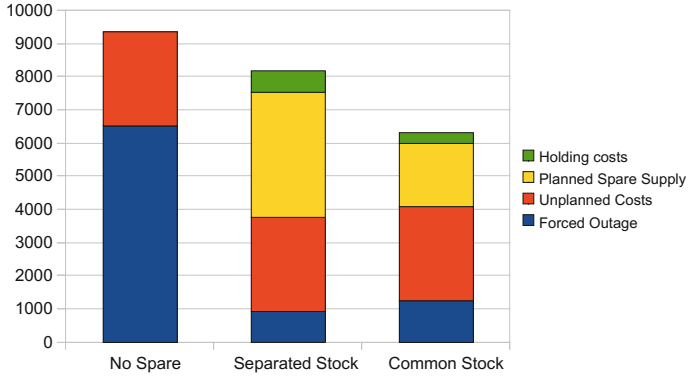


Fig. 5 Final cumulated cash flows for Company 2

3.4 Onshore Wind-Turbine

Nuclear and fossil energy are not influenced by the weather conditions, as hydraulic power depends on the water resource. But water may be stored in dams. Wind energy is another category of energy as its capacity is highly influenced by the local weather, and the wind cannot be stored. In order to provide realistic valuation of wind farms investments, it is therefore necessary to take into account a probabilistic model describing the wind conditions along the year. The impact of wind is subtle as high winds generate more energy by increasing the turbine load but on the other side high loads imply higher aging of operating components. The relationship between the wind characteristics (speed, turbulences…) and the component aging is not well-known for the moment. But a relevant model should take this context into account. Moreover for farms experiencing severe wind conditions during given seasons, maintenance operations may be impossible to perform (for example maintenances requiring a crane). A failure during this period may lead to long

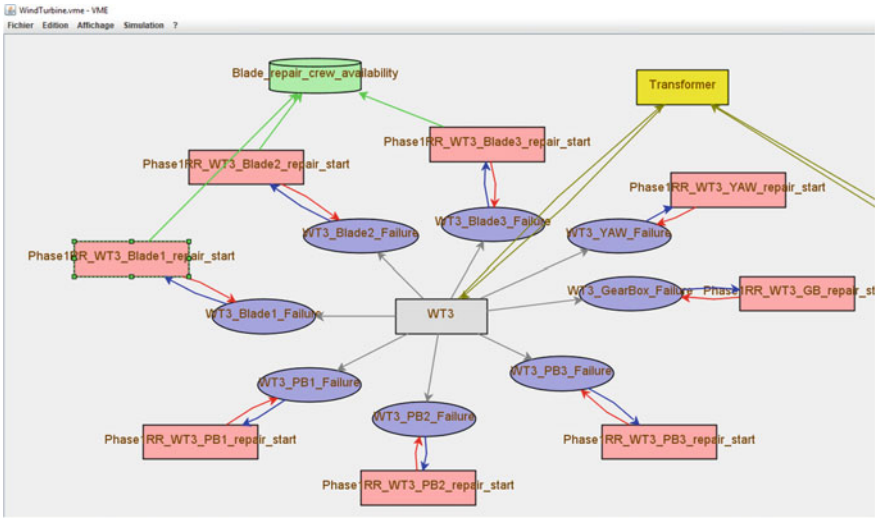


Fig. 6 Example of a wind-turbines farm model

unplanned unavailability. This analysis shows that an accurate asset management model for onshore wind farms has to consider the probabilistic behaviour of the wind and to define the impact the wind on the turbine components (generation, aging…). In this context, a relevant risk-informed decision making model has to combine two probabilistic modules: the weather module and the component reliability module. The biggest challenge is to define how these two modules interact with each other that is to say how the wind influence the reliability of the turbine components (Fig. 6).

3.5 Offshore Wind-Turbine

Offshore wind power share with onshore the same type of weather dependence. But the impact of weather is slightly different in this case. In one hand, wind regimes in the middle of the sea is often more regular than on the ground, allowing a better adjustment of turbines characteristics and hence, higher yields. Moreover, because of less restrictive constraints on landscapes and environmental impact, offshore wind turbines (OWT) are often higher and larger than onshore ones, improving production efficiency and making possible larger farms in terms of installed capacity. In the other hand, the marine environment induces harsher environmental conditions and thereby higher costs in terms of installation, inspection and maintenance operations. Indeed, the accessibility of the site for these operations is not only limited by the wind speed but also by the wave high and period, and potentially the current speed and direction. In addition, the intervention means required for offshore operations are much heavier than those need for onshore (jack-up vessels, shift work, etc.). For these reasons, logistical parameters are critical and

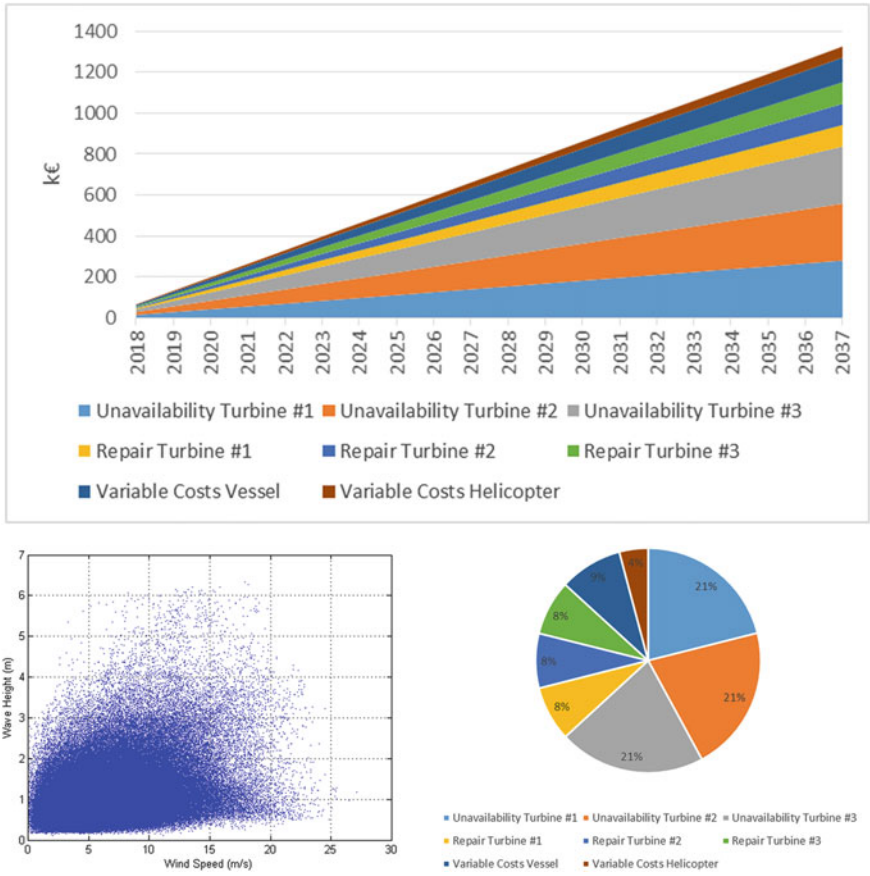


Fig. 7 Illustration of the different outputs provided by the ECUME O&M software

O&M costs are estimated between 15 and 30% of the overall production costs of power as it is only around 2% for onshore, reinforcing the importance of a dedicated asset management approach. To this aim, the ECUME softwares have been developed by EDF R&D [2]. These tools embed a weather-scenarios random generator, coupled with a logistic module, allowing Monte-Carlo simulations of offshore operations (Fig. 7).

4 Conclusion

This paper presents an overview of past and future development made by EDF R&D to support decision making regarding the engineering asset management of its plants. The generic framework, which has been developed for over fifteen years,

Table 1 Summary of model specificities for each type of generation

	Nuclear	Hydro.	Fossil	Onshore WT	Offshore WT
Supply chain	X	X	X	X	X
Assets reliability	X	X	X	X	X
Economic system	X	X	X	X	X
Complex availability model		X			
External covariates				X	X
Assets accessibility					X
Large number of assets	X				

has already been applied successfully to several power sources (nuclear, wind-turbines…) and adaptations are underway for of types of generation (fossil, hydraulic) to take into account all their respective specifities as they were described previously (see Table 1). If the three first lines of the table are common to all sources of power, as they are the base for any physical asset model (Supply chain, Reliability and Economic), the other lines summarize what differentiate the various generation type regarding asset management.

Usually Life-Cycle management relies on the optimization of the investments program. If the tools based on numerical evaluation of expected already have such optimization features, it is not that straightforward for tools based on Monte-Carlo simulation. For this reason future research will be focused on simulation-optimization issues, and how it can be applied to the next generation of EAM tools.

References

1. Amadi-Echendu JE et al (2007) What is engineering asset management? Dans: definition, concepts and scope of engineering asset management. Springer, pp 3–16
2. Douard F et al (2014) Probabilistic model of an offshore wind farm: from installation to operation. 19e Congrès de Maîtrise des Risques et Sûreté de Fonctionnement
3. Jones TW, Smith DJ (1982) An historical perspective of net present value and equivalent annual cost. Acc Historians J 9
4. Lair W, Ziani R, Mercier S, Roussignol M (2010) Modeling and quantification of aging systems for maintenance optimization. Reliability and maintainability symposium (RAMS)
5. Lonchampt J (2005) Méthode matricielle de quantification des risques associés à la maintenance d'un composant par un modèle semimarkovien. 6ème Congrès international pluridisciplinaire, Qualité et Sûreté de Fonctionnement, Qualita
6. Lonchampt J, Fessart K (2011) A complete probabilistic spare parts stock model under uncertainty. In: Advances in safety, reliability and risk management: ESREL 2011
7. Lonchampt J, Fessart K (2012) Investments portfolio optimal planning for industrial assets management: method and tool. In: PSAM11
8. Woodward DG (1997) Life cycle costing-theory, information acquisition and application. Int J Proj Manage 15(6):335–344
9. Zio E (2013) The Monte Carlo simulation method for system reliability and risk analysis. Springer, London

Reciprocating Compressor Valve Leakage Detection Under Varying Load Conditions

Panagiotis Loukopoulos, George Zolkiewski, Ian Bennett, Suresh Sampath, Pericles Pilidis, Fang Duan and David Mba

Abstract Reciprocating compressors are vital components in oil and gas industry though their maintenance cost can be high. Their valves are considered the most frequent failing part accounting for almost half the maintenance cost. Condition Based Maintenance which is based on diagnostics principles can assist towards decreasing cost and downtime while increasing safety and availability. Most common features utilised by reciprocating compressor diagnostics solutions are raw sensor vibration and pressure data. In this work non-uniformly sampled temperature-only measurements from an operational industrial reciprocating compressor, in contrast to experimental data commonly used, where utilised for valve leakage detection by employing principal components analysis and statistical process control. Analysis was done by using temperature ratios in order to filter out external effects like ambient temperature or speed. Result validate the success of the proposed methodology.

P. Loukopoulos (✉) · S. Sampath · P. Pilidis
School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, UK
e-mail: p.loukopoulos@cranfield.ac.uk

S. Sampath
e-mail: s.sampath@cranfield.ac.uk

P. Pilidis
e-mail: p.pilidis@cranfield.ac.uk

G. Zolkiewski · I. Bennett
Shell Global Solutions, Rijswijk, Netherlands
e-mail: george.zolkiewski@hotmail.com

I. Bennett
e-mail: Ian.Bennett@shell.com

F. Duan
School of Engineering, London South Bank University, London, UK
e-mail: duanf@lsbu.ac.uk

D. Mba
Faculty of Technology, De Montfort University, Leicester, UK

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_40

1 Introduction

Reciprocating compressors are essential components in oil and gas industry, requiring high reliability and availability [2]. Despite their popularity, their maintenance cost can be several times higher than that of other compressor types [17]. Bloch [2] notes that valves are the most common failing part (36%), making them the weakest component, accounting for half the maintenance cost [4].

In order to decrease downtime and cost, while increasing availability and safety, efficient maintenance is essential [2] since reciprocating failures can cause from production loss to human casualties [17]. Condition Based Maintenance (CBM) [6, 18], a policy founded on the diagnostics principle [6, 18] and has been increasingly popular over the years advocating that maintenance should be made when actually needed based on equipment's health status, is a step closer to this direction. Diagnostics has been an established area where several works have been created proposing various methodologies in order to improve CBM effectiveness, and can be categorised based on the information type used.

Qin et al. [16] used basis pursuit for de-noising. Wave matching was proposed for feature extraction where parameters of system's physical representation, optimized with differential evolution algorithm, were used as features, Support Vector Machine (SVM) was employed to determine health status of the compressor. Pichler et al. [15] transformed the data to time-frequency domain, computed their spectrograms and filtered them. Similarity of reference and test spectrograms was estimated. A threshold was created using one dimension logistic classification rule in order to determine the valve health state. This method was affected by varying load. In [10] they used the same transformation and created difference matrices between reference and test spectrograms, with two-dimensional autocorrelation filtering noise an highlighting any patterns. Distance of difference matrix and shape of autocorrelation served as features. Logistic Regression and SVM were compared regarding their classification capabilities. When both features were used accuracy of both methods increased. It was noted that this methodology can be applicable for varying load conditions. In [14] they used the proposed methodology with filtering applied twice for enhanced effect and extracting further features from autocorrelation of each filtered matrix. Nine types of valves were used along with various measuring configurations. Classifiers were compared considering features and valves used for training and testing. Ahmed et al. [1] extracted statistical features from time domain and used Principal Components Analysis (PCA) for dimensionality reduction. Statistical Process Control (SPC) was applied for fault detection on a reciprocating compressor. All works used experimental raw sensor vibration information.

Elhaj et al. [3] used visual analysis of time domain dynamic pressure and shaft instantaneous angular speed to detect suction and discharge valve leakage. They showed that either analysing each variable separately or combining their results, can yield good diagnostics performance. Pichler et al. [11] used gradient of expansion stroke from logarithmic scaled PV diagram and difference of suction and discharge

pressures of a cycle as features. A SVM was applied to determine health state of valve. They enhanced classification in [12] by linearising the classification boundary in order for failure data of valve types to be comparable. They compared their methodology in [13] with K-nearest neighbours, K-means, Gaussian data description, Gaussian mixture model along with a proposed classifier which can be used only for linearised classification boundary. When using data from all valve types for training and prior linearisation of classification boundary, SVM was superior, while after linearization proposed classifier was outperforming. When using data from a single valve type for training, after linearisation, proposed method was superior. These works employed experimental raw sensor pressure data.

Combination of several measurement types has also been the basis of some works. Guerra and Kolodziej [5] used temperature, pressure and vibration data, pre-processed them with a hamming window, transformed them to frequency domain, grouped them in frequency bins as features, and reduced them via PCA. Bayesian classification was applied to identify the status of the valve for varying load. Tran et al. [17] used vibration, pressure and current measurements, applied Teager-Kaiser energy operator to calculate envelope amplitude from vibration signals and wavelet transformation for de-noising the rest, extracted statistical features from time, frequency, and autoregressive domains, and reduced them with generalised discriminant analysis. Deep Belief Neural Network was used to identify machine's health state, outperforming a relevance vector machine and a back propagation neural network. Both works utilised experimental raw sensor information.

In the current work, non-uniformly sampled temperature measurements from an operational industrial, two-stage, four-cylinder, double-acting reciprocating compressor extracted from a server were used, instead of experimental information used in previous works. In order to compensate for the effect of external sources like speed, the ratios between Head End (HE) and Crank End (CE) valve temperatures were calculated and utilised to develop a PCA model, while SPC was employed to detect a valve leakage caused by a broken piece from the valve plate.

The rest of the paper is organised as follows. Section 2 analyses PCA and SPC fault detection procedure. Section 3 describes data acquisition procedure and temperature ratio calculation process. Section 4 presents results followed by a discussion. Section 5 contains concluding remarks.

2 Fault Detection Process

In this work PCA and SPC were used to perform valve leakage detection utilising temperature-only information.

2.1 Principal Components Analysis

PCA is a dimensionality reduction technique that projects a number of correlated variables in a lower space via a linear transformation (components), while preserving maximum possible variance within original set, creating a new group of uncorrelated, and orthogonal latent variables (scores) [7]. Let X be a $n \times p$ data matrix, its PCA transformation is given as:

$$X = TP' + R \tag{1}$$

where T is the $n \times k$ score matrix, the projection of X on the k-dimensional space with $k < p$. P is the $p \times k$ component matrix, the linear mapping of X to T. R is the $n \times p$ residuals matrix. Calculation of principal components (P) was done with use of singular value decomposition [7]. Selection of appropriate number (p) of components was done employing Cumulative Percentage of Variance (CPV) [7]. The information was centred and scaled to unit variance.

2.2 Statistical Process Control

SPC is used to monitor a univariate process for diagnostics purposes, which is considered to be healthy when its value lies within statistical limits decided by control chart used. For multivariate process, PCA is used to facilitate multivariate process control by reducing number of monitored variables and decorrelating them. After PCA model has been created, its scores and residuals can be used for SPC. Control charts employed in this work are Hotelling T^2 and Q residuals, most widely used ones regarding PCA/SPC [9]. Hotelling metric for score matrix T is [9]:

$$T^2 = \sum_{i=1}^{k} \frac{t_i}{s_i^2} \tag{2}$$

With t_i ith principal component scores, s_i^2 its variance. Control limits for this metric are given by Kruger and Xie [9]:

$$T_a^2 = \frac{k(n^2-1)}{n(n-k)} F_a(k, n-k) \tag{3}$$

where $F_a(k, n-k)$ is the $(100-1)a\%$ upper critical point of the F distribution with k and $n-k$ numbers of freedom. Q metric for residual matrix R is [9]:

$$Q = \sum_{i=1}^{n} (x_i - \hat{x}_i)^2 \tag{4}$$

With $\hat{x}_i$ reconstructed values of x_i. Control limits for this metric are given by Kourti [8]:

$$Q_a = g x_{h,a}^2 \tag{5}$$

where $g = \frac{var(R)}{2mean(R)}$, $h = \frac{2(var(R))^2}{var(R)}$ and $x_{h,a}^2$ is the $(100-1)a\%$ upper critical point of the x^2 distribution with h numbers of freedom.

The process of performing fault detection utilising PCA and SPC can be divided into two phases. Phase I can be considered the offline training step where healthy historical information is gathered, centred and scaled, and used to create a healthy PCA model while calculating T^2 and Q limits describing the healthy state. Phase II is the online implementation where new data are collected, normalised, and projected on the healthy PCA model. Their score and residual values are calculated and their T^2 and Q metrics are estimated and compared to the statistical limits from Phase I. In case of violation, these samples are out of control indicating a fault occurring thus performing detection. It could be noted that application of this methodology is fairly simple and straightforward. Furthermore, with dimensionality reduction procedure, number of variables monitored can be reduced since only 2 metrics T^2 and Q need to be monitored.

3 Data Acquisition

3.1 Data Preparation

Information employed in this work came from an operational industrial two-stage, four-cylinder, double-acting reciprocating compressor that has been used in various applications. The machine is instrumented with sensors collecting both process (temperature, pressure, etc.) and mechanically (bearing vibration, bearing temperature, etc.) related measurements, that stream continuously, via internet, to a central location for e-maintenance diagnostics. The large volume of data created, considering each sensor's sampling frequency, requires a huge amount of storage. Hence, a ruleset was created deciding which values should be stored, leading to non-uniformly sampled sets. To compensate, linear interpolation is used.

The fault mode under study was a valve failure. A ring valve was the defective component with cause of failure: broken valve plate leading to leakage. There were 13 defective cases available that all took place in the same cylinder within a period of one and a half years. Depending on case, the failing valve was either Head End (HE) or Crank End (CE) discharge valve. In all failures, valves were of same type, model, and manufacturer. Historical information of 16 temperature measurements, one for each valve, was extracted from a server with sampling period $T_s = 1s$ ($f_s = 1$ Hz), describing a period of roughly two and a half days regarding each failures, containing both healthy and faulty states.

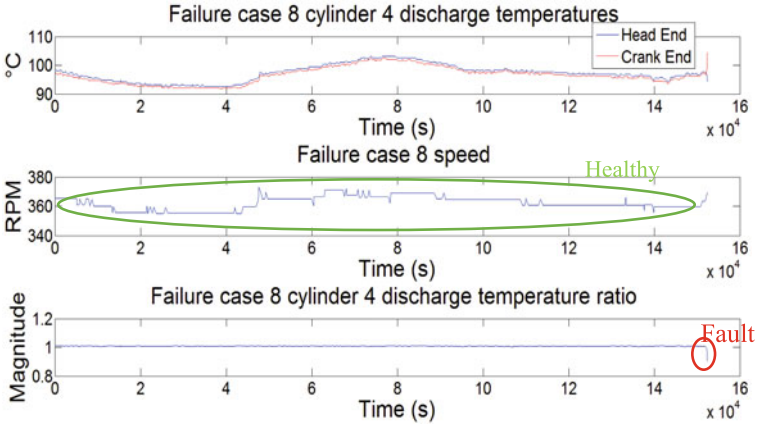


Fig. 1 Cylinder four discharge temperature ratio for failure case 8

3.2 Temperature Ratios

In order to mitigate the effect of external sources like speed variation or ambient temperature on the measurements, their ratios were employed instead. The ratios were calculated as follows: $r_i = HE_{kjtemp}/CE_{kjtemp}$, where $i = 1, \ldots, 8$ the number of ratios, j indicates whether the temperatures are for suction (1) or discharge (2), and $k = 1, \ldots, 4$ the number of cylinders. The temperature ratio of discharge temperatures in cylinder four for failure case eight can be found in Fig. 1. All three graphs comprising Fig. 1 describe both healthy (almost complete signal duration) and faulty (towards the end) states. The temperature values (top graph) vary over time, as does the speed (middle graph). In the lower graph it can be seen that by calculating the ratio the external effects can be filtered out as confirmed by its flat shape during the healthy period while the spike in the end corresponds to the faulty state thus providing a robust indicator.

4 Results

4.1 Phase I

From each failure case a significant amount of data describing the healthy state was extracted and merged into a single healthy set, that was centred and scaled to unit variance, and was used to create a PCA model with three principal components, characterising the healthy state. Then T^2 and Q limits were calculated, with a significance level of $a = 0.01$, discriminating the healthy and faulty states and can be found in Fig. 2. Analysis was done in Matlab.

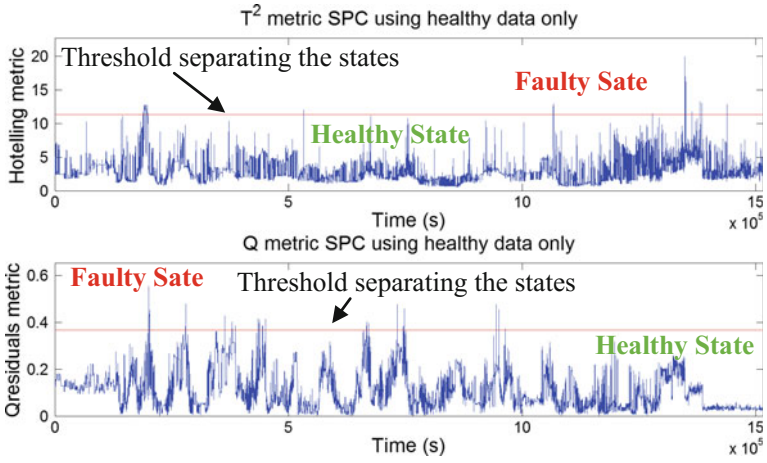


Fig. 2 T^2 and Q limits describing healthy state

From Fig. 2 it can be seen that there are some false alarms but they are to be expected since the data come from an operational compressor that had been used in various applications (compressing gases with diverse composition). False alarms can also be attributed to the statistical nature of this methodology since for $a = 0.01$, there are to be expected 1% of type I errors. In this analysis 1516000 samples were used thus, 15160 samples are expected to be out of control. T^2 false alarms were 8672 while Q ones were 7343 which both are way below the expected number.

4.2 Phase II

After creating the healthy PCA model and establishing the statistical limits describing the healthy state, the remaining data containing both healthy and faulty measurements were used for damage detection. Figure 3 contains SPC results for failure case one. The machine had been operating in the healthy region for almost all duration of the signal, with the fault appearing towards the end and being successfully detected by both metrics. The moment of detection could be considered as the moment a crack appears on the valve plate, with an increase of metrics' values reflecting the crack propagation process, and with the maximum value reached as the moment a piece of the plate breaks off indicating structural failure.

The successful performance of PCA/SPC fault detection for valve failure can be confirmed by also observing Fig. 4 where SPC results for case six are demonstrated. In this case as with the previous one, it can be seen that the machine had been healthy for a long period with the fault appearing towards the end and being detected by both metrics. Detection was performed successfully for all 13 cases,

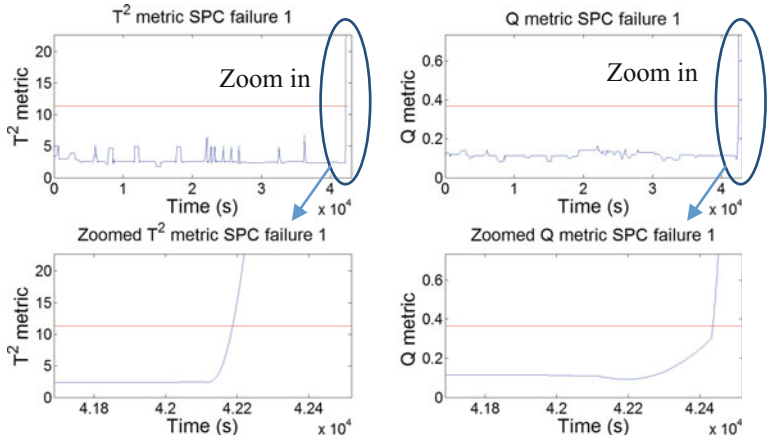


Fig. 3 Statistical process control results for failure case 1

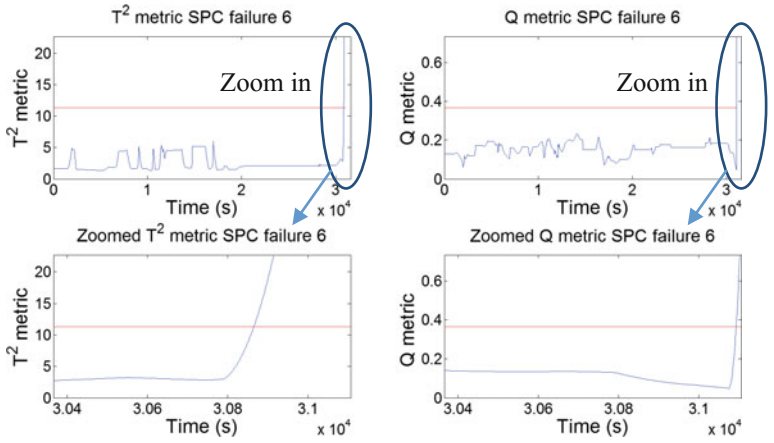


Fig. 4 Statistical process control results for failure case 6

confirming the validity of employing PCA/SPC for valve damage detection in collaboration with temperature only measurements in the form of the proposed ratio.

Going through the results, It can be observed that T^2 metric appears to be more sensitive as its statistical limit is violated prior to that of Q. This could be attributed to the fact that T^2 accounts for the variation within the system and the use of temperature ratios was able to remove successfully the effect of external sources thus reducing data variability. Consequently, T^2 was able to detect earlier any change attributed to an abnormality.

5 Conclusions

In this work, non-uniformly sampled temperature-only measurements were used in order to detect a leaking valve in an operational industrial reciprocating compressor, utilising principal components analysis and statistical process control. In order to mitigate the effect of external sources temperature ratios between head end and crank end valve temperatures were calculated and used for analysis. Results demonstrated that valve leakage was detected successfully, while use of ratio was able to significantly reduce number of false alarms, thus increasing detection efficiency.

Acknowledgements The authors would like to thank Amjad Chaudry and Jose Maria Gonzalez Martinez for their insight during analysis.

References

1. Ahmed M, Gu F, Ball AD (2012) Fault detection of reciprocating compressors using a model from principles component analysis of vibrations. J Phys Conf Ser 364. https://doi.org/10.1088/1742-6596/364/1/012133
2. Bloch HP (2006) A practical guide to compressor technology, 2nd edn. Wiley, Hoboken, New Jersey
3. Elhaj M, Almrabet M, Rgeai M, Ehtiwesh I (2010) A combined practical approach to condition monitoring of reciprocating compressors using IAS and dynamic pressure. World Acad Sci Eng Technol 63:186–192
4. Griffith WA, Flanagan EB (2001) Online continuous monitoring of mechanical condition and performance for critical reciprocating compressors. In: 30th turbomachinery symposium, Houston, Texas
5. Guerra CJ, Kolodziej JR (2014) A data-driven approach for condition monitoring of reciprocating compressor valves. J Eng Gas Turbines Power 136. https://doi.org/10.1115/1.4025944
6. Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech Syst Signal Process 20:1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012
7. Jolliffe IT (2002) Principal component analysis, 2nd edn. Springer series in statistics. Springer, New York. https://doi.org/10.1007/b98835
8. Kourti T (2005) Application of latent variable methods to process control and multivariate statistical process control in industry. Int J Adapt Control Signal Process 19:213–246. https://doi.org/10.1002/acs.859
9. Kruger U, Xie L (2012) Statistical monitoring of complex multivariate processes. Wiley, Chichester, UK. https://doi.org/10.1002/9780470517253
10. Pichler K, Lughofer E, Buchegger T, Klement EP, Huschenbett M (2012) Detecting cracks in reciprocating compressor valves using pattern recognition in frequency space. In: ASME 2012 conference on smart materials, adaptive structures and intelligent systems. ASME, pp 749–756. https://doi.org/10.1115/smasis2012-8052
11. Pichler K, Lughofer E, Pichler M, Buchegger T, Klement E, Huschenbett M (2013a). Detecting broken reciprocating compressor valves in the pV diagram. In: 2013 IEEE/ASME international conference on advanced intelligent mechatronics: mechatronics for human wellbeing. IEEE, pp 1625–1630. https://doi.org/10.1109/aim.2013.6584329

12. Pichler K, Lughofer E, Pichler M, Buchegger T, Klement EP, Huschenbett M (2013b) Detecting broken reciprocating compressor valves in pV diagrams of different valve types. In: ASME 2013 conference on smart materials, adaptive structures and intelligent systems. ASME. https://doi.org/10.1115/smasis2013-3009
13. Pichler K, Lughofer E, Pichler M, Buchegger T, Klement EP, Huschenbett M (2015) Detecting cracks in reciprocating compressor valves using pattern recognition in the pV diagram. Pattern Anal Appl 18:461–472. https://doi.org/10.1007/s10044-014-0431-5
14. Pichler K, Lughofer E, Pichler M, Buchegger T, Klement EP, Huschenbett M (2016) Fault detection in reciprocating compressor valves under varying load conditions. Mech Syst Signal Process 70–71:104–119. https://doi.org/10.1016/j.ymssp.2015.09.005
15. Pichler K, Schrems A, Buchegger T, Huschenbett M, Pichler M (2011) Fault detection in reciprocating compressor valves for steady-state load conditions. In: 11th IEEE international symposium on signal processing and information technology. IEEE, pp 224–229. https://doi.org/10.1109/isspit.2011.6151564
16. Qin Q, Jiang Z-N, Feng K, He W (2012) A novel scheme for fault detection of reciprocating compressor valves based on basis pursuit, wave matching and support vector machine. Measurement 45:897–908. https://doi.org/10.1016/j.measurement.2012.02.005
17. Tran VT, AlThobiani F, Ball A (2014) An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks. Expert Syst Appl 41:4113–4122. https://doi.org/10.1016/j.eswa.2013.12.026
18. Vachtsevanos G, Lewis F, Roemer M, Hess A, Wu B (2006) Intelligent fault diagnosis and prognosis for engineering systems. Wiley, Hoboken, NJ, USA. https://doi.org/10.1002/9780470117842

Acoustic Signature Based Early Fault Detection in Rolling Element Bearings

Amir Najafi Amin, Kris McKee, Ilyas Mazhar, Arne Bredin, Ben Mullins and Ian Howard

Abstract Early fault detection in rotary machines can reduce the maintenance cost and avoid unexpected failure in the production line. Vibration analysis can diagnose some of the common faults inside the rolling element bearings; however, the vibration measurement should be taken from a transducer that is located on the bearing or very close to the supporting structure, which is sometimes not feasible. This study compares acoustic and vibration signature-based methods for detecting faults inside the bearings. It uses both time and frequency based fault indicators (i.e. RMS, Kurtosis and envelope analysis) for investigating the condition of the system. Experiments were carried out on a belt-drive system with three different bearing conditions (normal, corroded and outer race fault). The experimental results show acoustic signature-based methods can detect the system's fault from close distances, and even for relatively far distance, some bearing conditions are still detectable.

A. Najafi Amin (✉) · K. McKee · I. Mazhar · A. Bredin · I. Howard
Department of Mechanical Engineering, Curtin University, Kent Street, Bentley, WA 6102, Australia
e-mail: amir.najafiamin@curtin.edu.au

K. McKee
e-mail: k.mckee@curtin.edu.au

I. Mazhar
e-mail: i.mazhar@curtin.edu.au

A. Bredin
e-mail: a.bredin@curtin.edu.au

I. Howard
e-mail: i.howard@curtin.edu.au

B. Mullins
Department of Health, Safety and Environment, Curtin University, Kent Street, Bentley, WA 6102, Australia
e-mail: b.mullins@curtin.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_41

1 Introduction

Fault detection in rotary machines is vital, and early fault detection can reduce the maintenance cost and avoid unexpected failure in the production line. This study focuses on fault detection in bearings. According to bearing manufacturers some of the faults such a ball fault requires immediate action, and some like inner race and outer race faults may give few weeks to few months till failure [11]; therefore it is crucial to monitor and detect bearing faults at the early stage.

Root-Mean-Square (RMS), Crest factor and Kurtosis values can be used for fault detection. Detecting faults through just using one of these values is not usually possible and combinations of them can be required. This also helps to find the stage of the fault in the system [2]. Principal Component Analysis (PCA) on historical data can show unusual variability and is useful for pre-processing of multi-fault diagnosis [7, 10]. Furthermore, neural networks and fault classification can be implemented to facilitate the fault detection [5, 18, 20]. Spectral Kurtosis (SK) is also a powerful tool which does not required historical data for detecting faults [15].

Localised faults in bearings generate periodic impulse forces each time the ball passes over the defect [6]. The frequencies of these impulses depend on the geometry of the bearing and the rotating speed. There are four common fault frequencies for a bearing. The Fundamental Train Frequency (FTF), which indicates the rotation rate of the cage or ball retainer. The Ball Spin Frequency (BSF), which shows the rotation rate of the balls or rollers. The Ball Pass Frequency Outer Race (BPFO) and the Ball Pass Frequency Inner Race (BPFI) show the rate at which a ball passes over a fault in outer or inner race. When the outer race is stationary, and the inner race is rotating with the shaft, these frequencies derive from:

$$FTF = \frac{1}{2}\left(1 - \frac{B_d}{P_d}\cos\theta\right)\omega, \tag{1}$$

$$BSF = \frac{P_d}{2B_d}\left(1 - \left(\frac{B_d}{P_d}\right)^2(\cos\theta)^2\right)\omega, \tag{2}$$

$$BPFO = \frac{n}{2}\left(1 - \frac{B_d}{P_d}\cos\theta\right)\omega, \tag{3}$$

$$BPFI = \frac{n}{2}\left(1 + \frac{B_d}{P_d}\cos\theta\right)\omega, \tag{4}$$

where n is number of bearing balls, B_d is ball diameter, P_d is bearing pitch diameter, θ is contact angle of the ball on the race and ω is the speed of shaft rotation [12]. For these mathematical models, it is assumed that there is no sliding between elements; however, by having sliding, the theoretical value for the inner race fault

will be slightly lower than the actual value and for the outer race fault will be higher [12].

Since there are a large number of frequency components in the raw frequency data, detecting fault frequencies from raw frequency components is extremely difficult [3]. The vibration response usually contains extraneous noise; however adaptive filtering can be used for cancelling some of these noises [18]. Wavelet transforms can also serve as a tool to filter the signal and extract the fault vibration signals and detect the fault size [17, 20]. The fault frequencies can also be extracted through the use of the envelope demodulation technique of high frequency resonance responses [8, 15, 17]. For this procedure first the proper frequency band that contains the impulsive components excited by the defect in the bearing should be detected [15, 17]. Although SK initially was used for detecting randomly occurring signals [3], it can also be used for detecting the proper frequency band for envelope analysis [15, 16, 19].

Stress waves (SW) and acoustic emissions (AE) signals are also useful for condition monitoring of rotary machines [9, 13, 14]. The AE sensors have also been found to be more effective than accelerometers for early fault detection [4]. Still due to attenuation of the signal, for getting rich data from the measurement, the transducer needs to be located on the bearing or very close to the supporting structure [1, 4], which is sometimes not feasible. This study uses acoustic data and tries to detect faults inside a bearing. This allows the use of non-contact measured acoustic data from reasonable distances from the source of the noise for the fault detection.

2 Experimental Setup

Experiments in this research were carried out on a belt-driven fin fan test rig at Curtin University. As Fig. 1 shows, the test rig consists of an electric motor which drives a fan's shaft using a V-belt. Two bearings support this shaft and the fan rotates along a horizontal axis. Three accelerometers measure the vibration of the motor, the driven bearing and the fan bearing. In addition, two high quality microphones at the same position measure acoustic noise from two opposite directions (driven side and fan side).

A National Instrument Compact DAQ system captures and digitizes data from the accelerometers and microphones, and transfers the data to MATLAB. During the test, data from all channels were digitized with 51,200 samples per second for each channel for a duration of thirty seconds.

The focus of this study was on detecting faults in bearings. For this purpose, different faults were separately induced to the bearing at the fan side. This bearing has pitch diameter of 34.5 mm and 8 balls. Each ball has 7.92 mm diameter and contact angle of 0 degree. Three different bearing conditions were investigated, healthy bearing, bearing with outer race fault and corroded bearing. For introducing the outer race fault, a hole was drilled in the outer race.

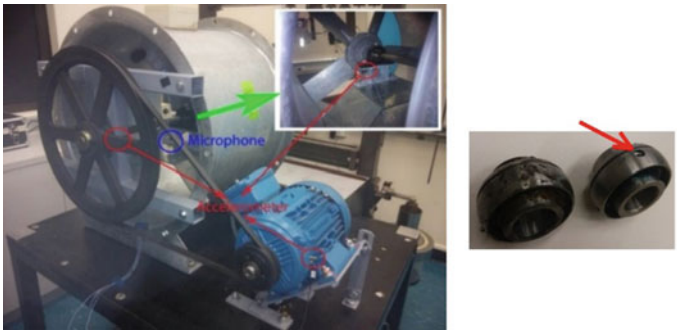


Fig. 1 Test rig (left) and faulty bearings (right)

3 Experimental Results

Fault detection for bearings usually focuses on localized faults which provide periodic impulses. This research investigated both localized and non-localized fault detection. First, the system with the healthy bearing was tested to find the base line for the RMS spectra and frequency spectra. In this test, the motor rotated at 13 Hz (780 rpm) and the fan's shaft rotated at 2.5 Hz.

For investigating the non-localized fault, the bearing at the fan side was replaced with an artificially corroded bearing. Figure 2 compares the RMS spectra of the healthy bearing and corroded bearing tests in 1/3rd Octave bands. This figure shows there to be distinguished amplification in both acceleration and acoustic RMS signature for frequencies between 200 and 1000 Hz.

For investigating the localized fault, the bearing at the fan side was replaced with a bearing with an outer race fault (Fig. 1). For this bearing at three different shaft speeds, the acceleration and acoustic data were captured. Equation 3 shows how the

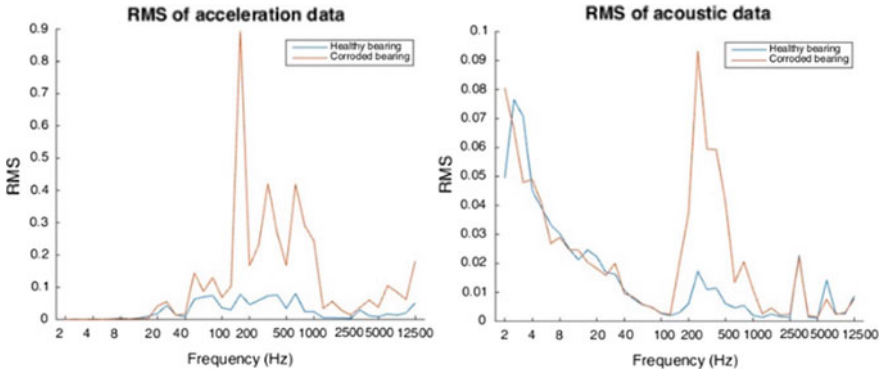


Fig. 2 RMS comparison of normal and corroded bearing

Table 1 Tests speeds

Speed of motor (Hz)	Speed of shaft (Hz)	Ball pass frequency of the outer race fault (Hz)
13	2.5	7.6
17	3.2	10.0
20	3.8	11.7

outer race fault frequency can be derived from the geometry of the bearing. Table 1 shows the speed of the shaft for these tests, and the expected fault frequencies.

Figure 3 shows the acoustic SK for a healthy bearing and bearing with outer race fault. As this figure shows, the acoustic frequency band of 200–500 Hz has high Kurtosis which has been chosen for envelope analysis filtration. Figure 4 shows the result of the acoustic envelope analysis for the three shaft speeds. In these figures, the related frequency of outer race fault for each shaft speed and its harmonics are clearly visible. These results show how the acoustic measurement at relatively close distance can be used for detecting localized bearing fault.

The acoustic data in these tests was captured with a microphone very close to the source of noise (near field); therefore, it could be assumed that the acoustic waves come directly from the source to the microphone. Furthermore, acoustic measurements were made for 8 different angles of orientation of the test rig, at distances of 2 and 4 m (Fig. 5), to investigate the effect of distance on acoustic fault detection. The RMS results for the corroded bearing for these measurements still showed distinguished amplification for frequencies between 200 and 1000 Hz; however, for the bearing with the outer race fault, the fault frequencies were not detectable any more. Since for these tests, the acoustic measurement was placed in the far field, the acoustic reflection inside the lab could easily interfere with the measurement. As a result, the periodic signals were faded out due to high noise levels.

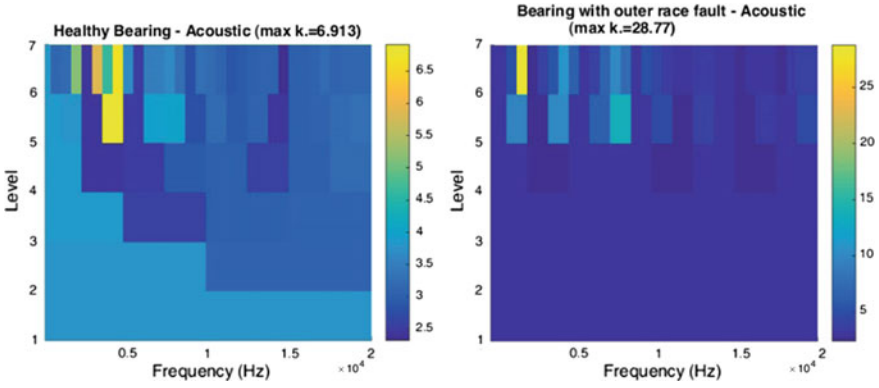


Fig. 3 Acoustic Spectral Kurtosis for a healthy bearing based on acoustic measurement

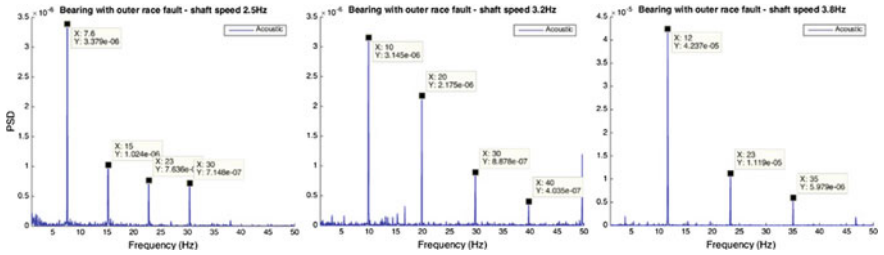


Fig. 4 Envelope analysis result for bearing with outer race

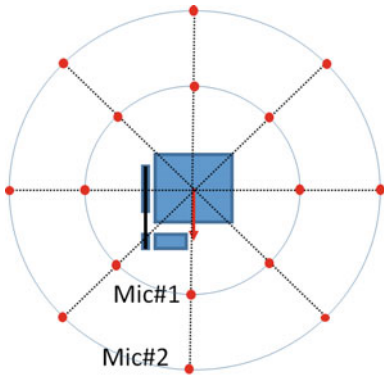


Fig. 5 Microphones positions in respect to the test rig at distance of 2 and 4 m

4 Discussion

For fault detection and condition monitoring of rotary machines, acceleration, SW and AE signals are commonly used [9, 13, 14]. This study uses the common fault detection techniques on acoustic measurement data and shows the measured sound from close distance can be used for detecting some localized and non-localized faults in bearings. In contrast to vibration measurement, where the transducer should be mounted on the bearing or on the supporting structure, for the acoustic measurement, the microphone can have more distance from the source of noise. In addition, the accelerometers must be rigidly mounted while the microphone must not be mounted and should be free of vibration input.

For extracting the fault frequency and its harmonics using envelope analysis, the filtering usually starts at several KHz [15, 18]; however, the result of this study shows with acoustic measurement, the whole frequency band for filtering can be below 1 kHz. Therefore, for digitizing and capturing data, lower sample rates could be sufficient and cheaper devices can be used.

This study investigates a bearing with outer race fault and shows that the fault frequency related to this condition and also its harmonics are easily visible with envelope analysis on both acceleration and acoustic data. Tests with different shaft

speeds show that the frequency band for envelope analysis for these speeds was not very sensitive to the shaft speed. Furthermore, tests on different fan blade pitch angles showed that there were no significant changes in RMS spectra or result of envelope analysis as well.

This study only investigated one type of localized fault in the bearing, and the rest of other known localized faults should be investigated; however, due to the common nature of these faults to generate harmonic impulses, similar results are expected to be achieved.

This study first used high quality microphones and showed that acoustic measurement at close distance can be used for detecting bearing faults. It then repeated the tests with low quality and cheap microphones. The comparison result between these two types of microphone showed that only for low frequencies (below 40 Hz), there was a major difference in result. Since the selected frequency band for RMS spectra and envelope analysis started at 200 Hz, these low-quality microphones were found to be very suitable for acoustic measurement and bearing fault detection.

5 Conclusion

This study investigated three bearing conditions and the following conclusions can be drawn from the current research:

- For a corroded bearing which has non-localized fault, the RMS spectra in 1/3rd Octave band showed major changes in both acceleration and acoustic signatures for frequencies between 200 and 1000 Hz. The result from acoustic measurement at both near and far distance to the source, confirmed this change in acoustic RMS signature.
- For a bearing with outer race fault, the SK showed high kurtosis values for frequencies between 200 and 500 Hz. Therefore, this frequency band was used for envelope analysis. For this bearing condition, the envelope analysis result on both acceleration and acoustic data at close distance showed the fault frequency and its harmonics. The tests on different shaft speeds and different fan blade pitch angles also confirmed that acoustic measurement at close distance can be used for detecting localized faults in bearings.
- For the far distance measurement, the fault frequency of the bearing with outer race fault was no longer visible due to interference of the acoustic reflection inside the room with the original signal.

Acknowledgements This work has been supported by Woodside Energy, CISCO and Curtin University.

References

1. Amini A, Entezami M, Papaelias M (2016) Onboard detection of railway axle bearing defects using envelope analysis of high frequency acoustic emission signals. Case Stud Nondestr Test Eval 6:8–16
2. Bhende AR, Awari GK, Untawale SP (2014) Comprehensive bearing condition monitoring algorithm for incipient fault detection using acoustic emission. J Tribol 2:1–30
3. Dwyer RF (1984) Use of the Kurtosis statistic in the frequency domain as an aid in detecting random signals. IEEE J Oceanic Eng 9
4. Gu DS, Choi BK (2011) Machinery faults detection using acoustic emission signal—from microdevices to helioseismology. INTECH Open Access Publisher
5. Hariharan V, Srinivasan P (2009) New approach of classification of rolling element bearing fault using artificial neural network. J Mech Eng ME 40:119–130
6. Hemmati F, Orfali W, Gadala MS (2016) Roller bearing acoustic signature extraction by wavelet packet transform, applications in fault detection and size estimation. Appl Acoust 104:101–118
7. Kamiel B, Forbes G, Entwistle R, Mazhar I, Howard I (2005) Impeller fault detection for a centrifugal pump using principal component analysis of time domain vibration features. Department of Mechanical Engineering, Universitiey Muhammadiyah Yogyakarta, Indonesia
8. Kamiel B, Mckee K, Entwistle R, Mazhar I, Howard I (2005) Multi fault diagnosis of the centrifugal pump using the wavelet transform and principal component analysis. In: Proceedings of the 9th IFToMM international conference on rotor dynamics. Springer, pp 555–566
9. Mba D, Bannister RH, Findlay GE (1999) Condition monitoring of low-speed rotating machinery using stress waves Part 1 & 2. Proc Inst Mech Eng Part E: J Process Mech Eng 213:153–185
10. Mckee KK, Forbes G, Mazhar I, Entwistle R, Hodkiewicz M, Howard I (2015) Cavitation sensitivity parameter analysis for centrifugal pumps based on spectral methods. Engineering Asset Management-Systems, Professional Practices and Certification. Springer
11. Quistgaard O (2012) Using selective envelope detection (SED) for the early detection of faults in rolling-element bearings. In: Kjær B (ed) BPT0046-EN-11. Brüel & Kjær
12. Roque AA, Silva TAN, Calado JMF, Dias JCQ (2009) An approach to fault diagnosis of rolling bearings. WSEAS Trans Syst Control 4:188–197
13. Sandoval HMU, Ramirez CAP, Méndez JEQ (2013) Acoustic emission-based early fault detection in tapered roller bearings. Ingeniería E Investigación 33:5–10
14. Sato I (1990) Rotating machinery diagnosis with acoustic emission techniques. Electr Eng Japan 110:115–127
15. Sawalhi N, Randall RB (2004) The application of spectral kurtosis to bearing diagnostics. In: Proceeding of ACOUSTICS, November 2004 Gold Coast, Australia, pp 393–398
16. Sawalhi N, Randall RB (2013) Localized fault detection and diagnosis in rolling element bearings: a collection of the state of art processing algorithms. In: 15th Australian international aerospace congress
17. Shi DF, Wang WJ, Qu LS (2004) Defect detection for bearings using envelope spectra of wavelet transform. J Vib Acoust 126:567–573
18. Wang H, Cheni P (2008) Fault diagnosis for a rolling bearing used in a reciprocating machine by adaptive filtering technique and fuzzy neural network. WSEAS Trans Syst Control 7:1–6
19. Wang Y, Xiang J, Markert R, Liang M (2016) Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: a review with applications. Mech Syst Signal Process 66–67:679–698
20. Yaqub M, Loparo KA (2016) An automated approach for bearing damage detection. J Vib Control 22:3253–3266

Localization of Bluetooth Smart Equipped Assets Based on Building Information Models

Mahtab Nezhadasl and Ian Howard

Abstract Indoor positioning systems are utilized to locate physical objects indoors without using GPS. Their applications include but are not limited to industry, business, and healthcare. This paper provides an analysis of a model and simulation of an indoor localization method which tracks physical assets relying on Bluetooth Smart. The system receives the desired building's floor plan and the materials of all walls and surfaces from the Building Information Model. The walls and surfaces have their own particular radio frequency (RF) absorption efficiency and transmission loss; when any propagated wave signal reaches a barrier, some of the signal will be reflected, some will be absorbed, and the rest will be transmitted through the barrier. This study implements the floor plan of a building and simulates the reflection and transmission of all signals (building's RF fingerprint map). To do so, the system generates a mesh for each Bluetooth reader, and calculates the level of received signal strength indicator (RSSI) for any points on the mesh. For each of these points, the simulation shows the propagation of RF signals in all directions and finds the summation of signals that may reach the reader to find the RSSI of that point.

1 Introduction

In the era of Internet of Things (IoT) and ubiquitous computing, it is crucial to know and access the current location information of devices and people and the ability to track them. The process of discovering the physical location of items is sometimes referred to as localization, positioning, location estimation, location

M. Nezhadasl (✉) · I. Howard
Curtin University, Bentley, WA, Australia
e-mail: mahtab.nezhadasl@postgrad.curtin.edu.au

I. Howard
e-mail: I.Howard@exchange.curtin.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_42

identification, or geolocation identification [20]. While the global positioning system (GPS) is employed for detecting the location of objects and individuals outdoors (where the satellites signals can be easily detected), the real-time location systems (RTLS) are capable of providing location estimation of targets indoors. The RTLS, also known as "indoor GPS", are electronic systems with the ever-increasing numbers of applications in industry, healthcare, retail, military, smart homes, and so on.

Over the past decade, the world has witnessed a variety of indoor localization methods from both industry and academia with the majority of them focusing on wireless technologies [10]. Such technologies include but are not limited to Wi-Fi, radio frequency identification (RFID), wireless local area network (WLAN), cellular phone networks, ZigBee, Ultra-wideband (UWB), Bluetooth (classic or smart), and hybrid [5, 16]. The utilization of RF-based technologies has two main advantages in comparison to the other indoor positioning systems: they cover larger area besides requiring less hardware. These are all because of their ability to penetrate through indoor obstacles such as furniture, walls and even human bodies [5]. To estimate the location of any desired target, RF-based RTLS solutions need to measure one or more properties of the received signal from an emitter to a receiving device. These signal properties are especially vulnerable to the surrounding environment and the distance of the travelled signal between an emitter and receiver [16]. RSSI is one of these properties that can be easily and inexpensively measured, and available in commercial off-the-shelf products. In addition, it has been used in several wireless sensor networks (WSN) to estimate the location of any radio device [15].

Fingerprinting, also known as Scene Analysis, is a commonly used localization algorithm with fair accuracy [16]. It performs a priori analysis of RSSI at every location inside the desired environment to develop a RSSI fingerprint database (radio map). Then it utilizes the developed map to find the closest match or matches for any device's RSSI to estimate its location [2, 8, 11, 12, 16, 18, 19]. Since any barrier like walls can affect the RSSI, it is worth considering the attenuation of the radio waves for getting accurate estimations. In fact, when any propagated wave signal reaches a barrier, some of the signal will be reflected, some will be absorbed, and the rest will be transmitted through the barrier [5]. In this paper, our simulation of the RSSI fingerprint-based indoor localization model has been presented that uses both the floor plan map of the environment and the ray tracing technique to improve the localization's accuracy by reducing the environmental effects.

The rest of the paper is organized as follows: Sect. 2 presents the work related to the current research. Section 3 gives the details of our model and simulation of an indoor localization method which tracks physical assets relying on Bluetooth Smart. Finally, Sects. 4 and 5 provide the discussion and concluding remarks respectively.

2 Related Work

RSSI is one of the signal properties that is easy to measure, and can be utilized in the form of client-based or network-based localizations [7]. In the client-based scheme, the object listens to emitters' signals and by knowing the position of them, it tries to locate itself. On the other hand, the network-based method is a reverse procedure and the readers listen to the emitted signal from the object so that they could locate the object [7]. It is worth mentioning that the first approach is only applicable for objects with computational capability, while the second one can be easily used for locating simple tags. Our simulation model has built on the network-based approach. Both of these methods need to calculate the distance by measuring the RSSI. The general equation for calculating the power of RSSI at distance of d from a receiver is:

$$P(d) = P(d_0) - 10 \times n \times \log_{10}\left(\frac{d}{d_0}\right), \tag{1}$$

where d_0 is the reference distance from sender which is usually 1 m, $P(d_0)$ is the power at distance of d_0 in dB, and n is the attenuation exponent [7]. For outdoor area without any barrier, n can be assumed to be 2. So the distance can be easily calculated by measuring the strength of the receiving signal. However, when there is a barrier between the sender and receiver, this calculation becomes challenging especially when the barrier is not stationary [7, 16]. The common way for dealing with this problem is to generate a fingerprint map for the signal strength, and the measured signal can be compared with the generated map to locate the object [12, 16, 19]. This map can be generated by actual physical measurement (site survey) at different locations or can be generated with the ray tracing simulation [7, 8, 19]. Either way this map changes by time due to the change of objects inside the room and the movement of people, and as a result, location detection becomes challenging [5, 10, 15].

Similar to our work, some studies [8] have employed both the floor plan and material of the surrounding environment to obtain indoor space localization. For instance, ARIADNE [7] could position targets by using a 2D floor plan, the knowledge of construction materials, and the ray tracing technique. It applied a clustering algorithm for searching inside its radio map.

In addition, new smart phones are equipment with several sensors such as magnetometer and accelerometer that enable calculating the movement of the device for indoor positioning system [6, 13, 17]. A study [19] shows that by using this data from these sensors, the fingerprinting method can be improved.

On the other hand, Bluetooth Smart technology, also known as Bluetooth Low Energy (BLE), is an upgrade to the standard/classic Bluetooth. It has been specifically designed for the IoT applications due to its small size, cheap price, and low energy consumption and implementation complexity. Also, it has the ability to be attached to physical assets in the form of small tags that makes it an ideal

candidate for the asset tracking purposes [1, 15]. Several studies [1, 3, 4, 9, 14, 15] have applied the BLE technology in indoor environments for the localization purposes. However, none of them utilized the ray tracing technique combined with the RF fingerprinting method. To address this issue, we have proposed a simulation model.

3 Modelling the BLE Positioning System

This study implements a C# code for simulating a RTLS, based on RSSI fingerprinting which tracks physical assets relying on Bluetooth Smart. It tries to locate any BLE transmitter by eight static BLE receivers. For generating the fingerprint map, the code uses the information of the walls and readers. For the readers, the position in a 2D plane should be defined. This code also sees each wall as a straight line in a 2D plane with its start and end point coordinates. In addition, some other characteristics need to be considered for the walls including their thickness, absorption coefficient, and transmission loss.

For building a fingerprint map, the program first generates an empty table for the entire defined area of interest based on the desired tolerance for each reader. In every cell of the table, it calculates the estimated RSSI for the reader for that point by using the ray tracing technique. To do so, it calculates and adds the power of all possible rays from that point to the reader. When a ray emits from a sender, it could encounter the reader directly or indirectly via reflecting or passing through obstacles. For the rays that reach the reader directly, Eq. 1 can be used to calculate the estimated RSSI by knowing the distance between the reader and desired point. For the other rays, the code calculates the first wall that the ray encounters and finds the RSSI at the point of contact. By using the absorption coefficient and transmission loss of the wall, it calculates the percentage of the power for the reflected ray and also for the ray passing through the wall. For each of these two new rays, it repeats this procedure until the power of the ray drops to a certain value or the ray reaches the reader. The RSSI value of each cell in the fingerprint map is the summation of all rays from that point that has reached to the reader. Figure 1 shows this procedure for generating the fingerprint map.

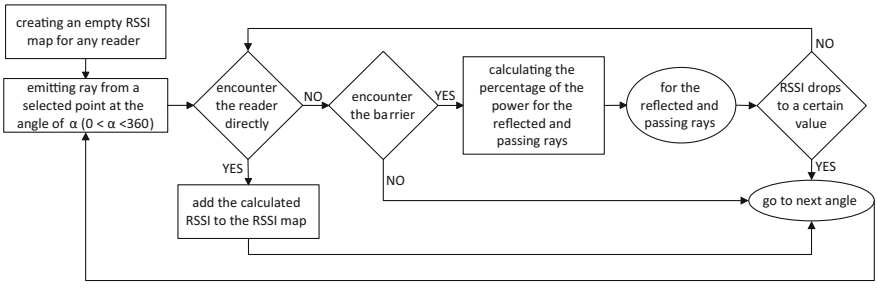


Fig. 1 The block diagram for generating the RSSI map

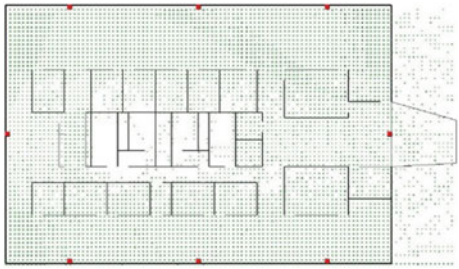


Fig. 2 The visible points to the readers

Figure 2 shows the overall coverage area of the eight stationary readers in this simulation. The figure illustrates all the eight static readers in the form of solid red squares, and the green dots as the visibility of points to the readers. The size of these dots depend on the number of the readers that can see them; the bigger the size of a dot, the more readers involved.

By having these fingerprint maps for each point on the map and for each reader, the calculated RSSI would be known. In the simulation mode, the user can select any point on the floorplan and based on these known RSSIs, the code estimates the possible points for the object. For this purpose, it finds all the points on the map of each reader that have similar RSSI of the selected point. Then the points that have been confirmed by higher number of readers, are assumed to be the possible location of the object. The simulation shows these points and the readers that see them on the display screen.

In addition, the code uses Eq. 1 and by assuming that there were no barriers or walls, it estimates the possible locations based on the RSSI of the selected point. The result of this situation can be compared to the result from the ray tracing and fingerprint maps. Figure 3 demonstrates an example where a point is visible to at least three readers. The red squares are the readers that could see the point (three red squares), the empty black square shows the selected point (clicked area on the screen which presents the location of a tag), blue circles present the estimated result by the ray tracing fingerprint, and finally the red circles exhibit the result of ignoring the attention and reflection effects of the walls. As this figure displays, since there is no barrier between the selected point and the readers, both methods show accurate results (the blue and red circles overlap).

On the other hand, Fig. 4 shows an example of one point that is behind walls. As this result shows, the fingerprinting map technique can easily estimate the location of the point of interest (the two blue circles), while by ignoring the effect of the walls, the result is no longer valid (the red circles in the centre of the floorplan). Again, the empty black square shows the selected point which is the clicked area on the screen presenting the location of a tag.

Figure 5 shows the result for the point that is only visible to one reader based on the ray tracing simulation. As this figure displays, the exact location of the point of interest is no longer known, and also by neglecting the effect of walls, the result shows higher error.

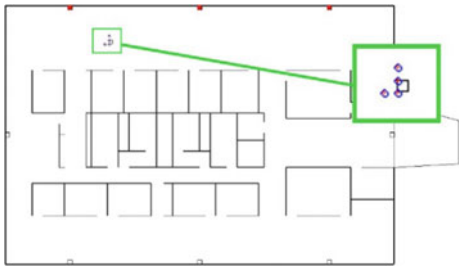


Fig. 3 The example of the point visible to at least three readers with no barrier

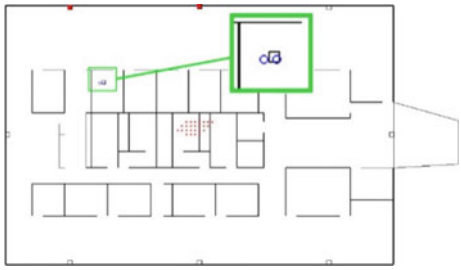


Fig. 4 The example of the point behind barriers

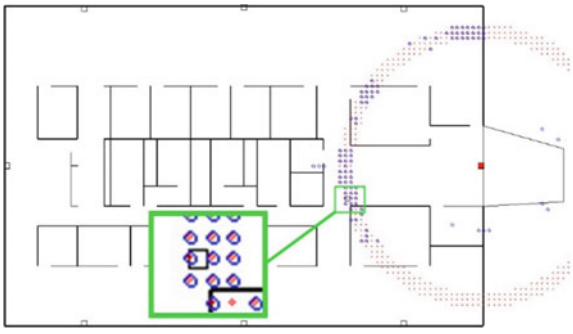


Fig. 5 The example of the point that is visible only to one reader

4 Discussion

Where there is no barrier between the sender of a signal and receiver, Eq. 1 calculates the power of RSSI at a known distance (d) from the sender. Therefore, by knowing the RSSI, the distance between the sender and receiver can be calculated. For a single receiver, this calculation leads to a circle with the diameter of d for the position of the sender. By having two receivers, the calculated locations are reduced to only two points. Furthermore, by having three receivers and using the trilateration method, the calculation leads to only one point which would be the exact location of the sender. Normally in indoor positioning systems (IPS), there are barriers between any sender and receiver; however, still for accurate locating at least three readers should receive the signal.

Also, using Eq. 1 alone in IPS is not sufficient, and the effect of barriers also need to be considered. A barrier can reflect some of the RF waves and may allow some part to pass through it. Therefore, it affects the RF field around the sender. The ray tracing technique can be utilized to effectively calculate the effect of barriers on the RF waves. In this study, the maximum broadcast range for the sender was assumed to be 100 m. Also, the readers for the simulation were positioned close to the external walls with the distance between each two in the row of about 12 m. In an outdoor setting without any barrier, all of the readers are able to see the entire area, but the fingerprint maps for this simulation show that there are areas that none of the readers are able to see. Moreover, due to the attenuation of the signal when passing through a wall, most of the floorplan area would be visible only to one or two readers which is not sufficient to accurately locate the position of the sender.

The aim of the IPS is to locate objects and track their movement indoors. For the purpose of tracking objects, knowing the type of object and its movement characteristics such as the maximum possible speed and normal movement pattern would help the system to eliminate the unrealistic estimated locations of the object and improve the accuracy of the IPS result.

Finally, our proposed simulation was designed based on the Bluetooth Smart technology. However, some other wireless technologies such as Wi-Fi can be used instead, considering the range of the selected wireless technology.

5 Conclusion

In this paper, we implemented a simulation model that estimates the location of any BLE beacon using some stationary receivers. Also, this work makes a contribution towards the RTLS, and has the potential of improving the accuracy of indoor localization. Our simulation model employs both the floor plan of the area of interest and the ray tracing method to reduce localization errors where there are some obstacles in the selected environment. The followings are some of the results we found:

- This study shows that ray tracing is a suitable technique for generating the fingerprint map for IPS.
- Although the broadcast range of BLE tags are about 100 m, the result of this study shows that even the distance of 24 m between readers is not sufficient for having an accurate IPS system.
- Knowing the type of the object and its movement characteristics would help in tracking the object more accurately by eliminating the unrealistic estimated locations of the object.

Acknowledgements This work has been supported by the Curtin University and Commonwealth's support through Curtin Research Scholarship (CRS) and Research Training Program (RTP) respectively.

References

1. Bisio I, Sciarrone A, Zappatore S (2016) A new asset tracking architecture integrating RFID, bluetooth low energy tags and ad hoc smartphone applications. Pervasive Mob Comput 31:79–93
2. Chapre Y, Mohapatra P, Jha S, Seneviratne A (2013) Received signal strength indicator and its analysis in a typical WLAN system (short paper). In: 2013 IEEE 38th conference on local computer networks (LCN). IEEE
3. Elbakly R, Youssef M (2016) A robust zero-calibration RF-based localization system for realistic environments. In: 2016 13th annual IEEE international conference on sensing, communication, and networking (SECON). IEEE
4. Fafoutis X, Mellios E, Twomey N, Diethe T, Hilton G, Piechocki R (2015) An RSSI-based wall prediction model for residential floor map construction. In: 2015 IEEE 2nd world forum on internet of things (WF-IoT). IEEE
5. Farid Z, Nordin R, Ismail M (2013) Recent advances in wireless indoor localization techniques and system. J Comput Netw Commun
6. Hussin Z (2014) Fast-converging indoor mapping for wireless indoor localization. In: Seventh annual PhD forum on pervasive computing and communications
7. Ji Y, Biaz S, Pandey S, Agrawal P (2006) ARIADNE: a dynamic indoor signal map construction and localization system. In: Proceedings of the 4th international conference on mobile systems, applications and services. ACM
8. Jiang Z, Zhao J, Han J, Tang S, Zhao J, Xi W (2013) Wi-Fi fingerprint based indoor localization without indoor space measurement. In: 2013 IEEE 10th international conference on mobile ad-hoc and sensor systems (MASS). IEEE
9. Kriz P, Maly F, Kozel T (2016) Improving indoor localization using bluetooth low energy beacons. In: Mobile information systems
10. Lymberopoulos D, Liu J, Yang X, Choudhury RR, Handziski V, Sen S (2015) A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned. In: Proceedings of the 14th international conference on information processing in sensor networks. ACM
11. Motamedi A, Soltani MM, Hammad A (2013) Localization of RFID-equipped assets during the operation phase of facilities. Adv Eng Inf 27(4):566–579
12. Padilla J, Padilla P, Valenzuela-Valdés J, Ramírez J, Górriz J (2014) RF fingerprint measurements for the identification of devices in wireless communication networks based on feature reduction and subspace transformation. Measurement 58:468–475
13. Schulcz RO, Varga GA, Tóth LAO (2010) Indoor location services and context-sensitive applications in wireless networks. In: International conference on indoor positioning and indoor navation. Zurich, Switzerland
14. Schwarz D, Schwarz M, Stückler J, Behnke S (2014) Cosero, find my keys! Object localization and retrieval using bluetooth low energy tags. In: Robot soccer world cup. Springer
15. Thaljaoui A, Val T, Nasri N, Brulin D (2015) BLE localization using RSSI measurements and iRingLA. In: 2015 IEEE international conference on industrial technology (ICIT). IEEE
16. Torres-Solis J, Falk TH, Chau T (2010) A review of indoor localization technologies: towards navigational assistance for topographical disorientation. INTECH Open Access Publisher

17. Wang S, Wen H, Clark R, Trigoni N (2016) Keyframe based large-scale indoor localisation using geomagnetic field and motion pattern. In: IEEE/RSJ international conference on intelligent robots and systems (IROS). Daejeon, Korea
18. Wu C, Yang Z, Liu Y (2015) Smartphones based crowdsourcing for indoor localization. IEEE Trans Mob Comput 14(2):444–457
19. Yang Z, Wu C, Liu Y (2012) Locating in fingerprint space: wireless indoor localization with little human intervention. In: Proceedings of the 18th annual international conference on mobile computing and networking. ACM
20. Zàruba GV, Huber M, Kamangar F, Chlamtac I (2007) Indoor location tracking using RSSI readings from a single Wi-Fi access point. Wirel Netw 13(2):221–235

Assessing Total Cost of Ownership: Effective Asset Management Along the Supply Chain

Amir Noorbakhsh, Carla Boehl and Kerry Brown

Abstract Firm purchasing models are an important tool for effectively managing business costs over the different time horizons and business cycles of the organisation. However, there is a lack of a clear understanding of the optimum method for supporting purchasing decisions for businesses and less knowledge of the costs over the entire supply chain of a product or service. Total cost of ownership (TCO) method helps the purchaser to compare different goods or service providers with regard to total cost rather than choosing a supplier due to their lowest initial price alone. The mining industry is an important sector in which to investigate the TCO as it is a valuable sector to the economy and the long-time horizons and large-scale projects require careful planning and management. This paper investigates models for purchasing and assesses the optimal performance issues accorded to purchasing and firm costs in the mining sector by the various methodologies. The notion and use of TCO as a basis for reviewing supplier selection models will be explored and investigated in detail. This review outlines the impediments of feasible TCO analysis execution and reveals common obstacles that must be attentively considered in the supply chain across mining industries assets.

1 Introduction

The overall objective of this paper is to assess various purchasing models and considers all obvious or hidden cost estimations over the entire life cycle of the mining assets. Further, it examines total cost of ownership along the supply chain

A. Noorbakhsh (✉) · C. Boehl
Western Australia School of Mines: Minerals, Energy and Chemical Engineering, Curtin University, Bentley 6102, WA, Australia
e-mail: amir.noorbakhshr@postgrad.curtin.edu.au

C. Boehl
e-mail: carla.boehl@curtin.edu.au

K. Brown
School of Business and Law, Edith Cowen University, Joondalup 6027, WA, Australia
e-mail: k.brown@ecu.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_43

and asset management as a valuable support for making or buying decisions of both products and services. Total Cost of Ownership (TCO) is a concept widely used in many businesses and industries around the world. TCO is a technique which provides valuable information for suppliers to make the best decision in relation to buying particular goods and services. TCO is a complicated, nested concept which not only determines the most critical or integral costs in the acquisition, possession, use and subsequent disposal of a good or service but also can specify elements such as order placement/replacement, supplier's qualification, shipping/hauling, acquiring, audit, refusal, displacement, downtime caused by failure, and costs of disposal.

Evaluation of actual TCO is challenging due to the cost implied calculation, ability to collect sufficient data and difficulty in estimating both obvious/direct and hidden/indirect machinery and services costs. In the modern supply chain perspective, transactions are not limited to just buyer and seller but contain third or even fourth party such as service providers and value added sellers. The paper analyses models of purchasing for the mining industry, in which manufacturers of mine machinery sell products and related services to mining companies, who operate them in the mine operation. This research assesses the total cost of ownership model as a tool in the mining supply chain concept to examine and improve purchasing decisions by foreseeing different attributes such as acquisition cost, quality, and the reliability of supplier's delivery. Mining companies require a thorough and highly developed TCO model to encompass all product life cycle costs and also be adaptable to supply chain principals and perspectives. This TCO model is aimed at providing insights into the real cost in the supply chain across mining industries assets.

The purpose is to develop an actual and implementable TCO model which can be tested with extensive collected real data and can be adaptable to the different supply chain actor's perspective. The aim is to analyse the actual firms' practices of the TCO model and develop an alternative framework for TCO. We argue that an efficient TCO model can increase suppliers' performance and improve cost justification according to TCO rather than just the price. This study examines TCO various cost parameters and proposes the procedures that support the total cost of ownership systematic application in mining industries.

2 Research Method

The first step in this research is a comprehensive literature review around Total Cost of Ownership (TCO). The literature review will reveal the limitations and construct a pre-framework for the TCO model along the supply chain and asset management. Much of the TCO literature involves only one case study firm. Hence, lack of sufficient data has been encountered. An exploratory research method is proposed for this research, involving surveys of companies and manufacturers within mining industries and case studies of mining organisations and their supply chains. Data

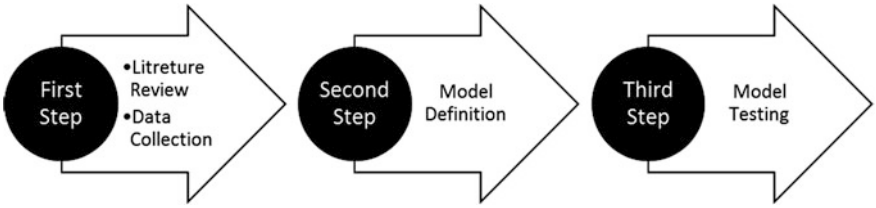


Fig. 1 Research method steps

collection will be accomplished through interviews with purchasing practitioners, supply chain employees or managers, maintenance service providers.

Understanding all the direct and indirect costs paid by the industries which is the main pillar of the TCO model will take place at this step. A very small discrepancy in cost issues could have significant effect on the estimation of actual cost that would lose million dollars (Fig. 1).

Careful literature review and data collection enable our model to be defined. All relevant costs including direct or indirect should be identified and classified to be calculated and evaluated. The model will be tested and implemented with real data over a specific duration of time and then the result will be checked and analysed precisely.

3 Review of Purchasing Methods

One of the pivotal activities of any competitive business is purchasing which accounts for 60–70% of total expenditures in manufacturing [24]. Supplier selection and actual cost purchasing strategies are key strategic issues for businesses to remain competitive [16]. Actual cost of purchasing including direct and indirect cost are often ignored by most of the traditional purchasing and supplier selection methods because of their frameworks which are based on quoted price selection [31].

Prior research found 50–90% of production costs can be attributed to purchasing function costs [11, 20, 34]. In 2009, Micheli, Cagno and Di Guilio conducted a study focused on the company's internal relationship and suppliers because of reasons such as product lifecycle reduction, products and services complication increasing and complexity of the process in enterprises [28].

Total cost of ownership (TCO) is one of the purchasing tools in many enterprises which is based on true cost of buying of peculiar products or services [17]. The aim of TCO is to specify the actual costs of buying and owning of products and services. These actual costs can comprise order placement, order replacement, supplier selection, shipping, logistic, inspection, return, and disposal. Total cost of ownership is an important tool to support strategic cost management due to considering the impact of the purchase decision on business costs as well as other cost parameters [20].

There are other methods which are close to TCO but have some weak points which make these unpopular among industries. Life-cycle costing ignores pre-purchase costs and target determination of the life cycle costs of the assets including operating, maintaining, and disposing and generally based on fixed assets [18]. Another method is zero-base pricing which is close to TCO but just considers the cost of running the business and a pricing framework for suppliers [7]. Cost-based supplier performance evaluation ignores the internal cost of business and only focuses on suppliers' external cost [5].

There have been some models which emphasise TCO. Handfield and Pannesi are among the first authors who believe that the overall life cycle of the product can be unrelated to the component of life cycle [23]. Modified cost ratio method is used by Carr and Ittner for identification of key factors which are the reason for costs increasing [9]. Caniato et al. classified TCO models to standard and ad hoc models [8]. According to the authors, standard models encompass all goods and services, whereas ad hoc models are just for a particular product. Agrawal and Graves define general standard TCO model for cost estimation [2] whilst Fornasiero, Zangiacomi, and Sorlini define an ad hoc model which particularly focused on truck's life cycle [22].

Ellram in her research develops the dollar-based model to value base model with an emphasis on nonmonetary measures. Hence, costs are classified to components, capital assets, maintenance and services which are not only used for purchasing departments, but also for other organisational departments [18]. Another TCO model focusing the monetary quantification of all financial and non-financial attributes is widely developed and implemented by Morssinkhof et al. [32]. In order to optimise the supply chain, Cavinato proposes a developed version of the TCO model based on reducing the costs and increasing the value even for the very last customer [10]. There are other researchers in the field of firm purchasing optimisation which first start by cost categorising based on product level, then defining the overall purchasing process model and finally calculating TCO by implementing a three-axes matrix [12–15]. Buy class model is another type of TCO model which is based on the type of the buy class/type of purchase and developed by Ellram [17].

This model can include both standard and ad hoc TCO models. Caniato et al. propose nine classification variables for TCO models [8].

- Goods and/or services
- Supplier and supply chain
- Efficiency and/or effectiveness
- Dollar and/or Value
- Ad hoc and/or standard
- Costing method
- Period
- Supplier
- Flowchart and/or matrix

Fig. 2 Adopted TCO calculation [30]

Generally, there is no standard formula for calculating TCO due to variety of procurement specific nature but it can be calculated as below [30] (Fig. 2).

Eventually, Total Cost of ownership can be determined as the following equation [26]:

$$TCO = A + PV \sum_{i=1}^{n} (Ti + Oi + Mi - Rn) \quad (1)$$

where

TCO	Total Cost of Ownership
A	Acquisition cost
PV	Net present value
Ti	Training cost in year i
Oi	Operating cost in year i
Mi	Maintenance cost in year i
R_n	Recapitalisation value in year i

4 Discussion

The TCO method takes into account extra expenses such as cost of order implementation, expenditure associated with searching activities and supplier qualification, insurance, warranties, and quality costs, exchange possibilities and downtime caused by failures [1]. Theoretically, TCO concept is extensively accepted but prior research has shown practical restriction in applying this method [33]. Ellram and Siferd surveyed 103 people who were professional in purchasing and indicated that 24% of them did not apply TCO model at all. While just 18% of purchasing professionals had formal TCO model, 58% of them used informal TCO model to examine the purchasing items [19]. Ferrin and Plank revealed that among 62% of their surveyed companies, TCO is used in less than 40% of the purchases [21] which supports Ellram's statement about the difficulty of TCO methodology application [18]. Majority of surveyed companies which used formal TCO model indicated difficulties in accurately forecasting costs such as overhead costs, labour costs and etc. [29].

Wynstra and Hurkens did investigate number of Dutch companies and also interviewed plenty of purchasing managers about the TCO model applications; their

research showed that although managers believed that TCO is a valuable purchasing method, most of the TCO applicants didn't have proper experience in TCO execution [36]. There are some papers tried to investigate the difficulty of TCO implementation in the enterprises and discussed about the root causes of this difficulty. According to Ellram, the most critical issue is resource allocation including lack of advanced computerized information systems [17]. Another type of resource allocation barrier is shortage of experienced and well trained personnel to collect and analyse the accurate data [29].

Shortage of readily available data which are really difficult to thoroughly collect in companies become a significant impediment for TCO implementing in every organization [27]. Ellram indicated general resistance to change as one of the obstacles which numerously encountered in her research [20]. Hence, the cultural impedance for fully understanding the TCO, requires collaboration in all firms' personnel. Based on previous research, it is crystal clear that implementing operatable TCO model needs well equipped information system and well trained personnel who exactly know how to gather, analyse and simplify the precise and actual data. Preparing Training manuals and classes, skilled personnel employment, consultation with experts, and simplification of operating method can be helpful for a flexible and operatable accurate TCO model.

As mentioned before, purchasers require TCO analysis for examination of the supplier's offers and performance. On the other hand, TCO analysis can be helpful for convincing the purchasers to shift their valuation system from the initial price based to the total cost of ownership. TCO application is a complex concept which needs spending time to be operated smoothly but there are some reports which stated problem in execution of the TCO model; some customers believed that TCO was a time consuming process and they didn't have enough willingness to share data with other [3, 33]. Both buyers and suppliers firms need experts to collaborate closely in surveying and modelling for execution of TCO analysis [25]. Another discussion is that theoretical TCO application overlooks total revenue estimation; Wouters indicated that formal TCO analysis by alone is inadequate for purchasing decision, hence total value of ownership was suggested by them to expand TCO analysis [35]. Herrera Piscopo et al. believed that an optimised and developed total cost of ownership could include total value in TCO calculation [25].

Although TCO has some limitations, it remains a very impressive application in the assessment of purchasing alternatives value [20, 25, 29, 35]. As discussed before, TCO application recognised as a valuable purchasing tool for supplier selection and operated in many purchasing firms, but TCO analysis rarely employed by suppliers firms as a selling tool based on very limited experimental researches have of this method in industrial practice [4, 6, 25, 27]. There is a lack of academic and empirical research in mining firms in using of TCO analysis hence, this research will examine TCO method by contribution of mining suppliers and purchasers.

5 Conclusion

The purpose of this paper was to assess total cost of ownership through the analysis of published TCO empirical case studies and discuss the reasons for limited use of TCO method in practice. The results of this investigation show number of important limitations need to be considered. First, Resource allocation including lack of reliable information systems and experienced personnel. Second, Shortage of readily available data. Third, cultural problem like general resistance to change. Fourth, complexity of the TCO model. Finally, contribution of both purchasing and supplying firms in sharing their data. The limitations identified in this paper are not exclusive of each other but rather can have a compounding effect (interactive effect) on understanding and valuing the TCO.

Lack of a proper TCO model can lead to an enterprise losing competitiveness and profitability in the market due to poor cost estimation and poor purchasing decision making. What separates this TCO model from the typical TCO model is the ability to determine the actual cost and to optimize the supply chain relation result in minimizing cost of product and services. The model not only can drive supplier improvement and identify performance but also forecast new items performance according to historical data.

References

1. Afonso P (2012) Total cost of ownership for supply chain management: a case study in an oem of the automotive industry. In: Advances in production management systems. Competitive manufacturing for innovative products and services. Springer
2. Agrawal A, Graves RJ (1999) A distributed systems model for estimation of printed circuit board fabrication costs. Prod Plann Control 10:650–658
3. Anderson JC, Kumar N, Narus JA (2007) Value merchants: demonstrating and documenting superior value in business markets. Harvard Business Press
4. Anderson JC, Narus JA, Van Rossum W (2006) Customer value propositions in business markets. Harvard Bus Rev 84:1–4
5. Beamon BM (1999) Measuring supply chain performance. Int J Oper Prod Manag 19: 275–292
6. Brown RJ (1979) A new marketing tool: life-cycle costing. Ind Mark Manage 8:109–113
7. Burt DN, Norquist WE, Anklesaria J (1990) Zero base pricing: achieving world class competitiveness through reduced all-in-costs: a proactive handbook for general managers, program managers, and procurement professionals. Probus Professional Pub
8. Caniato F, Ronchi S, Luzzini D, Brivio O (2015) Total cost of ownership along the supply chain: a model applied to the tinting industry. Prod Plann Control 26:427–437
9. Carr LP, Ittner CD (1992) Measuring the cost of ownership. J Cost Manag 6:42–51
10. Cavinato JL (1992) A total cost/value model for supply chain competitiveness. J Bus Logistics 13:285
11. De Boer L, Labro E, Morlacchi P (2001) A review of methods supporting supplier selection. Eur J Purchasing Supply Manag 7:75–89
12. Degraeve Z, Labro E, Roodhooft F (2000) An evaluation of vendor selection models from a total cost of ownership perspective. Eur J Oper Res 125:34–58

13. Degraeve Z, Labro E, Roodhooft F (2004) Total cost of ownership purchasing of a service: the case of airline selection at Alcatel Bell. Eur J Oper Res 156:23–40
14. Degraeve Z, Roodhooft F (1999) Improving the efficiency of the purchasing process using total cost of ownership information: the case of heating electrodes at Cockerill Sambre SA. Eur J Oper Res 112:42–53
15. Degraeve Z, Roodhooft F, van Doveren B (2005) The use of total cost of ownership for strategic procurement: a company-wide management information system. J Oper Res Soc 56:51–59
16. Dogan I, Aydin N (2011) Combining Bayesian networks and total cost of ownership method for supplier selection analysis. Comput Ind Eng 61:1072–1085
17. Ellram L (1994) A taxonomy of total cost of ownership models. J Bus Logistics 15:171
18. Ellram LM (1995) Total cost of ownership: an analysis approach for purchasing. Int J Phys Distrib Logistics Manag 25:4–23
19. Ellram LM, Siferd SP (1993) Purchasing: the cornerstone of the total cost of ownership concept. J Bus Logistics 14:163
20. Ellram LM, Siferd SP (1998) Total cost of ownership: a key concept in strategic cost management decisions. J Bus Logistics 19:55
21. Ferrin BG, Plank RE (2002) Total cost of ownership models: an exploratory study. J Supply Chain Manag 38:18–29
22. Fornasiero R, Zangiacomi A, Sorlini M (2012) A cost evaluation approach for trucks maintenance planning. Prod Plann Control 23:171–182
23. Handfield RB, PannesI RT (1994) Managing component life cycles in dynamic technological environments. Int J Purchasing Mat Manag 30:19–27
24. Heberling ME (1993) The rediscovery of modern purchasing. Int J Purchasing Mat Manag 29:47–53
25. Herrera Piscopo G, Johnston W, Bellenger DN (2008) Total cost of ownership and customer value in business markets. In: Creating and managing superior customer value. Emerald Group Publishing Limited
26. Hines T (2014) Supply chain strategies: demand driven and customer focused. Routledge
27. Hurkens K, Valk W, Wynstra F (2006) Total cost of ownership in the services sector: a case study. J Supply Chain Manag 42:27–37
28. Micheli GJ, Cagno E, Di Giulio A (2009) Reducing the total cost of supply through risk-efficiency-based supplier selection in the EPC industry. J Purchasing Supply Manag 15:166–177
29. Milligan B (1999) Tracking total cost of ownership proves elusive. Purchasing 127:22–23
30. Ministry of Business IEONZ (2013) Total cost of ownership: an introduction to whole-of-life costing. In: Suppliers GAA (ed) Government Procurement Branch: Ministry of Business, Innovation & Employment of New Zealand
31. Mohammady Garfamy R (2006) A data envelopment analysis approach based on total cost of ownership for supplier selection. J Enterp Inf Manag 19:662–678
32. Morssinkhof S, Wouters M, Warlop L (2011) Effects of providing total cost of ownership information on attribute weights in purchasing decisions. J Purchasing Supply Manag 17:132–142
33. Rosenback M (2013) Antecedents and obstacles to total cost of ownership analysis in industrial marketing: a case study. In: 29th IMP conference, Atlanta, Georgia
34. Weber M, Hiete M, Lauer L, Rentz O (2010) Low cost country sourcing and its effects on the total cost of ownership structure for a medical devices manufacturer. J Purchasing Supply Manag 16:4–16
35. Wouters M, Anderson JC, Wynstra F (2005) The adoption of total cost of ownership for sourcing decisions—a structural equations analysis. Acc Organ Soc 30:167–191
36. Wynstra F, Hurkens K (2005) Total cost and total value of ownership. Perspektiven des Supply Management. Springer

Developing a New DTIMS Predictive Model to Reduce Long Term Routine Maintenance

Phillipa O'Shea, Hui Chen, Hamish Featonby and Benjamin Orpilla

Abstract The NZ Transport Agency's movement to the Network Outcomes Contract (NOC) procurement model requires suppliers to perform a maximum level of network renewal (Base Preservation Quantity—BPQ) with their periodic renewals program, which establishes a renewals budget whilst routine maintenance is funded from a lump sum. NOCs have a focus on delivering level of service outcomes via Operational Performance Measures (OPM's). Therefore for modelling the objective function shifts from being focused on Pavement Condition Index to being focused on Maintenance Efficiency (Periodic and Routine). The Downer NZ Model was developed to address this and is designed to model different 'family' (class, traffic, location, etc.) of roads, using historical maintenance data to predict the likelihood and quantum of future maintenance required for a particular road section. By using the 'family' of roads approach allows for flexibility, to establish different deterioration rates depending on the key influencers of performance, such as the impact of drainage condition which is well documented in numerous papers to have major influence on pavement performance. The model is compatible with the latest dTIMS V9.5. Downer have utilised this model on all of their current NOC contracts and are currently in the testing validation stage. Initial hit rate analysis indicates that the model is generating strategies that align with the objective of Right Treatment at the Right Time in the Right location.

P. O'Shea (✉) · H. Featonby
Downer New Zealand, Palmerston North, New Zealand
e-mail: Phillipa.o'shea@downer.co.nz

H. Featonby
e-mail: hamish.featonby@downer.co.nz

H. Chen
Downer Australia, Melbourne, Australia
e-mail: hui.chen@downergroup.com

B. Orpilla
Downer New Zealand, Dunedin, New Zealand
e-mail: benjamin.orpilla@downer.co.nz

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_44

1 Introduction

With the introduction of New Zealand Transport Agency's (NZTA) Network Outcome Contracts (NOC), a new model was required to assist in the development of renewal programmes that are focused on operational outcomes rather than pavement performance outcomes.

A team of Asset Engineers from both Australia and New Zealand, with key experts Peter Mortimer, Paul O'Docherty, Amy Wade and Dr Nabin Pradhan were assembled with a purpose of "developing a model that selects the most appropriate treatment length sections for pavement resurfacing and rehabilitation, so as to maximise asset life and minimise the pavement maintenance activities undertaken on the entire network, while maintaining the required network outcomes". Several workshops were conducted and this paper details their findings and the resultant model.

2 Model Conceptual Design and Framework

The development of the model took into account the following considerations:

- Developing a generic modelling framework, with links to existing data systems, that can be customised for implementation on a range of NZTA networks
- Robust model that can be utilised for future pricing of both NZTA NOC contract areas as well as local government contracts
- Refinement of the model to ensure the renewal programme outcomes meet the NZTA Contract Specification for BPQ and forecast condition
- Ability to explore multiple maintenance strategies and measure outcomes of these strategies e.g. constrained v unconstrained
- Ability to establish the risk of different maintenance strategies to allow commercial decision making based on sound analysis e. BPQ's v BPQ-10%.

2.1 Overall Objective

Based on the documents referenced it was determined that the overall objective of the model is to be able to robustly justify asset renewal programs using evidence based methodology that follows the philosophy of delivering 'The Right Treatment in The Right Place at The Right Time with The Right Risk'. In order to achieve this the model must contain the following aspects:

- the forward maintenance projection must:
  - Be based only on cost elements that will be influenced by the alternative strategy being evaluated

  - Be based on actual costs
  - Be drawn from a robust maintenance cost record

- A minimum of three strategies must be compared under the new NPV method (rather than the previous "do minimum" and rehabilitation comparison). These are:
  - Scenario 1: Do Minimum
  - Scenario 2: Heavy Maintenance
  - Scenario 3: Full Renewal
- Economic Efficiency is the ratio of Net Present Values (NPVs) of the 30 year whole life costs and the Works costs over the contract period
- It is expected that at least 5–19% of the surface area (depending on road classification) of the treatment length will have been previously repaired or genuinely in need of repair before a positive NPV will be achievable.

2.2 Treatment Strategy and Operational Considerations

In conjunction with the overall strategy other key considerations that are relevant to the model design we taken from Downer's treatment selection strategy and operational considerations, such as:

- Clustering of maintenance treatment works to achieve economies of scale
- Drainage maintenance can delay or prevent pavement works
- Renewal treatment selection is aligned with Road Classification
- Use "families" of pavement characteristics to assist with tactical treatment selection using similar characteristics such as subgrade strength/sensitivity, traffic loading, winter maintenance, rainfall and base course properties
- Review the timing of renewal treatments to:
  - Rehabilitate rather than undertake large amounts of routine maintenance, where not economic to do so
  - Consider a heavy maintenance strategy for lower classification roads
- Utilising data from the Job Management System to inform on future maintenance needs, as the data has already been collected
- Implementing a holding strategy for treatment lengths that are programmed for an area wide treatment (AWT) in the current year and implementing a priority maintenance strategy for treatment lengths that are programmed for an area wide treatment (AWT) in the next two years when budget constraints don't allow for the full treatment.

3 Deterioration Modelling

3.1 Modelling Approach

The detailed design of the modelling approach focused on the use of Downer NZ 'All Faults' data, which represents surface and pavement defects such as potholes, cracking, flushing, etc. This data is collected and recorded into Downer's Job Management System (JMS) by network inspectors within each contract. The driving factors to use this data as the main driver in determining the performance of the model are:

- Utilises robust maintenance records and defect data that is continuously updated and represents the most up to date condition of the network
- Minimises the cost of data collection as the data is often collected already as part of each contract, therefore not requiring additional cost of condition rating.

Based on the overall framework of the conceptual design of the model, an empirical-probabilistic approach was adopted. This form of modelling looks at the historical performance (empirical data) of the pavement to determine the probability of the pavement reaching a certain condition state in the future [1]. This form of modelling was adopted to account for the variable nature of pavement performance and the occurrence of surface and pavement defects. Using this approach to modelling also allows for the use of different 'family' (class, traffic, location, etc.) of roads using historical maintenance data to predict the future likelihood and quantum of maintenance required for a particular road sections. It does not require a lot of data to start with and will learn with more data.

The 'family' of roads approach also allows for flexibility to establish different deterioration rates depending on the key influencers of performance, customise the model to any type of network and to investigate all aspects of the road corridor such as the impact of drainage condition.

3.2 Family of Roads and Defect Occurrence Probability

It was determined that for the first iteration of the model, the family groupings were based on the Network hierarchy and surface type, as this provided the best spread of data within each grouping and sufficient historical data to form meaningful trends.

For each family of roads a defects Transition Probability Matrix (TPM) was calculated using historical maintenance data (up to most recent periodic treatment) separated into five condition states for pavement and surface defects separately. Where:

- Pavement Defects = Depression, rutting, shear failure
- Surface Defects = Flushing, isolated cracking, scabbing

This was done as different types of defects required different treatments and also represents failure mode within the pavement. The condition states were based on extent of defect (% area of pavement section), the condition states are shown in Table 1.

Table 2 shows an example of a TPM. The matrix is used to calculate the future estimated condition state based on the probabilities and the current condition state as well as to calculate the future probabilities of each condition state.

Using the TPMs derived for each family of roads, the defects progression is calculated using the following equation for each year.

$$FCS = \sum_{n=1}^{5} (n \times p_n) \tag{1}$$

where:

FCS Future Condition State
n Condition State
P Probability of current condition state being rating 'n'

3.3 Long Term Pavement Performance Parameters

To complement the defects probability model, a number of empirical models were used for rutting and texture to assist with treatment triggering and defining intervention levels. The empirical models are based on historical deterioration rates calculated from historical laser profilometer data collected for the network,

Table 1 Defects condition states

	Condition state				
Defect extent	1	2	3	4	5
	$\leq 0.2\%$	>0.2% and $\leq 2\%$	>2% and $\leq 6\%$	>6% and $\leq 15\%$	>15%

Table 2 Defects transition probability matrix example

Current condition	Future condition state				
	1	2	3	4	5
1	0.88	0.10	0.02	0	0
2	–	0.75	0.20	0.05	0
3	–	–	0.70	0.21	0.09
4	–	–	–	0.80	0.20
5	–	–	–	–	1.00

currently collected annually for state highway networks and to a lesser extent for Local Authorities.

It was decided that a simple empirical approach to modelling for long term pavement performance parameters such as the above minimises the data requirements to run the model. Also by incorporating these parameters into the overall model design allows for room in the model to utilise more complex deterministic forms of modelling should they be required in future iterations, while preserving it's linkages to the treatment selection and intervention strategy.

4 Treatment Selection and Intervention Strategy

The treatment selection and intervention strategy is structured to satisfy the overall model framework and the objective of minimising long term routine maintenance costs. For the implementation of the model, there is a need to determine how the treatment strategies integrate together to enable the triggering of treatments to be logical.

To facilitate this concept, the treatments have been set up in the model in a hierarchical structure that is generally related to the pavement condition. In the development of this structure, there was need to consider the subsequent treatment particularly in relation to the holding treatments.

Figure 1 shows the simplified hierarchy of the treatments in relation to pavement condition as well as three distinct treatment strategies which apply to both spray seal and asphalt surfaced pavements.

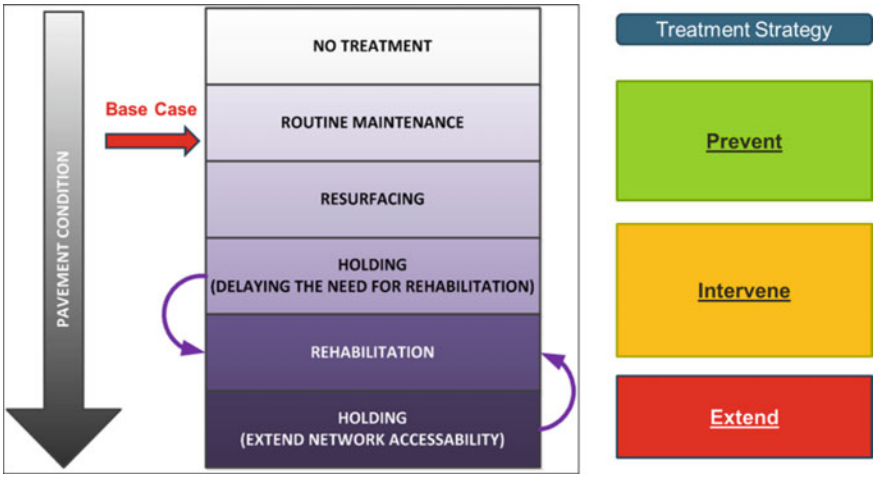


Fig. 1 Simplified hierarchy of treatments

While the purpose of the treatments is clearly identified, the treatment strategies are used to create the alignment between the concepts set out in the overall model framework to be able to justify treatment selection though the NZTA's Net Present Value requirements to consider Do Minimum, Do Something and Full Renewal through the model.

5 Optimisation—Maintenance Cost Minimisation

The philosophy of maintenance cost minimisation as the objective function for optimisation was achieved through a comparative approach between the routine maintenance effort before (Do Minimum) and after candidate treatment strategy has been applied, this is described below:

- Do Minimum—defined as the cumulative annual cost of rectifying pavement and surface defects over time, as modelled by the family of roads and probabilistic approach described earlier. No periodic treatments are carried out in this case where only routine maintenance is conducted for satisfactory levels of safety for road users.
- Treatment—periodic treatments are conducted when triggered. The effect on routine maintenance effort by the treatment is incorporated and compared.
- The model components were input into the dTIMS asset modelling and optimisation software, to model pavement performance, produce candidate treatments and optimise treatment selection on network level basis.

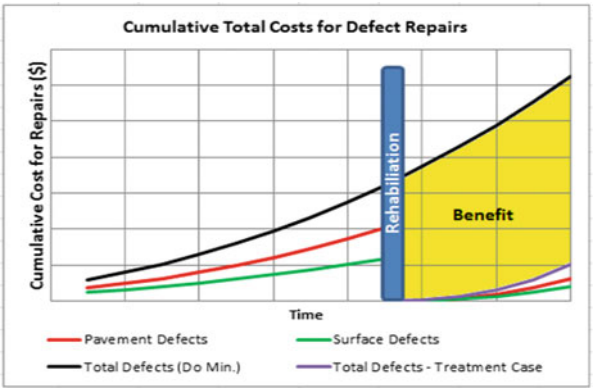
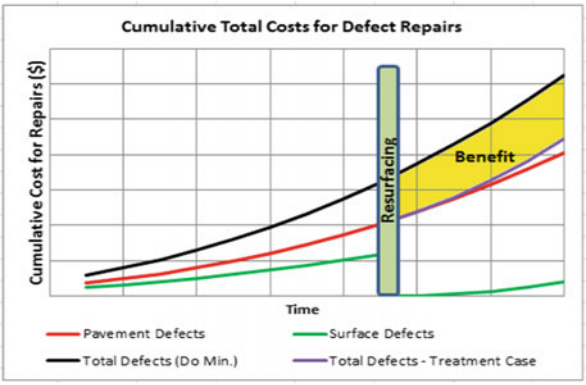
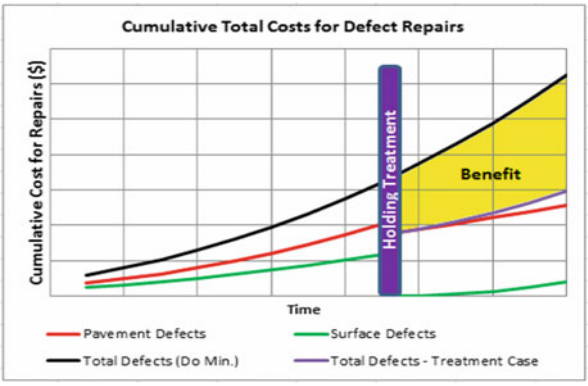
5.1 Treatment and Maintenance Interaction

The interaction between the treatments and routine maintenance is important to understand as this is the basis for determining the comparative routine maintenance effort which is in turn used for determining maintenance costs and the selection of treatments to use in the optimisation function. A schematic representation of the effect on routine maintenance for the generic treatments of resurfacing, holding treatment and rehabilitation are shown in Table 3.

6 Model Validation and Calibration

Field validation is required to determine the appropriateness and credibility of the model outputs against onsite engineering judgement. By comparing the generated strategies against the confirmed forward works programs and seeing why treatments are generated in real life, an iterative process can be taken to refine model outputs to

Table 3 Treatment and Maintenance Cost Interactions

Rehabilitation – Removes all pavement defects and surface defects – Has largest benefit – Most costly	
Resurfacing – Removes all surface defects – Has lowest benefit – Least costly	
Heavy maintenance treatment – Removes all surface defects – Removes some pavement defects – Mid range benefit – Mid range cost	

account for network specific and local engineering factors improving the credibility and accuracy of the model outputs.

Three stages of field validation have been performed for the model:

- Stage 1—Model development—Field validation conducted on South Canterbury NOC network for purposes of model development and initial refinement.
- Stage 2—Model refinement—Field validation conducted on Central Waikato NOC (CWNOC) and hit rate analysis against 2017/18 finalised FWP for purpose of model refinement across different networks.
- Stage 3—Implemented further refinements identified in Stage 2 and performed hit rate analysis against 2017/18 FWP again to see if it has improved

It is important to note that this was a 'blind' analysis since the Asset Engineer establishing the FWP didn't have access to the model output, therefore no bias was solicited.

6.1 Stage 2—Field Validation and Hit Rate Analysis Results

The results from the Stage 2 field validation produced a hit rate of 48%. There were a number of key findings and improvements identified to increase the hit rate of the model:

- Calibration of treatment intervention levels and limits—as different networks have their unique environment as well as local asset management practices. This highlights the need to calibrate treatment intervention on a per network basis
- First Reseal on rehabilitation sites—because the model is optimised based on defects, sites that have been recently rehabilitated and requiring the first reseal were not selected to limit risk and poor performance of subsequent resurfacing
- Repaired and open defects—The model relies on current open 'All Faults' data (i.e. defects requiring repair) to determine the condition state of the pavement. However, it was found that both historical (maintenance) and current 'All Faults' needs to be utilised to provide better treatment selection.

6.2 Stage 3—Model Refinement and Hit Rate Analysis Results

Based on the findings in Stage 2, the model was modified and refined to address these and a second output was produced and compared against the same approved confirmed FWP used in Stage 2. The results from the Stage 2 and 3 field validations

Table 4 Hit rate analysis results—Stage 2 v 3

Parameter	Stage 2%	Stage 3%
Total sites	–	–
Exact match	19	28
Matched within 3 years—optimised	21	23
Matched within 3 years—trigger based	7	17
Total hit rate	48	68

are shown in Table 4. The hit rate analysis methodology used is the same as those used for other roads model analysis.

The results show that there is marked improvement in all categories, especially with sites that are matched within 3 years of the optimised program jumping up to a combined hit rate of 51% from 40%. Two further improvements were identified to further increase the hit rate of the model:

- Optimisation objective function—incorporate long term performance parameters such as rutting, texture and roughness into the optimisation
- Trigger refinement—There are still gaps that have been identified in the model where sites are triggered in some years but not triggered in subsequent years.

7 Summary

The Downer model was developed primarily as a tactical tool to assist asset engineers and managers in effecting best use of resurfacing and renewal treatments, while improving efficiency in pavement and surface maintenance. This is achieved by predicting future maintenance costs and identifying high risk sites, from both defect and condition data. The main benefits for using a probabilistic model is that it is data lite and can learn and grow with more condition data. The model allows variability in road failure which is suited for operational asset management.

As a tactical tool, we see the model utilised in Downer-led road maintenance contracts to predict resurfacing and renewals with accuracy in alignment with engineering judgement.

The strategic benefit is in providing more certainty in FWP development and therefore maintenance cost prediction, within a lump sum environment.

7.1 Where to Next for the Model?

- From the validation exercises conducted, model improvements will be implemented to improve the triggering and optimisation of the model. This forms part of the continuous improvement process for the model within the Downer Integrated Maintenance Management System (IMMS) framework
- Process improvements will also be carried out around related activities that impact on the model run and output (i.e. All faults) as referenced in the IMMS
- Network calibration processes will be established to compensate for network specific behaviour which includes understanding achieved seal lives and chip grades
- Utilise established Pavement Classifications to develop further families and therefore TPM's.

7.2 Where Do We Run the Model?

- We will continue to run the model on our State Highway contracts as part of the Annual Plan delivery.
- Tararua, Whanganui, Hamilton and Waikato Alliance contracts will have runs carried out during this financial year. Local authority contracts can utilise this model having had an initial High Speed Data (HSD) run, without the investment and time required for a mature HSD set.

7.3 Looking into the Future

- We will develop the modelling of side drainage assets alongside pavement and surfacing.
- Integrating the Downer Model with Fuse—Deighton Operations (Downer's job management system). In the future, this will provide Fuse users the capability to run the model on the fly.

Reference

1. Costello SB, Snaith MS, Kerali HGR, Tachtsi LV, Ortiz-Garcia JJ (2004) Stochastic model for strategic assessment of road maintenance. In: Proceedings of the institution of civil engineers, Transport 158, November 2005

Efficient Evaluation of Internal Concrete Damage of Steel Plate-Bonded RC Slabs

Norihiko Ogura, Hitoshi Yatsumoto, Takahiro Nishida and Tomoki Shiotani

Abstract The Hanshin Expressway serves as an economic and industrial artery in Osaka, Japan. The 250 km long expressway network forms an important part of the infrastructure, being travelled by about 700,000 vehicles per day. After over 50 years in service since 1964, most of the structures are in need of large-scale reconstruction or rehabilitation. However, many challenges must be addressed to implement the renewal project efficiently. There are various types of works required, ranging from repair and strengthening to reconstruction of steel and concrete bridges. Other problems include considerably long work periods for the large-scale project, very limited space available for the work in a highly dense area, and adverse influence to the road traffic during the work. It is important to adopt methods and techniques that provide efficient cost reduction with minimum impacts to the public. Many of the reinforced concrete slabs on the Hanshin Expressway had been retrofitted with steel plates on the bottom surfaces to improve load carrying capacity and durability. In this study inspection and evaluation methods were developed for evaluating internal concrete damage without removing the existing steel plates.

N. Ogura (✉)
Engineering Department, CORE Institute of Technology Corporation, Tokyo, Japan
e-mail: ogura.nori@coreit.co.jp

H. Yatsumoto
Engineering Department, Hanshin Expressway Company Limited, Osaka, Japan
e-mail: hitoshi-yatsumoto@hanshin-exp.co.jp

T. Nishida · T. Shiotani
Laboratory on Innovative Techniques for Infrastructures, Department of Civil and Resources Engineering, Kyoto University, Kyoto, Japan
e-mail: nishida.takahiro.6e@kyoto-u.ac.jp

T. Shiotani
e-mail: shiotani.tomoki.2v@kyoto-u.ac.jp

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_45

1 Introduction

Many road bridges in Japan were built to old design standards, and reinforced concrete (RC) slabs of them are thinner than those of today. In order to improve their load carrying capacity and durability for an extended use in good condition, many of the old RC slabs have been strengthened by bonding steel plates to their bottom surfaces as shown in Fig. 1 (steel plate bonding method).

This method was first introduced in 1960s and has been applied to many bridge slabs. On the Hanshin Expressway, an important urban expressway network in the Kyoto-Osaka-Kobe area, about 70,000 panels (about 2–3 m on each side per panel) have been retrofitted with steel plates. However, the steel plate-bonded RC slabs have been in service over 30 years at maximum since the repair, and some of them are showing signs of deterioration.

The problem with the RC slabs repaired by this method is the steel plates covering the bottom surfaces which do not allow direct observation of internal concrete damage. There are some nondestructive testing techniques proposed for investigating the internal condition of the steel plate-bonded RC slabs. For example, hammer tapping test and thermography are commonly used to evaluate debonding of steel plates, and impact vibration test is used to evaluate rigidity of the while slabs. However, it is difficult to determine internal concrete damage of the RC slabs by any method [1].

This study proposes a technique developed for inspecting and evaluating internal concrete damage in existing steel plate-bonded RC slabs without restricting the traffic on them [2, 3].

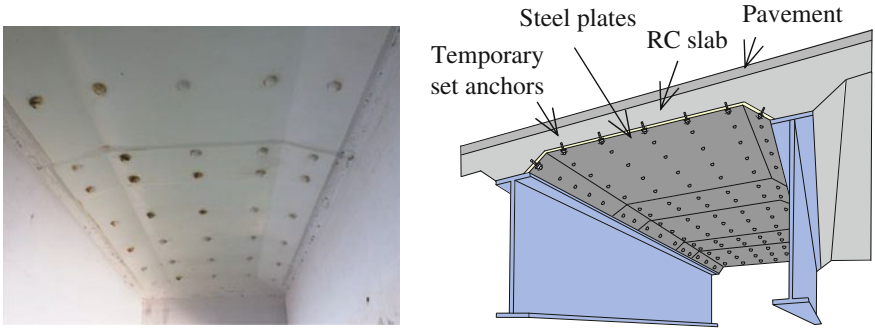


Fig. 1 A picture and a schematic of an RC slab bonded with steel plates

2 Internal Concrete Damage Investigation

2.1 Background of the Anchor Bolt Sensing Technique

Epoxy resin is used to bond the steel plates to the RC slabs in this method. Anchor bolts are driven into the RC slabs from the bottom as temporary retainers that hold the steel plates in place until the adhesive develops its full strength. Their depth of drive is 50–60 mm to ensure firm hold during the construction, which means they are in full contact with the inside of the concrete slabs. The authors focused on the anchor bolts which were present in all RC slabs bonded with steel plates, and developed a sensing technique using them effectively as ultrasonic probes.

2.2 Outline of the Internal Damage Detection System

Figure 2 shows the outline of the proposed anchor bolt sensing technique. An impact elastic wave is generated by hammering the head of anchor bolt which is protruding about 20 mm from the surface of the steel plate with a steel ball. Propagation of the generated wave is analyzed to determine internal concrete damage present inside the RC slab. The anchor bolts are used as probes in the proposed system, and the sensor installed on the head of the anchor bolt on the receiving side detects the elastic waves propagated through the RC slab. To know the excitation time and waveform features, the sensor is also installed on the surface of the steel plate immediately next to the impact-side anchor bolt.

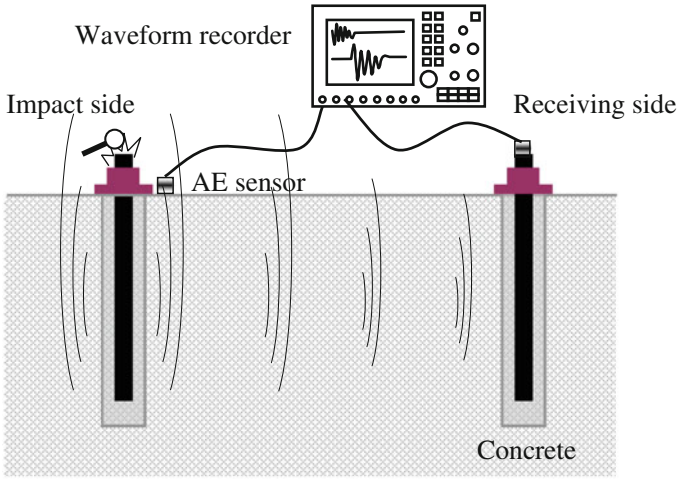


Fig. 2 Schematic diagram of measurement

Acoustic emission (AE) sensors (resonant frequency of 140 kHz) were used in the test for the measurement on both the impact and the receiving sides. Impact elastic waves were generated by hitting an anchor bolt at the head by using a steel ball with a diameter of 15 mm. At least three impacts were made per measurement, and a typical value was taken as the value to represent the measurement. It was ensured that none of the representative values were an outlier. Waveforms were recorded by the waveform recorder at a sampling rate of 0.2 μs with the number of samplings of 25,000.

2.3 *Verification Test Using Specimens*

Two reinforced concrete specimens ($210 \times 210 \times 2200$ mm^3) shown in Fig. 3 were prepared for the measurement by the anchor bolt sensing technique. One was a control specimen of sound concrete, and the other was a defect specimen of concrete with internal damage. The defect part was created by placing poor quality concrete to a depth of 20 mm from the interface with the steel plate. The numbered points in the figure represent the anchor bolts to which the sensor was attached for measurement. Measurement was also taken on the unstrengthened RC slabs before bonding the steel plates. In order to investigate the effect of bonding condition, each specimen was dividing into two parts at the middle between Points 3 and 4, and complete debonding was reproduced on the right half (poor bond zone) as shown in the figure. This makes Measurement pair 2–3 to represent fully bonded condition, Measurement pair 4–5, poorly bonded condition, and, Measurement pairs 2–4, 3–5 and 2–5, mixed condition with a half-bonded and half-debonded steel plate.

Due to the presence of the steel plates, propagation velocity of the first-arriving P-waves, which is commonly used for such analysis, cannot provide proper evaluation of internal concrete damage of the steel plate-bonded slabs. Instead, the authors investigated using the propagation of the Rayleigh waves excited at the anchor bolt which was oscillated by the impact of the steel ball. Wavelet transform was applied to the measured waveforms to analyze the magnitude of energy of the propagated waves. Figure 4 shows the results for Measurement pair 2–5 as an example of the experiment using the specimens. Time is on the horizontal axis, and frequency is on the vertical axis. The contour colors represent the spectral intensity. The closer to red, the higher the spectral intensity. Both models exhibited high spectral intensities in an approximate range from 500 to 2000 μs in time or at around 10 kHz in frequency. The high spectral intensities found after 500 μs at around 10 kHz in the contour graphs were considered to show the arrivals of the P-, S- and Rayleigh waves, in this order in accordance with their propagation velocities. Multiple reflection occurs in the anchor bolt when hit at the head, which likely makes the bolt to behave as a single oscillating body and excite the waves. Table 1 shows the resonant frequency of the longitudinal waves in the anchor bolt calculated by using the length and mechanical property values of the bolt. The frequency value for each mode varies depending on boundary conditions. With fixed-fixed

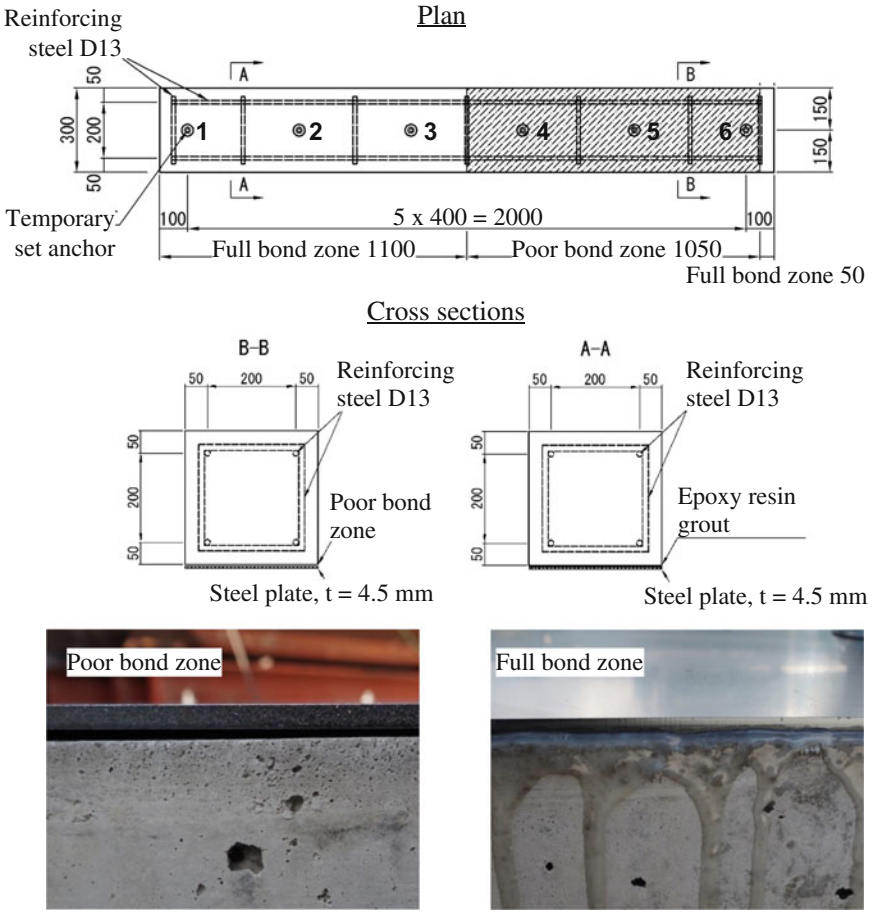


Fig. 3 Outline of the specimens

ends assumed for the first mode, the resonant frequency would be 10 kHz in the first mode and 20 kHz in the second mode. The theoretical solutions of resonant frequency suggest that significant oscillation generally occurs in the anchor bolt at around 10 kHz, although the theoretical values can vary depending on boundary conditions at the ends of the anchor bolt. This allows to conclude that the results for Measurement pair 2–5 with significant changes in spectral intensity at around 10 kHz well represent the propagation of the waves excited at the anchor bolt. Consequently, a special focus was placed on this frequency band around 10 kHz in the following analysis.

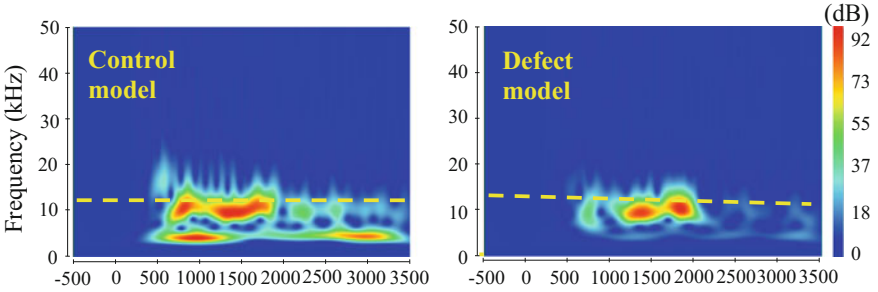


Fig. 4 Wavelet analysis results (measurement pair 2–5)

Table 1 Resonant frequency of the longitudinal waves in the anchor bolt

Oscillation modes	Fixed-fixed ends (kHz)	Fixed-free ends (kHz)
1	10.3	5.1
2	20.6	15.4
3	30.9	25.7

$$f_{\mathrm{n}} = \frac{n}{2L}\sqrt{\frac{E}{\rho}}$$

Fixed-fixed ends: n = 1, 2, 3, …
Fixed-free ends: n = 1/2, 3/2, 5/2, …

As a typical example, change with time in spectral intensity at 12 kHz was extracted from the frequency band in focus for further analysis on the wavelet analysis results. Figure 5 shows the diagrams of spectral intensity at 12 kHz shown in Fig. 4. The peak of spectral intensity appeared at around 800 μs in the control model, and at a delayed time around 1800 μs in the defect model. Considering the large energy of the Rayleigh waves, the peak of spectral intensity can be taken as the arrival of the Rayleigh waves which should be later than those of P- and S-waves. With this applied, arrival of the Rayleigh waves was later in the defect model than in the control model. Analysis on other measurement pairs resulted in a similar tendency in most cases where the arrival time was delayed in the defect model compared to the control model. Figure 6 shows comparative data of transmission velocity for Measurement pair 2–5 between the control and defect models. Unlike in the analysis focused on 12 kHz shown in Fig. 5, a frequency band between 10 and 15 kHz was covered in this analysis, with other typical frequencies shown for comparison. Transmission velocity was found to be lower in the defect model than in the control model. Although further research was needed to identify the reason for the variation in the decrease between the different frequencies, it was obvious that

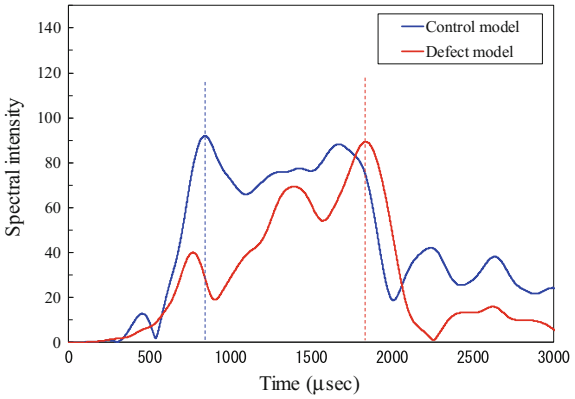


Fig. 5 Changes in spectral intensity at 12 kHz

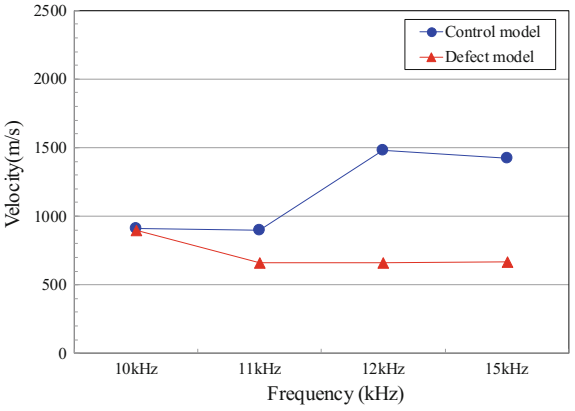


Fig. 6 Comparison of transmission velocity (measurement pair 2–5)

concrete with internal damage would exhibit a decrease in transmission velocity in the frequency band of the Rayleigh waves under the conditions of this experiment as proposed as an index of internal concrete damage in this study. Analysis on other measurement pairs showed similar results in most cases. As shown from the experimental results, it is possible to evaluate internal concrete damage of steel plate-bonded slabs by focusing on the characteristic frequency band of the Rayleigh waves and calculating the arrival time, or transmission velocity, of the waves.

3 Verification Test on a Steel Plate-Bonded RC Slab from an Existing Bridge

Inspection was carried out on an existing bridge in service for further verification of the anchor bolt sensing technique. The inspection included that by core drilling, and the results were compared with those by the anchor bolt sensing technique. Cores

Fig. 7 Road bridge in service (left: general view; right: a panel in focus)

were drilled at six locations in each of six panels (2 m × 3 m per panel) as shown in Fig. 7. This report describes the results with two of the six panels. Measurement was taken by using the anchor bolt sensing technique to record waveforms at propagation distances of 400–600 mm.

The information provided by the anchor bolt sensing technique is the transmission velocity of Rayleigh waves measured between the two anchor bolts that have traveled along a line between a pair of specific measurement points. In order to extend the linear information to a planar form, the data of limited points was spatially interpolated by Kriging method and transformed into a two-dimensional planar distribution.

Figure 8 shows the transmission velocity distribution in the steel plate-bonded RC slab in service measured by the anchor bolt sensing technique. The frequency in focus is 12 kHz. The red and blue dots (•) in the diagrams show the results of the inspection by core drilling. The red dots are the core holes with cracks found in the inside surfaces, and the blue ones are those without cracks.

It was found that transmission velocity tended to be low at locations where cracks were found in the inside surfaces of the core holes. In contrast, transmission velocity was as high as over 1000 m/s near the core holes without inside surface cracks. This suggests that a general correlation exists between internal damage and transmission velocity. On the other hand, some core holes had no inside surface cracks despite a low transmission velocity, or some had cracks despite a high transmission velocity, indicating a limitation of internal concrete damage evaluation using the transmission velocity. More detailed analysis will be made to identify the cause of the difference between the measurement by the anchor bolt sensing technique and the results of destructive test.

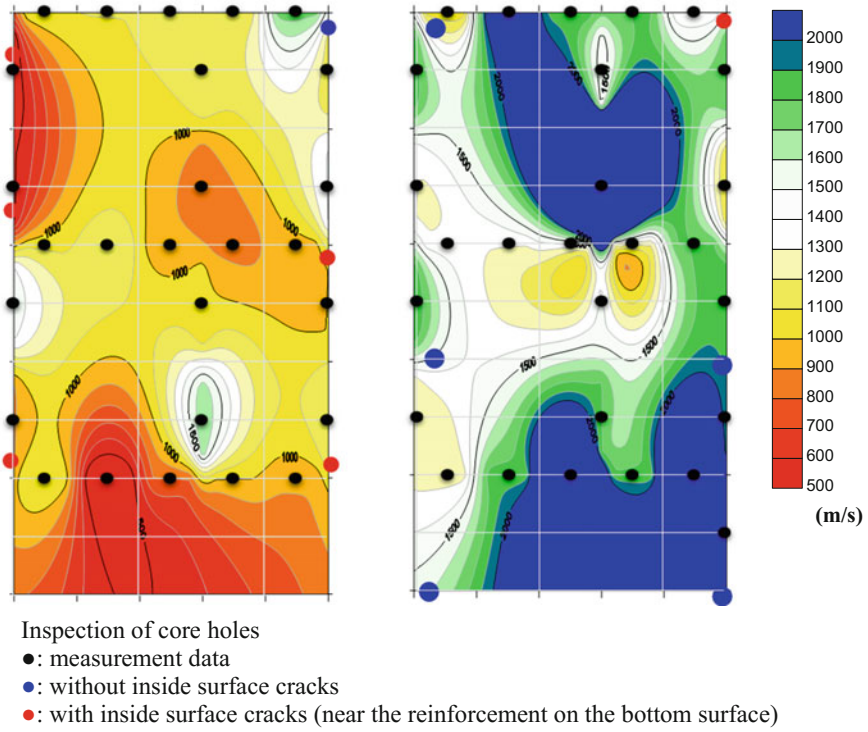


Fig. 8 Comparison between measurement by the anchor bolt sensing technique (transmission velocity) and inspection by core drilling

4 Conclusions

The anchor bolt sensing technique proposed in this study was found to be useful for evaluating internal concrete damage of steel plate-bonded RC slabs without removing the steel plates. The technique will enable efficient detection of those in need of repair or strengthening from existing steel plate-bonded RC slabs which are present in the order of several tens of thousands. Further research will be conducted to improve the detection accuracy of the system.

Acknowledgements This study was implemented in collaboration with the Laboratory of Innovative Techniques for Infrastructures, Kyoto University. Their great assistance and support to this research are gratefully acknowledged.

References

1. Sano M, Yamashita K, Matsui S, Horikawa T, Hisari Y, Niina T (2011) Evaluation of fatigue resistance and re-injection of resin for uplifted reinforced concrete deck slabs reinforced by steel plates bonding. J Jpn Soc Civ Eng Ser A1 (Structural Engineering & Earthquake Engineering (SE/EE)) 67(1):27–38
2. Ogura N, Yatsumoto H, Chang KC, Shiotani T (2015a) An ultrasonic method utilizing anchors to inspect steel-plate bonded RC decks. In: The 6th international conference on emerging technologies in non-destructive testing. CRC Press, pp 61–67
3. Ogura N, Yatsumoto H, Chang KC, Shiotani T (2015b) Evaluation of damaged RC decks with ultrasonics using the anchors in steel plate. In: JCI symposium on advanced NDE techniques for diagnosis and prognosis of concrete structures. Japan Concrete Institute, pp 27–32

Automated Bearing Fault Diagnostics with Cost-Effective Vibration Sensor

Agusmian Partogi Ompusunggu, Ted Ooijevaar, Bovic Kilundu Y'Ebondo and Steven Devos

Abstract *A non-continuous* condition monitoring approach through periodic vibration measurement has become a common practice in the industry. However, this approach can lead to serious misinterpretation, where rapidly growing faults, that might occur in rolling element bearings, could be missed. In contrast, *a continuous* condition monitoring approach offers a more optimal solution in which the bearing condition is continuously tracked. This way, catastrophic failures can be anticipated in advance thus allowing an optimal maintenance action. Despite its advantages, a continuous condition monitoring is not well adopted by the industry because of the high investment cost, where the sensor cost is a major factor. To remedy this gap, cost-effective vibration sensors are therefore needed. Yet, as shown in this paper, cost-effective vibration sensors inherently comprise some technical limitation, e.g. high background noise. These shortcoming needs to be solved before these sensors can be used in an industrial setting. This paper first discusses the selection of a cost-effective vibration sensor from the market and the sensor deployment for condition monitoring purposes. Afterwards, a novel diagnostics method that can deal with the high background noise of the sensor is presented. To demonstrate the feasibility of our approach, vibration signals acquired with a high-end accelerometer and the selected cost-effective accelerometer on an industrially representative gearbox were analysed. The results show that the diagnostics performance of the cost-effective accelerometer is comparable with the one of the high-end accelerometer.

1 Introduction

As the health of rotating machinery is significantly determined by the condition of rolling element bearings, monitoring the bearing condition therefore plays a vital role in the maintenance program of rotating machines. Vibration based condition monitoring is a well-established approach that has been employed by industries for

A. P. Ompusunggu (✉) · T. Ooijevaar · B. Kilundu Y'Ebondo · S. Devos
Flanders Make vzw, Celestijnenlaan 300, Leuven, Belgium
e-mail: agusmian.ompusunggu@flandersmake.be

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_46

many years in their maintenance program. The common practice of this approach is that vibration measurements are *periodically* recorded using *portable* vibration sensors (i.e. accelerometers) and measurement signals are analysed by an expert to interpret the bearing condition.

However, this common practice can lead to serious misinterpretation, where rapidly growing faults, that might occur in rolling element bearings, could be missed. In contrast, a continuous condition monitoring approach offers a more optimal solution in which the bearing condition is continuously tracked. In this way catastrophic failures can be anticipated in advance, thus allowing optimal maintenance action. Despite its advantages, a continuous monitoring program is however not well adopted by industry because of the high investment cost, where the sensor cost is a major factor. To overcome this gap, cost-effective accelerometers are needed.

Several publications concerning the use of low-cost MEMS accelerometer for condition monitoring purposes in general have become available in the literature [1, 4]. To the authors' knowledge, there are very limited works reported in the literature that discuss the use of cost-effective accelerometers for bearing condition monitoring [3, 5]. In the latter publications, it is not clear yet how the performance of cost-effective accelerometer compares with that of high-end accelerometer for bearing fault diagnosis.

In this paper, a selection of cost-effective accelerometers available on the market based on pre-determined technical requirements necessary for bearing condition monitoring purposes is first discussed. Further steps to deploy the selected accelerometers for condition monitoring purposes are then presented. Yet, as will be discussed later in this paper, the selected accelerometers inherently comprise some technical limitations, e.g. a high-background noise. A high background noise level can cause problems when applying the traditional diagnostics method based on the high frequency demodulation technique [8]. For reliable bearing monitoring, this shortcoming therefore needs to be solved. This paper presents a method for bearing fault diagnostics that can deal with the high-background noise limitation. In addition, the presented method also allows to provide an automated diagnostic decision, so human errors can be avoided, and affordable implementations of multi-smart sensor architectures can be achieved without employing expert users.

The remainder of this paper is organised as follows. Section 2 discusses the market survey for low-cost accelerometers suitable for bearing condition monitoring purposes. Section 3 presents a bearing diagnostics algorithm that can deal with the limitation of the selected sensor. Section 4 describes the experiment for validation. Results are discussed in Sect. 5 and conclusions are drawn in Sect. 6.

2 Market Survey, Sensor Selection and Sensor Deployment

High investment cost is one of the bottlenecks for adopting continuous condition based maintenance strategies in the industry. A major part of these costs is introduced by the sensors. Advancements in the field of MEMS accelerometers have

enabled opportunities for low-cost alternatives while maintaining basic-performance requirements for bearing condition monitoring purposes.

MEMS accelerometers offer many attractive attributes. They combine the economic benefit with, for example, a compact, a high sensitivity, a good resistance to shocks and an acceptable stability over a wide range of temperatures. However, their noise performance over higher frequency ranges is typically low, i.e. a noise density level in the order of mg/$\sqrt{\text{Hz}}$. Low noise MEMS accelerometers are available today with noise density levels anywhere from 10 to 100 μg/$\sqrt{\text{Hz}}$, but are restricted to a few kHz of bandwidth, as shown in Fig. 1. Other often reported limitations of the low-cost MEMS accelerometers are the long-term signal drift, bias offset and overall robustness of the sensor to industrial environments.

Based on the authors' experiences, a dynamic range of at least ±50 g and a ±3 dB frequency bandwidth up-to 10 kHz is desired for successful applications of bearing condition monitoring on industrial machinery. The market survey, finalised by 27-01-2017, over a wide range of commercially available sensors showed that the ADXL001-70 is the only low-cost (analog output) MEMS accelerometer that meets these performance requirements, see Fig. 1. The main limitation of the ADXL001-70 is the rather high noise density level of 3.3 mg/$\sqrt{\text{Hz}}$.

Proper hardware solutions were exploited to cope with the inherent limitations of the low-cost MEMS accelerometer that can affect the monitoring performance. To this end, a printed circuit board (PCB) and a tailored-packaging have been designed and produced. Figure 2 schematically illustrates the sensor deployment process to protect the MEMS sensor and enhance its overall robustness.

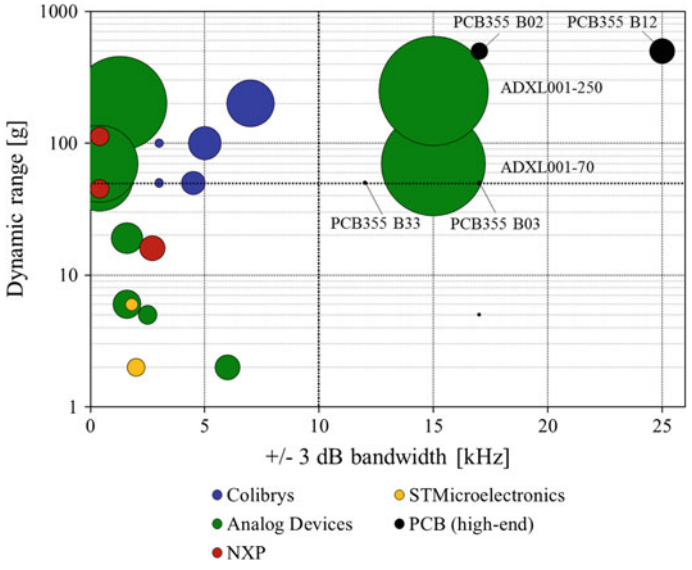


Fig. 1 A market overview of analog MEMS accelerometers. The dashed lines indicate the minimum requirements set for condition monitoring applications. The diameter of the circles indicates the noise density as specified in the datasheets

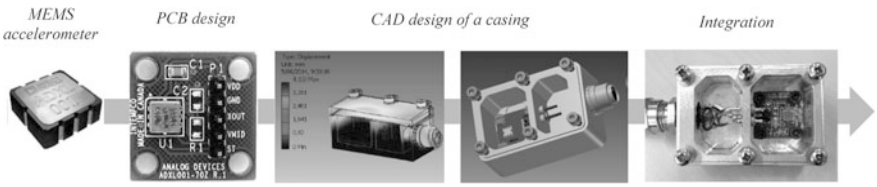


Fig. 2 Sensor integration process

3 Diagnostics Algorithm

The bearing fault diagnostics algorithm proposed in this paper is summarised in Fig. 3. In the first step, which is optional, the raw vibration signal is angularly resampled if the shaft speed variation is significant. Discrete components, which are dominant in vibration signals of rotating machinery, are removed from the raw signals by the phase-editing method described in Barbini et al. [2]. These discrete components are typically originated from gear-related signals and EMI-related signals. The filtered signal containing bearing-related information is then subsequently enhanced by the spectral subtraction and minimum entropy deconvolution (MED). The spectral subtraction aims for reducing random noise inherent in low-cost MEMS accelerometers, while the MED filtering aims for enhancing the impulsiveness of the bearing-fault related signal if present. The combined

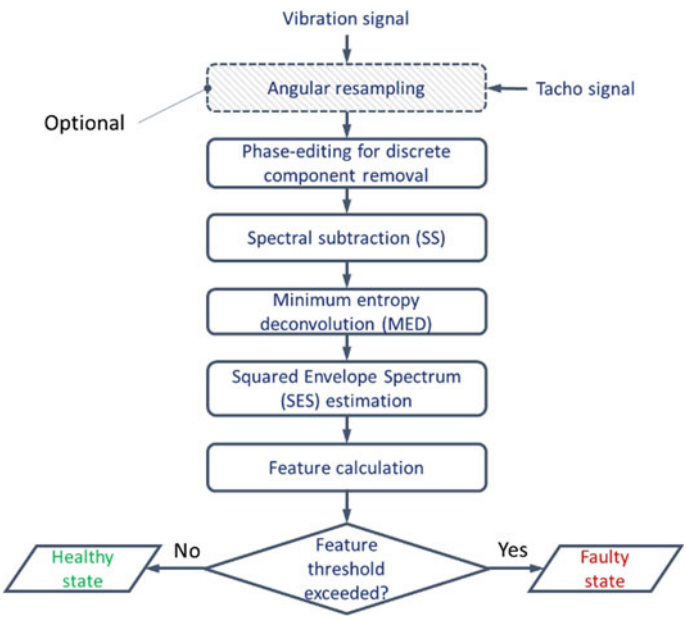


Fig. 3 Flowchart of the diagnostics method

processing steps will be described in detail in [6]. Bearing fault signatures are then identified through the squared envelope spectrum (SES) of the enhanced signal [8]. Furthermore, bearing fault features are computed from the SES for automated diagnostics decision.

The bearing fault feature F is defined as the summation of the normalised SES magnitudes $\overline{SES}$ of the first five harmonics of a bearing fault frequency of interest v_b. This definition can be formulated in the following equation:

$$F = \sum_{k=1}^{5} \overline{SES}[kv_b]. \tag{1}$$

The normalised SES magnitude $\overline{SES}$ is expressed in the following equation:

$$\overline{SES}[k] = SES[k]/SES[0] \tag{2}$$

with $SES[0]$ denoting the magnitude of SES at frequency zero. Finally, a diagnostic decision can be made automatically by comparing the computed feature value with a pre-determined threshold. A healthy state is reported if the feature value is smaller than the threshold, while a faulty state is reported if the feature value is larger than the threshold. Note that the threshold is determined based on the generalized extreme values (GEV) distribution using the feature values extracted from healthy states. The readers are referred to [7] for more details about the thresholding method.

4 Experiment

Figure 4 shows the photograph and the schematic-top-view of the gearbox setup used in this study. The test setup consists of (i) an induction electric motor, (ii) a gearbox and (iii) a magnetic brake. The motor is controlled by a variable-frequency-drive (VFD) with either a stationary mode or a transient mode (run-up/run-down). The motor speed can be controlled from 0 to 3000 rpm. The gearbox input shaft is connected to the motor through a flexible coupling, while the gearbox output shaft is directly coupled to the brake. The torque applied to the brake can be adjusted by the controller from 0 to 50 Nm.

As illustrated in Fig. 4(b), the gearbox is assembled with three-parallel shafts connected through contacting spur gear pairs. Note that the number of gear teeth is indicated in the figure by the number on each gear. Hence the total reduction factor from the input to the output shaft is equal to (100/29) × (90/36) = 8.62. The input shaft is supported by deep groove ball bearings MB ER-14K, while the other shafts are supported by deep groove ball bearings MB ER-16K. For simulating a healthy or faulty state on the gearbox, the left-side bearing housing that supports the second shaft (indicated by the oval dashed-line) is installed either by a healthy or a damaged bearing with an inner race fault. The theoretical fault frequencies of this bearing of interest (MB ER-16K) are listed in Table 1.

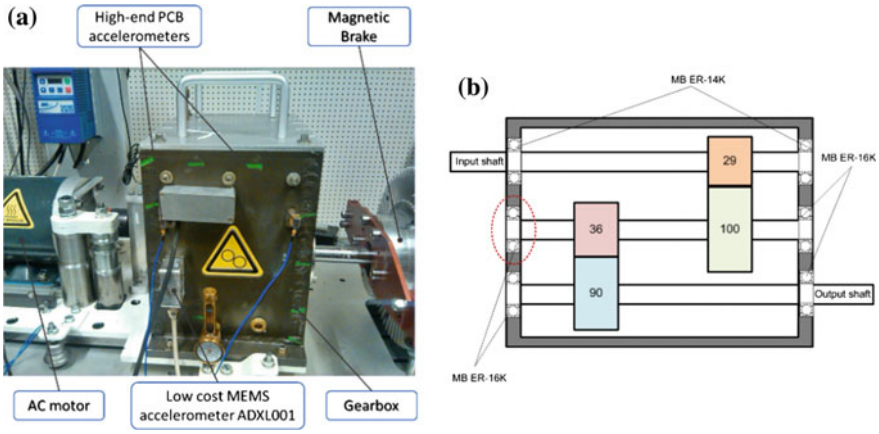


Fig. 4 Gearbox setup: **a** photograph, **b** gearbox layout

Table 1 Theoretical bearing fault frequencies of MB ER-16K bearing at the shaft speed of 60 rpm

Fault frequencies	FTF	BPFO	BDF	BPFI
Value (Hz)	0.402	3.572	4.644	5.430

For each healthy or unhealthy state, eight operating conditions were imposed on the gearbox setup, namely four different motor speeds of 2100, 2400, 2700 and 3000 rpm; and two different brake torques of 12.5 and 40 Nm. Because of the transmission ratio, the rotational speed of the second shaft is 29/100 lower than that of the motor speed, while the torque applied on the second shaft is 36/90 lower than that of the brake torque. Hence, for such imposed operating conditions, the rotational speeds of the second shaft are 609, 696, 783 and 870 rpm, while the torques applied to the second shaft are 5 and 16 Nm.

Different sensors, comprising high-end accelerometers (HE-PCB) and low-cost accelerometer (MEMS), were mounted on the gearbox housing, see Fig. 4(a). All vibration signals are acquired using an NI-DAQ system and then stored in a PC with a dedicated Labview program. For each operating condition, three runs are repeated. The data are then analysed using a script written in Matlab. Note that only vibration signals obtained with the high-end accelerometer (HE) and low-cost accelerometer (MEMS) mounted on the same location were analysed and will be discussed later in Sect. 5.

5 Results and Discussion

Figures 5 and 6 show typical raw vibration signals and the frequency spectra obtained from the high-end accelerometer (HE) and the low-cost accelerometer (MEMS) of the faulty and healthy gearbox. Note that the DC offset of the MEMS accelerometer signal has been removed here. As seen in the raw frequency spectra, the background noise level particularly at high frequency of the low-cost accelerometer is much higher than that of the high-end accelerometer.

It was identified that this high-frequency noise is a mixture of random noise and electromagnetic interference (EMI) noise. The random noise originates from the intrinsic property of the low-cost accelerometer, while the EMI noise originates from the motor controller, i.e. VFD. This EMI noise is indicated by the dominant peaks present at the carrier (switching) frequency of VFD and the harmonics in the

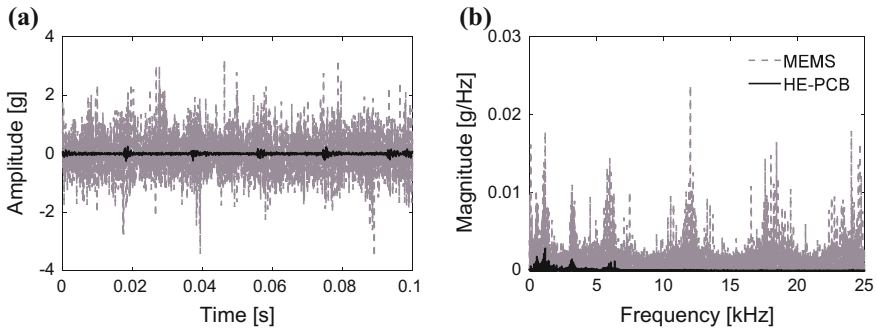


Fig. 5 **a** Raw vibration signals of high-end accelerometer (HE-PCB) and low-cost accelerometer (MEMS); **b** the raw frequency spectra, measured on the faulty gearbox under the motor speed of 2100 rpm and brake torque of 12.5 Nm

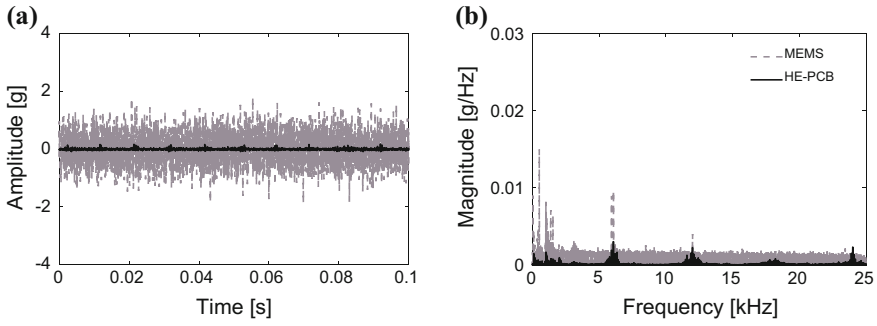


Fig. 6 **a** Raw vibration signals of high-end accelerometer (HE-PCB) and low-cost accelerometer (MEMS); **b** the raw frequency spectra, measured on the healthy gearbox under the motor speed of 2100 rpm and the brake torque of 12.5 Nm

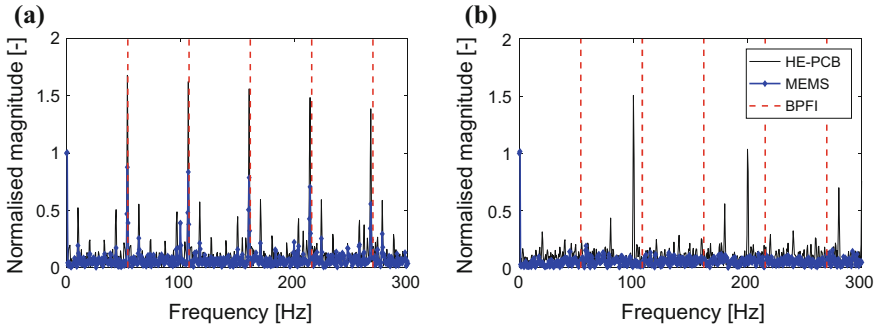


Fig. 7 The normalised envelope spectra estimated with the proposed diagnostics method. **a** Result of faulty gearbox and **b** result of healthy gearbox. The spectra are obtained from the analyses on the signals shown in Figs. 5 and 6

frequency spectrum of raw signal. In contrast to the measurement signal of the faulty gearbox, the high-end accelerometer signal is polluted by the EMI noise in the healthy gearbox measurement, as indicated by peaks at the harmonics of the VFD carrier frequency, namely around 6, 12, 18 and 24 kHz, see Fig. 6(b).

The diagnostics method discussed in Sect. 3 was applied to all the high-end and low-cost accelerometer signals measured at all the operating conditions as specified in Sect. 4. For demonstration, the resulting envelope spectra of the signals in Figs. 5 and 6 are shown in Fig. 7.

As seen in Fig. 7(a), the inner race fault signature, revealed by the peaks at the harmonics of BPFI (indicated by the dashed lines) and the sidebands around the peaks, is visible in the envelope spectra of both low-cost and high-end accelerometer signals. For this case, the fundamental BPFI is 53.5 Hz and the sideband spacing is about 10 Hz, which is equal to the second shaft rotational speed, see Fig. 4. One can see that the BPFI signature in the envelope spectrum of the high-end accelerometer signal is still preserved until the first five harmonics, while the BPFI signature in the envelope spectrum of the low-cost accelerometer signal is prone to attenuate after the first four harmonics. Notably, the dominant peaks at the harmonics of the doubled line frequency (100, 200 and 300 Hz) in the envelope spectrum of the high-end accelerometer signal indicate that the signal is polluted by EMI noise as reported in [9].

The feature values were computed from both the high-end and low-cost accelerometer signals measured in all conditions as presented in Fig. 8. As seen, the feature extracted from both accelerometers can clearly discriminate between the healthy and faulty state, independent of the operating conditions imposed in this study. This may suggest that a feature threshold separating the two classes can be set independent of the operating conditions. Furthermore, one can observe that the degree of separation of the low-cost accelerometer is visually larger than that of high-end accelerometer. The possible explanation is that the high-end accelerometer signals are polluted by EMI noise during the measurement in the healthy state, resulting in a skewed distribution.

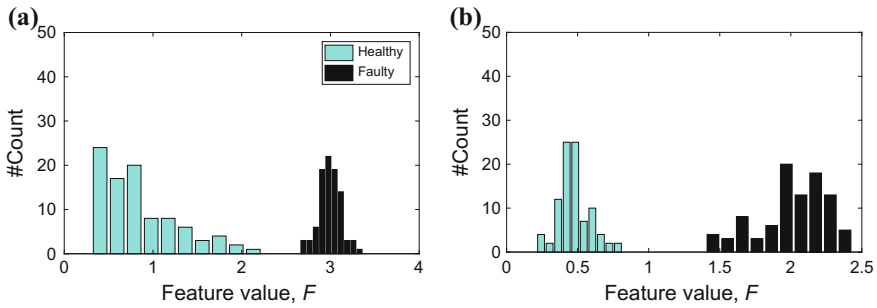


Fig. 8 Histograms of the feature values computed from the envelope spectra of healthy and faulty gearbox vibration signals of all operating conditions, **a** results of the high-end accelerometer signals and **b** results of the low-cost MEMS accelerometer signals

6 Conclusions

The feasibility of using low-cost MEMS accelerometer for bearing fault diagnosis on complex rotating machinery was investigated and is reported in this paper. It is shown experimentally that the selected low-cost MEMS accelerometer inherently comprises high frequency noise, which could be problematic when applying the traditional diagnostics method based on the high frequency demodulation technique. By using a novel automated diagnostics method, it was demonstrated in this paper that the diagnostic performance of high-end accelerometer is comparable with that of low-cost accelerometer.

Acknowledgements The results presented in this paper are the outcomes of the industrial research project VIBMON (Cost-effective vibroacoustic monitoring, project No. 150092), funded by Flanders Make vzw and VLAIO.

References

1. Albarbar A, Mekid S, Starr A, Pietruszkiewicz R (2008) Suitability of MEMS accelerometers for condition monitoring: an experimental study. Sensors 8:784–799
2. Barbini L et al (2017) Phase editing as a signal pre-processing step for automated bearing fault detection. Mech Syst Signal Process 91:407–421
3. Bechhoefer E, Wadham-Gagnon M, Boucher B (2012) Initial condition monitoring experience on a wind turbine. Minneapolis, Minnesota, USA, Prognostics and Health Management Society
4. Chaudhury SB, Sengupta M, Mukherjee K (2014) Vibration monitoring of rotating machines using MEMS accelerometer. Int J Sci Eng Res (IJSER) 2(9):J2013358
5. Maruthi GS, Hegde V (2014) Preliminary investigation on bearing fault analysis in three phase induction motor by MEMS accelerometer. IEEE, Bangalore, India
6. Ompusunggu A (2019) Spectral subtraction and minimum entropy deconvolution for enhanced vibration based bearing faults diagnosis. In preparation for publication

7. Ompusunggu AP, Y'Ebondo BK, Devos S, Petre F (2014) Towards an automatic diagnostics system for gearboxes based on vibration measurements. In: International conference on noise and vibration engineering (ISMA), Leuven, pp 2821–2835
8. Randall RB, Antoni J (2011) Rolling element bearing diagnostics—a tutorial. Mech Syst Signal Process 25:485–520
9. Smith WA et al (2016) Optimised spectral kurtosis for bearing diagnostics under electromagnetic interference. Mech Syst Signal Process 75:371–394

Integrated Modelling and Decision Support of Continuous Production Systems

Samuel Patterson, Paul Hyland and Talara Berry

Abstract The benefits that integrated modelling can provide to asset managers of production systems are clear and well known. This work presents a framework for formulating integrated Mixed Integer Linear Programming (MILP) models of continuous production systems. The framework promotes the creation of generalised, reusable models and domain expert engagement to reduce the upfront modelling effort required to generate value; reducing the payback period of modelling investment. Methodologies for using the framework to deliver support for strategic, tactical and operational asset management decisions are presented. Emphasis is placed on translating the model results to corrective actions for managers to make system wide improvements. The learnings from applying the concepts to two continuous production systems—an open-pit coal mine and a building energy plant—are then discussed. Finally, avenues to further verify the value and extend utility of the contributions are offered.

1 Introduction

To be effective when managing production system assets, considering decisions from a whole-of-system perspective is paramount. Being able to efficiently understand the system wide impact of even the most minor isolated decision can be a considerable advantage for asset managers in today's competitive global market.

Integrated modelling and analysis is a popular approach in research and industry for achieving overall improvement in a wide variety of complex production systems

S. Patterson (✉) · T. Berry
Synengco Pty. Ltd. & Queensland University of Technology, Brisbane, Australia
e-mail: sam.patterson@synengco.com

T. Berry
e-mail: talara.berry@hdr.qut.edu.au

P. Hyland
Queensland University of Technology, Brisbane, Australia
e-mail: paul.hyland@qut.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_47

[4, 6, 8, 11]. Here, integrated analysis refers to simultaneously optimising varying decisions from separate parts of a production system [12]. Continuous production systems are focused upon here as they allow the various subsystems to be connected to one another via material and energy flow connections [3, 10].

In the authors' previous work [9], an integrated Mixed Integer Linear Programming (MILP) model for optimising the energy efficiency of an open-pit coal mine was presented. The methodology used in that work involved constructing the integrated model out of individual subsystem modules, connected using a standardised, material flow interface. This approach was developed to reduce the effort of modelling mines that are different configurations of the same subsystems, which is a common trait of continuous production systems.

So that the benefits of the methodology presented in [9] can be realised by a wider audience of asset managers, this paper presents a more generalised modelling framework, exemplified using MILP, that promotes low cost model formulation with domain expert engagement in Sect. 2. Section 3 then presents methodologies to use the model for supporting strategic, tactical and operational decisions. Open-pit coal mine and a building energy plant examples demonstrate the worth of the contributions in Sect. 4. Finally, the concluding remarks in Sect. 5 set out an agenda for furthering the work in theory and practice.

2 Modelling Framework

Production systems that are made up of diverse assets can be difficult and costly to model using a single formulation structure. Here, we take a modular approach of splitting production systems into subsystems, modelling them separately and then integrating them. This way, modelling effort can be focused and suited specifically to the processes, behaviours and available data of each subsystem, independently of one another. These subsystem models can be also reused among many systems, which can reduce the upfront modelling costs significantly. This approach, however, can make connecting the subsystems in an integrated model difficult if their boundaries are also different to one another. For this reason a standardised interface (or connection) is proposed to enable modularisation [1]. Exploiting the continuous aspect of continuous production systems, flows or rates are used to connect subsystems [3, 10], for example mass flow, measured in kg/s, or energy flow, measured in kW (kJ/s).

This approach means that the production system's operation and capacity as a whole is implied by the individual operations and capacities of the subsystems that make up the system and the connectivity between them, much like real-life. This allows for a simple, generalised structure at the whole-of-system level, while still accounting for the emergent 'network effect' that comes from having connected subsystems with different processes, behaviours and capacities. This is a critical attribute for enabling managers to use the model for identifying bottlenecks and double-handling from a whole-of-system perspective.

2.1 Subsystem Modelling

Generalising upon the requirements explained in [9], a valid subsystem formulation must abide by the following requirements:

- boundary connection points are represented by a flow/rate variable
- the type of flow transferred over connection points is constrained internally
- a variable is provided for use in the objective.

The first requirement ensures subsystems can be connected together in any fashion necessary to represent a wide variety of production system layouts. The second requirement makes sure that the functional behaviour of subsystems is encapsulated within their formulation. This means no internal knowledge about a given subsystem is required at the whole-of-system level of the integrated model or within other subsystems that it may be connected to. The third ensures the overall objective can simply be a function of all subsystems' objective variables.

Several strategies for modelling subsystems are suggested with the reduction of upfront modelling cost in mind. Existing research literature can be a good source of mathematical models to replicate. Industry may also have resources that could be useful for defining the subsystem formulations; for example, control system models. The balancing speed and accuracy is required to ensure valuable results are achieved in a reasonable time. The availability of information may also impact the detail of the model. Focusing effort on potential bottleneck subsystems is also recommended to best consider their impact on the whole system.

The reusability of subsystem modules is also key [1]. Reusability can be achieved by keeping assumptions straightforward and creating generalised formulations. For example, a general storage tank subsystem formulation can be used across many subsystems in a single distillery plant; likewise, it could be used when across multiple distilleries. To make generalised models, empirical methods can be employed and fitted to operational data to represent a wider variety of processes. This can also help gain precision; by fitting the model with recent operational information, as opposed to theoretical, design or historical data, the solution can better represent the actual operation of the production system.

After starting with general formulations, it may then be necessary to modify particular subsystem models to suit the production system being modelled. This follows an 'agile' style of model development, whereby a running optimisation is always available, with extra effort undertaken where and when required. This ensures value is generated as soon as possible, meaning realised savings can fund further improvements, mitigating the risk of a large upfront modelling effort failing to produce the desired results.

2.2 Integration

The system layout is defined by sets of subsystems $q \in Q$; their inlet and outlet connection points, $i \in I_q$ and $j \in J_q$, respectively, along with a set of all connection points $k \in K = \cup_{q \in Q}\{I_q \cup J_q\}$; the connections between subsystem $c \in C$, that are tuples of outlet and inlet connection points $(c^{\text{out}}, c^{\text{in}})|c^{\text{out}} \in \cup_{q \in Q} J_q, c^{\text{in}} \in \cup_{q \in Q} I_q$; and finally the types of flow, $g \in G$, that can be transferred between subsystems. The time dimension of the optimisation is partitioned into multiple 'operating states', $t \in T$, that can have different lengths, H_t. An operating state is defined as the period of time in which no changes occur to the overall work being performed by the subsystems. The variable length of states gives the model the ability to have increased resolution when changes in operation may occur, and reduced resolution during stable operation. The continuous variable z_{tq} represents the contribution of each subsystem q during operating state t to the objective function, Eq. 1, which, for the examples in Sect. 4, is to minimise the sum of these variables. The continuous variable θ_{kgt}, represents the flow rate of type g being transferred over connection point k during operating state t. This is used in Constraint (Eq. 2) where the outlet and inlet connection point of each connection C are equated.

$$\text{Min} \sum_{\forall t,q} z_{tq} \tag{1}$$

$$\theta_{c^{\text{out}}gt} = \theta_{c^{\text{in}}gt} \tag{2}$$

Various tasks can be defined to specify the output required from the system. Here we present two tasks used for the examples in Sect. 4. The first, $b \in B$, sets the flow over a connection in Constraint (Eq. 3) that requires that there is between $\zeta_b^{\min}$ and $\zeta_b^{\max}$ of flow type ζ_b^{type} over the connection point $\zeta_b^{\text{con.p}}$ between the operating states ζ_b^{start} and ζ_b^{stop}. The second is an accumulated task, $a \in A$. This describes the requirement that the system produces between $\zeta_a^{\min}$ and $\zeta_a^{\max}$ of accumulated flow type (i.e. tonnes, if the flow is tonnes/hour), ζ_a^{type}, over the connection $\zeta_a^{\text{con.p}}$ between states ζ_a^{start} and ζ_a^{stop}. This is enforced with Constraint (Eq. 4); it assumes the units of H_t are the denominator of the flow $\theta_{\zeta_a^{\text{con.p}}\zeta_a^{\text{grade}}t}$.

$$\zeta_b^{min} \leq \sum_{t=\zeta_b^{start}}^{\zeta_b^{stop}} \theta_{\zeta_b^{con.p}\zeta_b^{type}t} \leq \zeta_b^{max} \tag{3}$$

$$\zeta_a^{min} \leq \sum_{t=\zeta_a^{start}}^{\zeta_a^{stop}} H_t \theta_{\zeta_a^{con.p}\zeta_a^{type}t} \leq \zeta_a^{max} \quad \forall a \tag{4}$$

To demonstrate using the framework structures to formulate a subsystem, a simple example subsystem module, $q_v \in V \subset Q$, is presented below. The subsystem has one internal binary decision variable, x_{gtq_v}, that is 1 when the subsystem is processing grade g and 0 otherwise. This is set by Constraint (Eq. 5), which also sets the minimum, $\pi^{min}_{q_v}$, and maximum, $\pi^{max}_{q_v}$, operating throughput. Constraint (Eq. 6) dictates that only one grade can be processed at a time and Constraint (Eq. 7) ensures flow is conserved between input and output. Finally, as required for each subsystem, Constraint (Eq. 8) sets the energy consumption, z_{tq_v}, as a function of a fixed cost, α_{q_v}, when operating and a variable cost, β_{gq_v}, proportional to the throughput. For more advanced examples of subsystem modules, see the model presented in [9].

$$x_{gtq_v}\pi^{min}_{q_v} \leq \sum_{\forall i} \theta_{igt} \leq x_{gtq_v}\pi^{max}_{q_v} \quad \forall g, t, q_v \tag{5}$$

$$\sum_{\forall g} x_{gtq_v} \leq 1 \quad \forall t, q_v \tag{6}$$

$$\sum_{\forall i} \theta_{igt} \leq \sum_{\forall j} \theta_{jgt} \quad \forall g, t, q_v \tag{7}$$

$$z_{tq_v} = H_t \alpha_{q_v} x_{tq_v} + H_t \sum_{\forall i,g} \beta_{gq_v} \theta_{igt} \quad \forall t, q_v \tag{8}$$

3 Decision Support Approach

To employ the framework to aid decision makers, a suite of methodologies are proposed. These have been designed with well-known continuous improvement strategies [13] in mind, from the perspective of strategic, tactical and operational decision makers, as represented in Fig. 1.

Model generation—By leveraging on the reusability and simplified connectivity that the modelling framework promotes, domain experts can help build the model

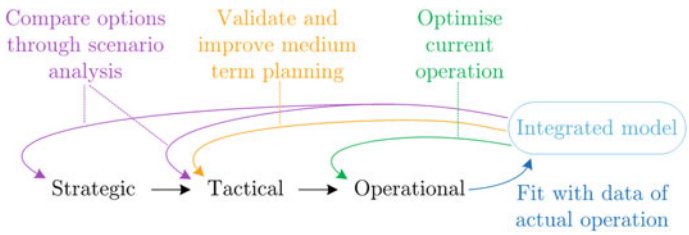


Fig. 1 Decision support methodology

of their production system in a facilitated modelling style [5] without needing to understand the specifics of the mathematical formulations that are employed. This has been designed with the intention of reducing the upfront effort of modelling, engaging end-users early to help them learn how the model works and allowing detailed modelling effort to be focused only where it is required. On top of this, linking the model parameters to existing data systems is recommended. This allows the model to better represent the production system's current operation, helping to improve the solution precision and reduce the burden of entering up-to-date parameter values for end-users.

Operational optimisation—In this methodology the model is solved using current operating conditions to provide operational decision makers with detail about the optimal way to meet planned targets; re-solving when new information is available allows decision makers to see the new, best way to meet the plan. The results of the model can also be used to improve communication between operators, allowing them to see how their subsystem is working as a part of the wider system and see how their decisions impact operators of adjacent shifts.

Plan validation—Since planning decisions, usually made at a tactical level, are input into the model as tasks, the model can be used to allow planners to see how their decisions effect the actual operation of the production system within its current capacity. By integrating the model into planning systems that are currently being used this methodology would help planners make plans that allow for more optimal operation. It can also be used as a tool to help planners of separate subsystems to communicate with each other.

Scenario analysis—This is for aiding both tactical and strategic decision makers, and may also be useful for operational decision makers. By making modifications to the model, re-solving and comparing their impact, decision makers can assess the quantitative impact of options that affect the production system. Tactical options such as asset maintenance and procurement can also be examined. For longer-term decision makers, system layout, operating cost structure or portfolio management scenarios can be assessed. Scenarios can also assess externalities that impact decisions. For example, changing system demand parameters for different market forecasts, or changing constraints to represent new regulations or social responsibilities. If the designed scenarios prove to be valuable and are required periodically the analysis can be easily automated.

4 Examples

4.1 Open-Pit Coal Mine

Model generation—As previously mentioned, the concepts presented here have been developed out of the work contributed in [9]. Taking into consideration that an open-pit coal mining system is regarded as a continuous production system [7, 14], four common subsystems were formulated: the excavation and haulage subsystem

assigns shovels to dig material from pits and trucks transport the material from pits to destinations, the processing plant subsystem converts raw materials into various types of product and waste material, the stockpile stores material to buffer production, and the belt conveyor transports material. By integrating these subsystems a model of a case study mine in Queensland, Australia was created. The model minimises the total energy use of the mine for a given set of shift targets for provided energy efficiency decision support. With only limited access to information and data about the mine's operation several useful analyses were performed using the model, demonstrating the ability for quick payback on modelling effort.

Operational optimisation—A lack of synchronicity between the extraction and coal handling and processing sides of the mine can cause unnecessary double-handling, which has a negative impact on energy efficiency. By using this model, decisions that are made about the workload of individual subsystems can be quickly assessed to see their overall impact on energy efficiency. For example, the demand analysis in [9] showed that increasing production of the extraction subsystem on its own by 40% lead to significant ($\sim$10%) reductions in energy efficiency, while increasing overall mine production by up to 40% was possible without reducing energy efficiency. On the other hand, reducing the overall production only had a minimal negative impact on energy efficiency, which indicates flexibility in the mines current operating point if production is balanced across the subsystems. This is useful knowledge for making cost-justified, operational decisions about increasing production after periods of unexpected lost production, such as weather events, or decreasing production when demand is unexpectedly lowered.

Plan validation—As the mine is operating with plenty of excess capacity, the plans are not tightly constrained and, as such, can vary significantly without impacting the overall production of the mine but can have a large impact on energy efficiency. By using the model to validate planning decisions, more consistent, energy-efficient plans can be generated. For example, the results of the model can be used to ensure the processing plant's production schedule aligns with ROM extraction to minimise wasteful re-handling at the stockpile. The truck analysis in [9] also demonstrated how maintenance planning could use the model to identify equipment that would result in a net gain in energy efficiency if they were taken out to be maintained. Using the model like this is valuable for assessing the impact that production and maintenance decision have on each other.

Scenario analysis—The possibility of aligning the mine's operation with the adjunct power plant (its only customer) is currently being investigated using scenario analysis. As can be seen in the process flow diagram in Fig. 2 of [9], there is a legacy bypass around the wash plant that is not being used as they do not distinguish between grades of raw coal. However, when the power station is at low loads, lower quality coal does not have as much of an impact on its efficiency. It may therefore be an overall improvement to not waste energy washing the coal that is used for low loads. Another option being explored is the potential for either replacing the stockpile between the crusher and wash plant, or adding a bypass around it so that crushed coal can be directly fed into the wash plant, reducing wasteful double-handling.

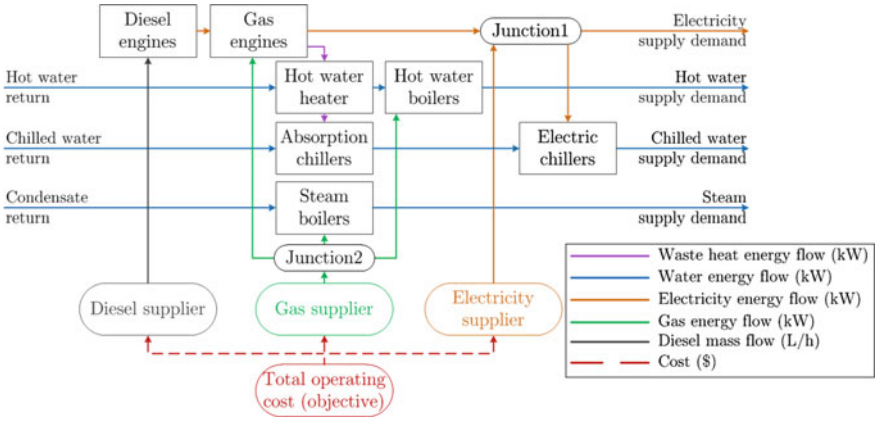


Fig. 2 Building energy plant process flow diagram

4.2 Building Energy Plant

Model generation—Another application currently being explored is a distributed, multi-generation system [2] that provides electricity, hot water, chilled water and steam services to a building in Brisbane, Australia. The development of the model was conducted with relatively limited access to information. Three subsystem models have been formulated to achieve an appropriate representation of the system. The main subsystem is a general 'equipment group' that converts fuel (e.g. electricity or gas) into any of the services. One or several equipment pieces exist in the group, each with an efficiency, that are activated one at a time to meet the service demand. Along with this equipment group subsystem, a general mass balance junction has been created to allow for the splitting and mixing of services. By connecting the various instances of groups and junctions to one another via energy flow connections the overall system can provide the required services. The third, conceptual, subsystem represents the conversion of diesel, gas and grid electricity supply to a cost-for-use in the objective value. The resulting model of the plant is shown in Fig. 2. In this example, the advantage of having generalised subsystem formulations is clear: the general equipment group has been used to represent seven groups with differing functions.

Operational optimisation—By using the integrated model of the plant and the operational decision support approach described in Sect. 3, the model is being recommended as a tool for building managers to control plant operation to meet demand with minimum cost. Further to this, a closed-loop control methodology, whereby the model results are directly passed to equipment controllers (removing the need for building manager oversight), could be used to ensure the lowest cost operation is always being achieved. For example, it is estimated that model-based control to reduce peak demand charges could result in more than $40,000 per year of savings for the building being studied.

Plan validation—Tactical planning of equipment maintenance can be aided by focusing on areas that represent the best overall efficiency return, at a time with minimum impact on operating cost while ensuring the redundancy required for delivering the critical services. For example, the model can be used to plan maintenance of the electric chillers in advance to coincide with times that the absorption chiller is doing most of the work, and vice versa. Planners can also assess the impact of unexpected equipment outages, the potential for selling electricity back to the grid on a curtailable load contract and can examine the impact of expected or unexpected changes in demand.

Scenario analysis—The model is currently being used by the building owners to negotiate better contracts given they currently have excess capacity. So far, this has resulted in identifying savings of between 10 and 20% on consumption cost through increased negotiation power. Several other scenarios can also be explored using the model, including: options for fixed or variable contracts, whether or not to decommission or sell any equipment, the potential of selling services to an adjacent building, and the impact of new technologies. The high resolution, integrated nature of the model means that these scenarios will give decision makers a detailed, system-wide understanding of the options in-front of them.

5 Conclusions

The contributions of this paper have been shown to be useful through two clear examples of distinct production systems. The framework helps reduce modelling effort and enables end-user engagement. The opportunities for using the model for operational, tactical and strategic decision support are significant and promising. The contributions deliver a repeatable process for researchers to apply to existing and new models in an integrated fashion to new problems and deliver valuable results with their application. For industry, the contributions are in a format that enables various decision makers to be engaged in the development of the model as well as benefit from the decision support methodologies.

While the concepts have been applied with success to two processing plants, additional work is required to verify their applicability and value, as well as further generalise their utility. Central to the design of the concepts is their practicality; future work should therefore be based on more applications to new, continuous production systems in different industries. Studies with decision-makers assessing the impact of external factors (e.g. market, regulatory, social) are also suggested.

To overcome potential limitations with using MILP when faced with complex adaptive systems, non-linearities and/or uncertainties, the application of new modelling paradigms and solution techniques within the framework is suggested for future research. For example, alternate mathematical programming structures, such as formations with non-linear or fuzzy constraints, or even other techniques, such as empirical modelling or system dynamics, could help when dealing with

complexities and uncertainties. To migrate the concepts towards having a more concrete use in industry, software is currently being designed.

The decision support suite explained in Sect. 3 is being used to describe the high-level functionality of software tools that deliver the contributions to end-users.

References

1. Baldwin CY, Clark KB (2000) Design rules: the power of modularity. MIT Press, Cambridge
2. Chicco G, Mancarella P (2009) Distributed multi-generation: a comprehensive view. Renew Sustain Energy Rev 13:535–551
3. Childerhouse P, Towill DR (2003) Simplified material flow holds the key to supply chain integration. Omega 31:17–27
4. Elkington T, Durham R (2011) Integrated open pit pushback selection and production capacity optimization. J Min Sci 47:177–190
5. Franco LA, Montibeller G (2010) Facilitated modelling in operational research. Eur J Oper Res 205:489–500
6. Guo Z, Shi L, Chen L, Liang Y (2015) A harmony search-based memetic optimization model for integrated production and transportation scheduling in MTO manufacturing. Omega
7. Jiu S, Zhou Z, Liu J (2013) The equipment maintenance scheduling problem in a coal production system. Int J Prod Res 51:5309–5336
8. Khan M, Jaber MY, Ahmad A-R (2014) An integrated supply chain model with errors in quality inspection and learning in production. Omega 42:16–24
9. Patterson S, Kozan E, Hyland P (2016) An integrated model of an open-pit coal mine: improving energy efficiency decisions. Int J Prod Res 54:4213–4227
10. Prajogo D, Olhager J (2012) Supply chain integration and performance: the effects of long-term relationships, information technology and sharing, and logistics integration. Int J Prod Econ 135:514–522
11. Relvas S, Boschetto Magatão SN, Barbosa-Póvoa APFD, Neves JRF (2013) Integrated scheduling and inventory management of an oil products distribution system. Omega 41:955–968
12. Sarmiento AM, Nagi R (1999) A review of integrated analysis of production-distribution systems. IIE Trans 31:1061–1074
13. Walton M (1988) Deming management method. Penguin, New York
14. Zuñiga R, Wuest T, Thoben K-D (2015) Comparing mining and manufacturing supply chain processes: challenges and requirements. Prod Plann Control 26:81–96

Risk Prioritisation for Cultural and Arts Infrastructure

Andrew Pham, Derren Foster, Christine Soo and Melinda Hodkiewicz

Abstract Asset management is about delivering value to the organisation through its assets. However value is difficult to define for social infrastructure. In a cost-constrained environment challenging decisions have to be made in Cultural and Arts portfolios on which buildings should be prioritised for funding. We present an approach to compare asset performance and risk within one such portfolio. The Western Australian Government's Culture and the Arts Portfolio includes agencies managing approximately 40 social infrastructure buildings such as museums, theatres, libraries and other buildings supporting culture and the arts. To establish an understanding of the assets in the portfolio semi-structured interviews were conducted with asset management experts. Using these results, an on-line survey and the literature, a risk prioritisation tool is developed. The tool was tested by application to three of the Portfolio's buildings. The project identifies seven key criteria for the prioritisation of the Portfolio's buildings. These are: safety and essential services; housing of collection items; condition of building systems; reputation; technical facilities; service delivery; and public exposure. The project rates buildings in terms of importance and vulnerability in each of the seven criteria. Information is presented using a risk prioritisation score; an importance-vulnerability matrix; and spider chart.

A. Pham (✉) · M. Hodkiewicz
School of Mechanical and Chemical Engineering, University of Western Australia, Perth, Australia
e-mail: 21109699@student.uwa.edu.au

M. Hodkiewicz
e-mail: melinda.hodkiewicz@uwa.edu.au

D. Foster
Infrastructure Planning and Support, WA Government Department of Culture and the Arts, Perth, Australia
e-mail: derren.foster@dca.wa.gov.au

C. Soo
Management and Organisations, Business School, University of Western Australia, Perth, Australia
e-mail: derren.foster@dca.wa.gov.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_48

1 Introduction

The Culture and the Arts Portfolio ('the Portfolio') consists of the Art Gallery of Western Australia, the Department of Culture and the Arts (DCA), Perth Theatre Trust, ScreenWest, the State Library of Western Australia and the Western Australian Museum. The mission of the Portfolio is to promote the culture and the arts sector within the community as well as preserve Western Australia's collections. Amongst its six constituent independent agencies the Portfolio is responsible for managing around 40 building assets ('assets') in metropolitan and regional areas. There is ongoing pressure for the Portfolio to improve the management of its assets in order to further the service it provides to the community. This poses a challenge for the Portfolio, as it receives limited government funding, and must find a way to utilise its funding efficiently and effectively.

The Portfolio has identified a number of key opportunities for the improvement of its asset management. One such opportunity is the development of a risk based process for the prioritisation of its assets. There is no widely accepted process for the prioritisation of social infrastructure such as the Portfolio's assets. This prioritisation is necessary in deciding which of the Portfolio's assets require a higher level of funding and attention.

2 Literature Review

While extensive research has been conducted on cost allocation and risk ranking decision support processes in the management of physical assets such as roads, pipelines and plant equipment, there is very little existing literature on the decision support tools for the management of arts and culture infrastructure. In the social infrastructure area research is focussed on areas social infrastructure in the form of hospitals, roads etc., social infrastructure as it relates to poverty or third world, and resilience of social infrastructure.

One of the main challenges in arts and culture infrastructure is the wide range of stakeholders including many different community groups each with their own perceptions of the value of the social infrastructure and/or the services it provides. These perceptions can be difficult to capture and quantify. The absence of data that is quantified and/or comparable across assets means that traditional cost allocation and risk ranking tools cannot be used. Qualitative processes which have been used in these types of situations include Martilla and James' [5] importance-performance analysis (IPA) and multi-criteria analysis (MCA) [2].

MacCrimmon [4] suggests three approaches for generating criteria for MCA. These are: an examination of relevant literature, analytical study and casual empiricism. Casual empiricism is the most relevant approach to this project, whereby a group of knowledgeable experts is observed to understand how their decisions are made. It is recognised that success in MCA depends on the number of

criteria being kept to a reasonable minimum, with the ideal number of criteria depending on the given situation [1]. Keeney and Raiffa [3] suggest that the chosen criteria also need to be a complete set, so that all aspects of the problem are covered. They must be (1) operational, so that the use of the criteria is practical, (2) decomposable, so that the evaluation process can be broken down, (3) non-redundant, so that there is no double counting of impacts, and (4) minimal, so that the problem is kept as small as possible.

A casual empiricism approach is used for this project involving the consultation of stakeholders and experts within the DCA Portfolio, while bearing in mind the five requirements of the criteria suggested by Keeney and Raiffa.

3 Process

The process by which the research is conducted is summarised in Fig. 1. Because the research involves human participants, ethics approval was sought and obtained. Data was collected through 11 semi-structured interviews with key personnel both internal and external to the Portfolio, ranging from 40 min to 80 min in length. Interviewees external to the Portfolio included experts from other organisations assisting with the management of the Portfolio's assets, such as the Department of Finance Building Management and Works, as well as asset management experts from around the world, such as the New South Wales Department of Treasury and Finance. Interviewees internal to the Portfolio included personnel involved in asset management from five of the six Portfolio agencies. A mix of internal and external interviewees was chosen in order to gain insight from a range of perspectives.

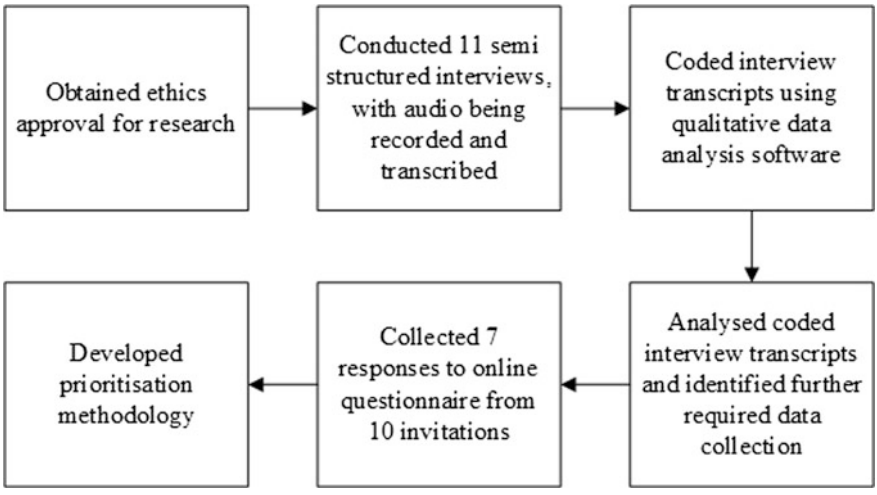


Fig. 1 Process map summarising the tasks conducted throughout the research

In four of the five Portfolio agency interviews, the interviewees included both the facilities manager of the agency as well as an employee in a strategic position such as a general manager or director. The purpose of concurrently interviewing these two roles was to collect both a strategic and operational response to the interview questions.

11 h of interview audio was recorded. Once the interviews had been transcribed they were coded using NVivo, a qualitative data analysis software package. Using NVivo, sections within the interview transcripts were coded according to the discussion topic. This allowed 29 potential prioritisation criteria to be identified.

An initial analysis of the interview results was conducted. The least mentioned criteria were discarded. Portfolio agency interviewees were then invited to complete an anonymous online questionnaire. From the 10 Portfolio agency personnel invited to complete the questionnaire, seven completed the questionnaire. The questionnaire allowed the respondents to rate the importance of the 16 remaining criteria from the interviews. This was done on a five point Likert scale, with a score of five allocated to criteria perceived to be extremely important and a score of one allocated to criteria perceived to be not at all important. The respondents' ratings of the criteria generated a score for each criterion, allowing them to be ranked by their score. All criteria scoring higher than a four (very important) were then used in the prioritisation process.

In order to develop the prioritisation process, a review of the literature as well as current and past DCA prioritisation techniques was conducted. Numerous consultations with the DCA were also held in order to gain an understanding of the client's need. A strong focus on simplicity, usability, transparency and understandability was expressed by the DCA. With input from both previous literature and DCA consultations, a process was developed. The key prioritisation criteria identified through the online questionnaire was then used as the criteria in the process.

4 Results and Discussion

4.1 Key Criteria

29 criteria for the prioritisation of an asset were identified from the 11 interviews conducted. From an analysis of the interview transcripts in NVivo, it was determined which of these criteria were mentioned most frequently in the interviews and were discussed in the most interviews. Table 1 shows both of these results for a random sample of the identified criteria.

From the online questionnaire following the interviews, the criteria rated by the participants on average as having a greater importance than "very important" (average rating of four) were: safety and essential services; housing of collection items; condition of building systems; reputation; technical facilities; service

Table 1 The number of times each criterion was mentioned and number of interviews the criterion was mentioned in

Criteria	Mentions	Interviews
Age	5	4
Environment	1	1
Heritage	4	3
Mechanical systems	12	5
Public exposure	13	8
Reputation	15	7
Stakeholders	7	7

Figures for only a random sample of the total mentioned criteria are displayed

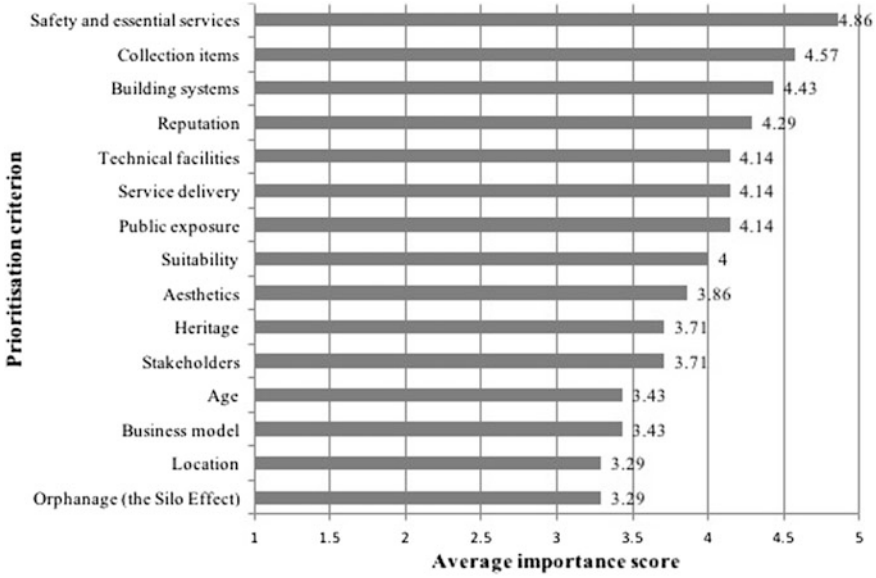


Fig. 2 Average ratings of each of the 16 criteria in the online questionnaire

delivery; and public exposure. The average ratings of each of the 16 criteria are shown in Fig. 2. Figure 2 shows seven criteria clearly rated above four and perceived by the respondents as having a greater importance than "very important".

4.2 Prioritisation Process

From the literature review, it was found that the decision technique most appropriate for the prioritisation process was a modified version of Martilla and James' [5] IPA combined with an MCA approach. These two techniques were deemed as

being the most practical in providing the DCA with a sufficiently sound prioritisation process.

A MCA approach to prioritising a given asset is taken by rating the importance and vulnerability of each criterion for the asset. The importance is defined as the extent to which the criterion is related to the mission of the Portfolio. The vulnerability is the extent to which the asset is performing poorly in the given criterion. For example, a performing arts venue would rate low on importance for "collection items" if it doesn't house any collection items. By rating an asset in each criterion in these two dimensions, priority is given to assets that are both essential to the Portfolio as well as in poor condition, as opposed to assets that are essential but in good condition, or any other combination of the two factors.

The ratings are processed and visualised in three different ways to provide insight on the level of prioritisation of the asset. The three ways are: a prioritisation score; an importance-vulnerability matrix (Fig. 3); and a spider plot (Fig. 4). Figures 3 and 4 have taken example data for illustrative purposes. The importance-vulnerability matrix can also be shown for each of the different importance criteria rather than as a combined score.

A prioritisation score for the asset is generated by summing the products of each criterion's importance and vulnerability matrix. For example, a given asset that rates four in the importance of "reputation" and three in the vulnerability of "reputation" is assigned a criterion score of 12 for reputation. The criterion scores are then summed to generate the prioritisation score for the asset.

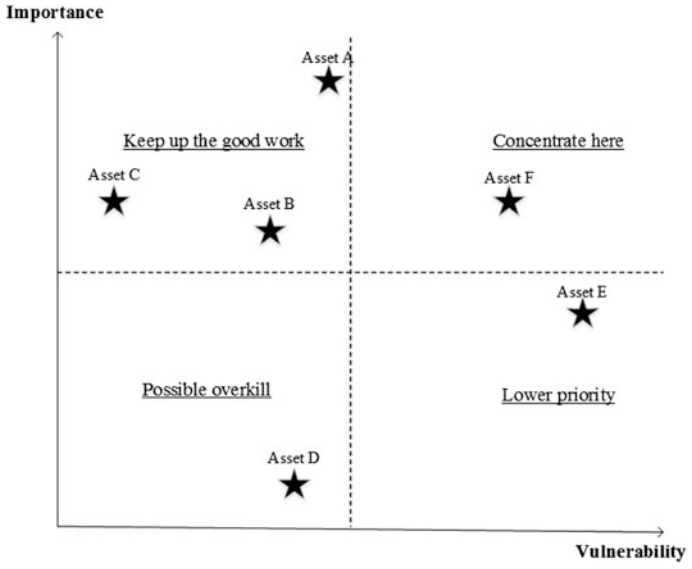


Fig. 3 An importance-vulnerability matrix providing a visual representation of the relative importance and vulnerability to failure of different assets

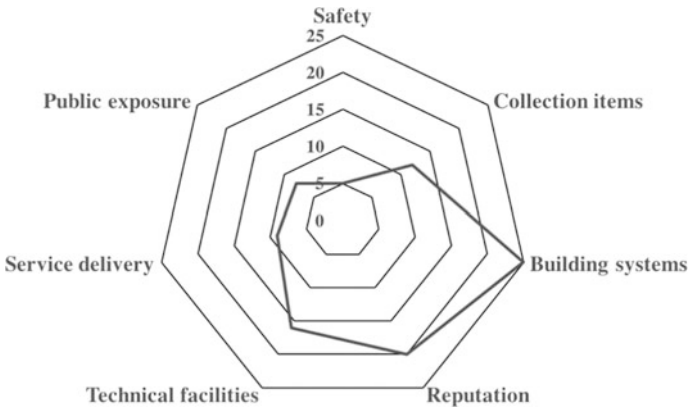


Fig. 4 A spider chart showing an asset's score for each of the criteria

The importance-vulnerability matrix shown in Fig. 3 allows for a comparison of the Portfolio's assets in terms of their relative importance and vulnerability to failure. Based on Martilla and James' [5] importance-performance matrix, it is split into four quadrants with each describing the asset management of the assets within it. This importance-performance matrix has been modified by inverting the horizontal axis, placing lower performing assets to the right of the graph as opposed to on the left as originally suggested by Martilla and James. This modification has been made in order to generate a higher prioritisation score for low performing assets.

In order to visualise the level of prioritisation of each asset in greater detail, a spider chart can be used as shown in Fig. 4. The spider chart allows identification of the criteria to be prioritised within an asset. Each of the asset's criterion scores are plotted, with the criterion score being the product of the criteria's importance rating and likelihood of failure rating. For example, Fig. 4 illustrates an example where the asset requires the most attention to its building systems and reputation.

5 Conclusions and Future Work

This study has identified seven key criteria to be used for the prioritisation of Portfolio's assets. From these key criteria, a process has been developed to provide the Portfolio insight into how funding can be allocated most effectively. This process involves rating the importance and vulnerability of each criteria for a given asset. The ratings can be used in three key tools to assist with funding allocation: a prioritisation score; an importance-vulnerability matrix; and a spider chart.

There is potential for the process to be developed further. No criteria weightings have been assigned in the process in order to maintain the simplicity of the approach. The inclusion of weightings is an opportunity for future work. The

process has been developed primarily in consultation with the DCA. There is an opportunity to gain feedback from the other Portfolio agencies on the process through a workshop, using the feedback to test, modify and improve the process. While a subjective element still remains in the prioritisation process from the rating process, there is a potential to eliminate this by using numerical inputs such as performance indicators in the prioritisation process.

Acknowledgements In addition to the academic supervisors and client mentor, the author would like to thank Carl Pekin and Michelle Nicholson from the Infrastructure Planning and Support team at the DCA for their assistance throughout the project. The author also acknowledges the research participants for their contribution to the project.

References

1. Blanchard BS, Fabrycky WJ (2006) Systems engineering and analysis, 4th edn. Pearson Prentice Hall, New Jersey
2. Greco S (2005) Multiple criteria decision analysis: state of the art surveys. Springer, New York
3. Keeney RL, Raiffa H (1976) Decisions with multiple objectives: preferences and value tradeoffs. Wiley, New York
4. MacCrimmon KR (1969) Improving the system design and evaluation process by the use of trade-off information: an application to northeast corridor transportation planning. The Rand Corporation, Santa Monica
5. Martilla JA, James JC (1977) Importance-performance analysis. J Mark 41(1):77–79

Analysing an Industrial Safety Process Through Process Mining: A Case Study

Anastasiia Pika, Arthur H. M. ter Hofstede, Robert K. Perrons, Georg Grossmann, Markus Stumptner and Jim Cooley

Abstract The application of data analytics has delivered significant value in a broad range of industries, but has frequently failed to bring about operational efficiencies or process safety improvements in asset-intensive sectors despite a profound increase in the volume of digital information that is stored by companies operating in this domain. Process safety management is a major concern of asset-intensive organisations, as the consequences of incidents can be catastrophic. In order to prevent incidents from occurring, organisations enforce prescriptive safety procedures frequently known as "Permit to Work" systems. These procedures can be thought of as a workflow consisting of a set of well-defined steps that have to be performed in a certain order to minimise safety risks. In this paper, we apply analytical tools from the area of process mining to shed light on how safety processes are actually executed. Process mining constitutes a relatively recent area of research at the intersection of data analytics and business process management (BPM) that is concerned with extracting insights from event logs that record process executions. We report the results of a process mining case study based on a Permit to Work system in an asset-intensive organisation in Australia.

A. Pika (✉) · A. H. M. ter Hofstede · R. K. Perrons
Queensland University of Technology, Brisbane, Australia
e-mail: a.pika@qut.edu.au

A. H. M. ter Hofstede
e-mail: a.terhofstede@qut.edu.au

R. K. Perrons
e-mail: robert.perrons@qut.edu.au

G. Grossmann · M. Stumptner
University of South Australia, Adelaide, Australia
e-mail: georg.grossmann@unisa.edu.au

M. Stumptner
e-mail: markus.stumptner@unisa.edu.au

J. Cooley
Origin Energy, Brisbane, Australia
e-mail: jim.cooley@originenergy.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_49

1 Introduction

Organisations in asset-intensive domains invest a lot of resources in the management of safety risks [9, 12], as the cost of safety incidents in these domains can be prohibitive. As a part of health and safety management strategies, organisations design safety processes, e.g. Permit to Work systems [2]. Such processes consist of steps which have to be followed by employees in a certain order to minimise safety risks. Safety processes are often complex and understanding their operational mechanics is crucial for risk identification and process improvement. Large volumes of data capturing the actual conduct of these processes are recorded; however, it is often a challenge to get actionable insights from this data [9, 12].

In this paper we propose the application of process mining to derive insights from the execution of safety processes. Process mining is concerned with the analysis of process execution data [15] and provides, among others, methods for discovering how processes are actually executed (process discovery), where deviations from expected process flows occur (process conformance) as well as techniques for process performance analysis and process comparison. Process mining algorithms take as an input an event log which contains information about process executions. Most process mining algorithms require that at least process instance identifiers, process activities and their timestamps are captured in the event log. Event logs may contain other information, for example whether an activity was started or completed, or information related to cost or location. Some information can be domain-specific, e.g. the requested loan amount for a mortgage application or the medicare number of a patient. Safety processes involve the conduct of activities in a certain order and when they are supported by information systems this history of work is typically recorded in an event log which can be analysed through the application of process mining techniques.

We present the results of a case study in which process mining was used to analyse the Permit to Work system in an Australian asset-intensive organisation. The Permit to Work system forms an integral part of the company's "Life Saving Rules", a set of clearly articulated requirements that all employees must follow to mitigate risks. Process compliance is considered to be critical in the company, and instances of non-compliance are treated as serious incidents in their own right. Our initial goal was general in nature, namely to investigate what insights about the process could be obtained through the application of process mining. We conducted data collection, interpretation, and analysis iteratively in collaboration with the process experts. During this engagement specific questions about the Permit to Work system emerged concerning process conformance, performance and evolution. These questions guided our subsequent analyses. The case study uncovered the process complexity, identified parts of the process where deviations from expected process flows occur as well as factors affecting process performance. Our findings highlighted the need for process simplification which could potentially make it easier to communicate the process, thus reducing mistakes in its conduct.

In the following sections, we first discuss related work, position our work and highlight our contributions (Sect. 2). We then describe the case study design, methods and findings (Sect. 3) as well as directions for future work (Sect. 4).

2 Related Work

Organisations have been using occupational health and safety management systems (OHS MSs) since the mid-eighties to minimise the number of health and safety (H&S) incidents [9]. An OHS MS is "a set of institutionalized, interrelated, and interacting strategic H&S management practices designed to establish and achieve occupational safety and health goals and objectives" [19]. Permit to Work systems are often a part of OHS MSs [2]. To measure and evaluate different aspects of OHS MSs companies use indicators: *lagging indicators* are used for performance measurement, *structural performance indicators* are used to evaluate "system compliance with a given specification form", and *operational performance indicators* are used to evaluate "effectiveness of internal system processes" [9]. Structural and operational performance indicators are also referred to as *leading indicators* [9]. Another purpose of leading indicators is "to flag potential problems" and "help uncover weaknesses in the organization's procedures or employee behaviour" [12]. The numbers of leading indicators tracked by companies can be very high (up to a few hundred) resulting in large volumes of data being recorded [9]. It is often a challenge to get insights from such data [9] and "turning data into action is a struggle for many organizations" [12].

Despite the challenges organisations face trying to make sense of safety-related data, Tan et al. [14] argue that they "may be in a position to make considerable progress by applying 'Big Data' analytical tools" in the oil and gas industry; and Podgórski [9] argues that companies should make better use of operational performance indicators to get "a picture of how processes operate at the shop-floor level". *However, it is not shown how these ideas can be operationalised.*

H&S incidents are often linked to a poor safety culture [6] which is often evaluated using questionnaires [2, 9]. However, this approach has been criticised because employees' answers can be affected by "social expectations about their behaviour" and frequent involvement of all employees in such questionnaires is costly [9].

Several applications of Big Data for risk mitigation were reported by Tan et al. [14]: the use of hands-free checklists by workers assembling equipment in the oil and gas industry to reduce the chance of making mistakes; the use of built-in sensors on assets to monitor manufacturing assembly operations and reduce operational risks; and the use of Big Data to detect risks in supply chains.

We, on the other hand, propose the use of process mining to gain insights into the execution and performance of H&S processes. Process mining is a research area at the intersection of data analytics and BPM and is concerned with extracting insights from event logs that result from the conduct of business processes [15].

Risk-aware BPM provides methods for dealing with risks in processes [13] including methods for mining process risks from event logs, e.g., a recommendation system that supports risk-informed decisions [1] and a method that helps to evaluate and predict overall process risk [8]. *Process mining techniques helped to gain insights into process efficiency in organisations across different industries; however, to the best of our knowledge, they have not yet been applied to safety processes.*

3 Case Study

3.1 Analysis Approach

The case study was conducted in an Australian energy company. We analysed the company's Permit to Work system, a safety process consisting of activities which must be performed by employees to minimise work hazards. The process is an important part of the company's risk management strategy. Failure to follow the process may pose significant health and safety risks.

Data collection, interpretation, and analysis were performed iteratively in collaboration with some company personnel with knowledge of both the process and associated information. We refer to these personnel as process experts. At the beginning of the case study we did not focus on any specific questions; our goal was to learn what insights about the process we can uncover through the application of process mining. We refer to this part of the case study as *Stage I* and discuss it in Sect. 3.2. After we presented our initial findings and demonstrated different process mining capabilities to the process stakeholders, they were excited to learn more about specific aspects of the process. Our subsequent analyses were guided by their questions. We refer to this part of the case study as *Stage II* and discuss it in Sect. 3.3.

The process is supported by different information systems which record the process execution data. We worked with the process experts to collect this data from different sources, interpret it, and create an event log suitable for process mining algorithms (in XES[1] format). The resulting event log contained 43,061 cases (i.e. process instances) which were started and completed during the period of 28/03/2008–8/12/2016. The data was analysed using the open source process mining framework ProM[2] and the process analysis tool Disco.[3]

[1] www.xes-standard.org.

[2] www.promtools.org.

[3] https://fluxicon.com/disco/.

3.2 Stage I

Goals and Methods

At the beginning of the case study the company did not have any specific questions and was interested to see what they could learn about the Permit to Work process by analysing its data. During this stage we applied a range of process mining techniques to shed light on different aspects of the process. To learn how the process is executed in practice, we used the Disco miner, a process discovery tool which mines a fuzzy model from an event log [4]. Process conformance analysis was used to identify where and how often actual process behaviour deviates from expected process flows. We applied an alignment-based process conformance technique [16] which uses as input an event log and a normative process model (i.e. a model which describes expected process behaviour). The technique creates the so-called alignments by relating events in the log to elements of the model and highlights discrepancies between the prescribed and the observed process behaviour. We first created a normative process model based on the process documentation which was subsequently updated based on input from the process experts. An alignment-based performance analysis approach [16] was used to identify parts of the process with the longest execution times. To check whether process behaviour and performance is changing, we used an extensible framework [7] which allows us to define measures of resource and process behaviour and to track their evolution over time.

Findings

Process discovery revealed that the Permit to Work system is a complex process with more than 2,300 process variants. Figure 1 demonstrates the discovered process model which represents all observed process behaviour and highlights the complexity of the process. A company representative was impressed by the insights revealed by the process discovery result. He noted that some process participants had often cited the process complexity as a leading cause of lost productivity and inadvertent process non-compliance, and that the discovered model provides evidence for the need to simplify the process.

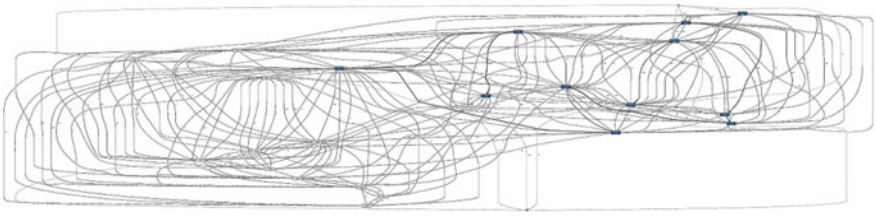


Fig. 1 Discovered process model of the Permit to Work process

Conformance checking identified those parts of the process where deviations from expected process flows occur. To further interpret the deviations, i.e. to learn whether they are related to skipped activities or activities executed in a wrong order, we checked those parts of the process where deviations occur using Disco. We discovered that many non-conformances occur at the beginning of the process and are related to risk assessment and approvals, for example, some approvals are skipped or in some cases authorisations are performed after a permit to work is withdrawn. We also learned that the frequency of some non-conformances is changing over time. Figure 2 shows an example of a process conformance issue (represented in terms of a fuzzy model) and its evolution over time. We can observe that in most cases (40,479) risk assessment was authorised (task *PTW_RA_Authorised*) after it was created (task *PTW_RA_Created*) which is expected process behaviour, while in 1,048 cases risk assessment was authorised before it was created which is unexpected. We can also observe that the frequency of such deviations is decreasing over time. The process experts could not explain this abnormal behaviour and further investigation is required to identify its causes.

Process performance analysis helped us to identify parts of the process and process variants with long execution times. For example, we learned that cases in which a permit was re-issued took on average 61.6 days while the mean duration of other cases was 33.8 days (mean durations of different variants of the process were checked using Disco). Similarly, cases in which testing was performed had longer execution times (mean duration 69.1 days) compared to other cases (mean duration 36 days). By checking process evolution, we learned that the process is getting more efficient over time. To get further insights into process performance, the process experts asked us to look at different types of work; we discuss these findings in Sect. 3.3.

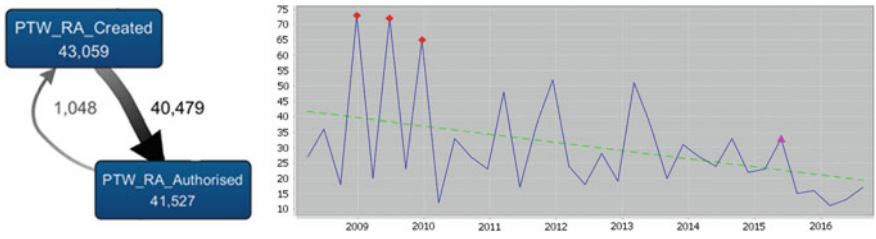


Fig. 2 Example of a process conformance issue and its evolution over time: risk assessment was authorised before it was created

3.3 Stage II

Goals and Methods

During the second stage of the case study we focused on specific areas of interest to the process stakeholders which emerged from Stage I. The stakeholders were interested to learn more about conformance issues of the process and their evolution over time. They asked us to analyse specific process behaviours related to Key Performance Indicators (KPIs) defined by the company. They were also interested to learn whether there are any differences in the execution and performance for different types of work. There is an attribute in the process execution data capturing type of work and the process experts grouped its values into four cohorts: *major outages, major forced outages, short-term maintenance* and *medium-term maintenance*. As per the company's request, during the second stage of the case study we analysed more recent data (cases started and completed during the period of 1/1/2011–8/12/2016).

In addition to process mining techniques described in Sect. 3.2, during the second stage of the case study we applied the ProM plug-in "Process Profiler 3D"[4] [18]. This plug-in takes as input an event log and a process model and visualises differences in the execution and performance of different process cohorts. We used it to compare the aforementioned different types of work.

Findings

The process experts provided us with a list of process-related KPIs they were interested in. Here we show examples of findings for three KPIs: KPI 1 "Request for access must be approved seven days before the scheduled start date"; KPI 2 "Permit to work must be issued before the scheduled start date"; and KPI 3 "Permit to work must be withdrawn before the due date". Scheduled start date and due date are not process steps but they are recorded in the event log. The presence of this information allowed us to apply the same approach for checking process evolution [7] we used during the first stage of the case study. Figure 3 shows that the percentages of cases with KPI violations are different for the three KPIs: around 30% for KPI 1, 70% for KPI 2 and 50% for KPI 3 (on average across the whole period). Figure 3 also demonstrates that the level of violations does not change significantly for KPIs 1 and 2, but it is increasing for KPI 3. Overall, the number of non-conformances with respect to these KPIs is high; causes of this behaviour are yet to be investigated.

Process comparison identified differences in the execution and performance of the four types of work. For example, we learned that permit suspension is never repeated for *major forced outages*, but it can be repeated for other types of work. We discovered that there are significant differences in performance for the four

[4] https://svn.win.tue.nl/repos/prom/Packages/ProcessProfiler3D.

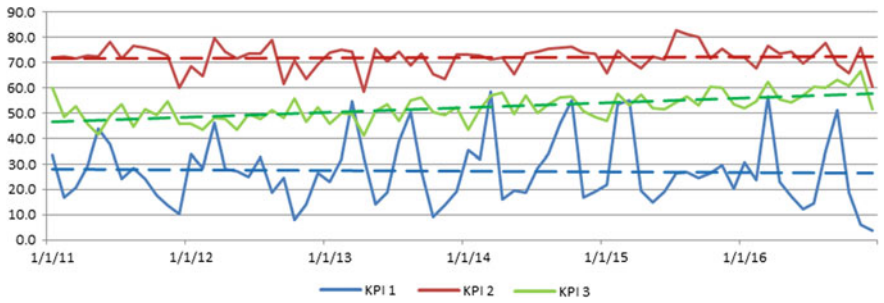


Fig. 3 Percentage of cases started per month in which a KPI was violated

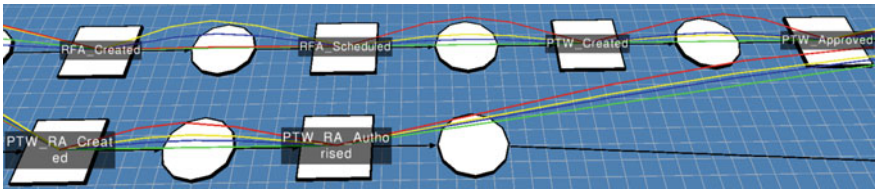


Fig. 4 Mean waiting times for the four cohorts: major outages (red), major forced outages (green), short-term maintenance (blue) and medium-term maintenance (yellow)

cohorts, with mean case durations ranging from 7.5 days for *major forced outages* to 58.1 days for *major outages* (mean case durations were checked using Disco). Execution times also vary for the four cohorts in different parts of the process. Figure 4 shows mean waiting times for the four cohorts in a part of the process (arc colours represent different cohorts and arc heights reflect mean waiting times). While mean waiting time is the highest for *major outages* in most parts of the process (red arcs), this is not the case for mean waiting time between tasks *RFA_Created* and *RFA_Scheduled* which is the longest (7.7 days) for *medium-term maintenance* (yellow arc).

4 Lessons Learned and Future Work

The application of process mining techniques allowed us to uncover the complexity of the permit to work process, identify those parts of the process where deviations from expected process behaviour occur as well as process variants with the longest execution times, discover differences in the execution and performance for different types of work, and learn that some process behaviour is changing over time. The company is considering an improvement initiative for the process aimed at

simplifying it. This could potentially lead to a decrease in process conformance issues and to an improvement of process safety in the company.

The case study uncovered some data quality issues; for example, some process-related data is not recorded (e.g. activity start times) or is partially recorded (e.g. information about employees involved in the process). Accurate capture of all process execution data would enable a more elaborate analysis, e.g. we could consider activity durations or analyse behaviour of employees handling the process.

We also identified some shortcomings of alignment-based process conformance and performance analysis techniques. The conformance checking technique used can identify deviations from expected process behaviour, however further investigation is required to interpret these deviations. The performance analysis technique used can visualise execution times of different process parts; it could be enhanced by visualising performance of different process execution paths.

A step forward in the evolution of process safety is transforming the knowledge gained through the application of process mining into actions. Process mining helped us to identify some process conformance issues; however, we did not investigate the causes of these issues and ways to overcome them. As mentioned earlier, process complexity could be a factor affecting both process performance and conformance, and the company is considering a simplification initiative for the process. Lack of proper process communication or inconsistencies in process representations could be other possible sources of conformance issues. To improve process communication, it is necessary to capture the real world processes and depict how they relate to stakeholders and the affected artefacts, e.g. in the form of lifecycles [11]. Inconsistencies in process representations can be detected [3] and corrected, e.g. with the help of process model repair [10]. Techniques for process configuration [5, 17] can be used to support consistent representation of multiple process versions.

5 Conclusions

Safety processes aim to minimise safety hazards in asset-intensive organisations. A lot of data is recorded about the conduct of these processes, and whilst this data is extensively used to produce KPIs to measure business performance, it remains a challenge to obtain actionable insights about the processes themselves from this data. In this paper we reported the results of a case study in which process mining was applied to analyse process execution data of a Permit to Work system in an Australian energy company. The case study revealed different types of process insights, thus providing an initial indication of the potential of the application of process mining to safety processes in asset-intensive organisations. Specifically, the case study results highlight the complexity of the process and areas of poor process performance, providing evidence for the company for the need to simplify the process with a view to improving its reliability and efficiency.

References

1. Conforti R, de Leoni M, La Rosa M, van der Aalst WMP, ter Hofstede AHM (2015) A recommendation system for predicting risks across multiple business process instances. Decis Support Syst 69:1–19
2. Flin R, Mearns K, O'Connor P, Bryden R (2000) Measuring safety climate: identifying the common features. Saf Sci 34(1):177–192
3. Grossmann G, Mafazi S, Mayer W, Schrefl M, Stumptner M (2015) Change propagation and conflict resolution for the co-evolution of business processes. Int J Cooperative Inf Syst 24(1)
4. Günther CW, Rozinat A (2012) Disco: discover your processes. In Demonstration track of the 10th international conference on business process management, pp 40–44. http://ceur-ws.org/Vol-940/paper8.pdf
5. Mayer W, Stumptner M, Killisperger P, Grossmann G (2011) A declarative framework for work process configuration. Artif Intell Eng Des Anal Manuf 25(2):143–162
6. Nordlöf H, Wiitavaara B, Winblad U, Wijk K, Westerling R (2015) Safety culture and reasons for risk-taking at a large steel-manufacturing company: investigating the worker perspective. Saf Sci 73:126–135
7. Pika A, Leyer M, Wynn MT, Fidge CJ, ter Hofstede AHM, van der Aalst WMP (2017) Mining resource profiles from event logs. ACM Trans Manage Inf Syst 8(1) (article 1)
8. Pika A, van der Aalst WMP, Wynn MT, Fidge CJ, ter Hofstede AHM (2016) Evaluating and predicting overall process risk using event logs. Inf Sci 352:98–120
9. Podgórski D (2015) Measuring operational performance of OSH management system—a demonstration of AHP-based selection of leading key performance indicators. Saf Sci 73:146–166
10. Polyvyanyy A, van der Aalst WMP, ter Hofstede AHM, Wynn MT (2016) Impact-driven process model repair. ACM Trans Softw Eng Methodol 25(4):28
11. Schrefl M, Stumptner M (2002) Behavior-consistent specialization of object life cycles. ACM Trans Softw Eng Methodol 11(1):92–148
12. Sinelnikov S, Inouye J, Kerper S (2015) Using leading indicators to measure occupational health and safety performance. Saf Sci 72:240–248
13. Suriadi S, Weiss B, Winkelmann A, ter Hofstede AHM, Adams M, Conforti R, Fidge C, La Rosa M, Ouyang C, Pika A, Rosemann M, Wynn M (2014) Current research in risk-aware business process management—overview, comparison, and gap analysis. Commun Assoc Inf Syst 34 (article 52)
14. Tan KH, Ortiz-Gallardo VG, Perrons RK (2016) Using Big Data to manage safety-related risk in the upstream oil & gas industry: a research agenda. Energy Explor Exploit 34(2):282–289
15. van der Aalst WMP (2016) Process mining: data science in action. Springer, Berlin. http://www.springer.com/978-3-662-49850-7
16. van der Aalst WMP, Adriansyah A, van Dongen B (2012) Replaying history on process models for conformance checking and performance analysis. Wiley Interdisc Rev Data Min Knowl Discovery 2(2):182–192
17. van der Aalst WMP, Dumas M, Gottschalk F, ter Hofstede AHM, La Rosa M, Mendling J (2010) Preserving correctness during business process model configuration. Formal Asp Comput 22(3–4):459–482
18. Wynn MT, Poppe E, Xu J, ter Hofstede AHM, Brown R, Pini A, van der Aalst WMP (2017) ProcessProfiler3D: a visualisation framework for log-based process performance comparison. Decis Support Syst 100:93–108
19. Yorio PL, Willmer DR, Moore SM (2015) Health and safety management systems through a multilevel and strategic management perspective: theoretical and empirical considerations. Saf Sci 72:221–228

Enterprise Risk Profiling Using Asset Transaction History and Condition Measurements

R. A. Platfoot

Abstract Modern asset information systems collect significant quantities of variable-quality data covering operational performance work and financial transactions, as well as measurement history. This paper presents methods to utilise asset configuration data and transaction data such as work history and to provide insight as to where risk of asset failure is currently peaking across an extensive asset portfolio. By ranking the likely potential for failure based on undesirable condition results which are reasonably current, it is possible to generate detailed watch lists to consider and then escalate decisions on the appropriate timing for cost effective intervention. The methods of reporting not only need to manage multi-variate drivers of risk but also the means to effectively communicate where intervention should be focused across the portfolio.

1 Introduction

There is an increasing need to effectively communicate the true state of risk in a complex asset portfolio, [3]. The work reported in this paper considers a mix of logical and probabilistic risk models [7], where models are applied as a quantification of the current state to allow comparison and identification of key risk issues, [2]. Probabilistic models can also provide a current state representation and they are used to manage unstructured data where there is no other guidance as to the severity of risk to be determined from the data magnitudes [1].

This work addressed value adding large quantities of asset information data drawn from standard Enterprise Asset Management (EAM) systems tracking work on the assets, costs and measurements undertaken as part of routine maintenance [6]. In the management of large asset portfolios which can number more than 50–100,000 assets, new methods are needed to locate and then report risks which are both current and potential. The reporting aspect is crucial since asset managers must

R. A. Platfoot (✉)
Covaris Pty Ltd., Sydney, Australia
e-mail: r.platfoot@covaris.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_50

manage their investment portfolios, answer Board-level challenges and need intelligence which makes sense given their own level of technical expertise and background.

A key concern is the management of data quality associated with asset information systems [4]. The concerns include managing extensive data sets, managing poor quality data (e.g. nulls, misallocated, unallocated etc.), and providing executive reports which precisely support decision making. Hence robust systems for data management and reporting are called for, representing a break from traditional statistical graphical reporting. This is consistent with modern research into the optimisation of cognitive-focused reporting dashboards [8].

2 Nature of Asset Information

In this work asset information has been classified into three classes as shown below. This choice of classification is based on the type of information system in which the data is collected plus the classic business function (e.g. operations, maintenance, finance) which is responsible for the management and use of that system (Table 1).

In this paper, we are considering data from the first two classes. The management of this data can be complex and an example of a generic but widely applicable data structure is shown in Fig. 1.

Master data is utilised to register assets in the system along with how their health will be assessed using measurements. Given the need to manage thousands of possible measurement points which control millions of measurements, this is a complex and precise exercise. Data entities which capture asset information include

Table 1 Classification of asset information

Classification	Examples	Nature
Incident	Logged downtime in operation	Low frequency time stamped May be associated with a system or specific asset All data has significance Possible tolerance on outer bounds
Transaction	Work order Condition monitoring reading Inspection result Purchase order Capital project registration	Low frequency time stamped Usually allocated to an asset or system Complex information associated with the transaction All data has significance Measures of effort in terms of cost and labour
Real time	SCADA output Relevant to the internet of things (IoT)	High frequency time stamped Large volumes of data of low significance Possible tolerance on outer bounds

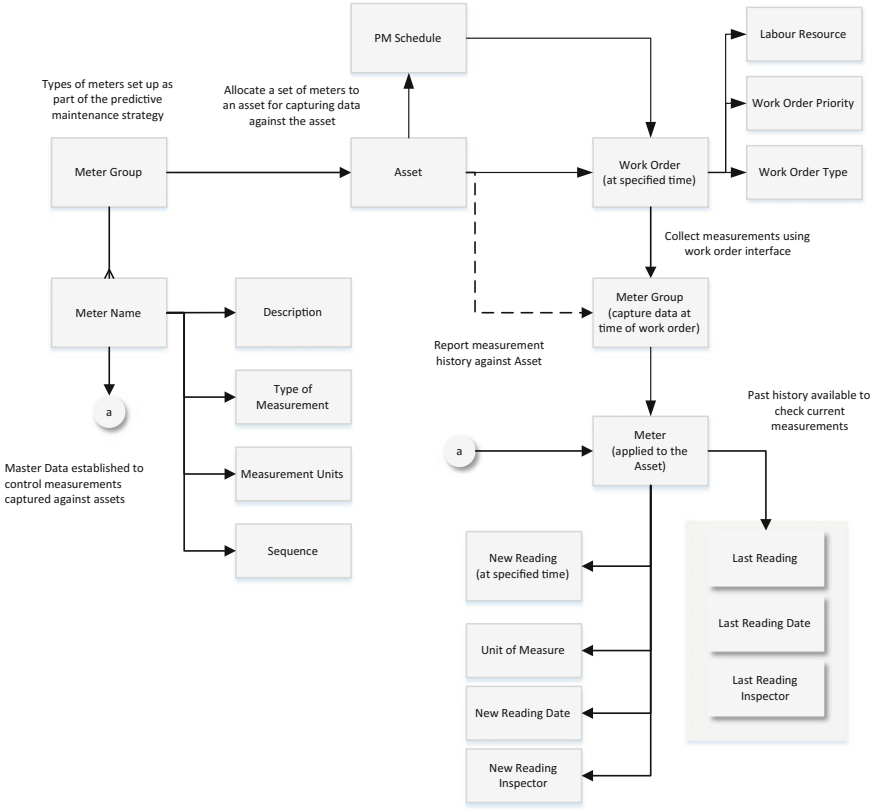


Fig. 1 Managing asset information data

the Work Order and the New Reading, both highlighted in the figure. Supporting data such as the unit of measurement of the sensor, inspector and so on supplement the asset-relevant information.

3 Risk Assessment—Work Transactions

Risk-based reports may be used to locate assets of concern in large portfolios, utilising the information described in Sect. 2 regarding work order transactions. These reports may be integrated using a range of risk metrics in a balanced scorecard approach such as suggested by Mkandawire, Ijumba and Whitehead [5].

Risk metrics which are possible include Weibull forecast of corrective maintenance with probabilities of work trending up, and high priority work (or notifications which are yet to be resolved). The second set represent risk since the rate of work requires high levels of urgent response. The reporting format for each of these

risk profiles uses spider charts based on risk ranking of specific assets, where 100% represents maximum risk.

The example below shows a drill down approach from a top-level system: in this case, a cooling system within a complex industrial facility. In accordance with good practice recommended for dashboards [8], a drill-down capability allowed the most significant system to be further examined. Along with other assets, the cooling towers were calculated to have a high risk based on relative frequency of high priority work. Hence, they were assets of interest to be further considered. Their key problems are shown in the second graphic to be the fans and their variable speed drives.

In these diagrams the assets are ranked by relative frequency of high priority work, so that worst risk example is an asset (or group of assets) which have a relatively high population of urgent work indicating advance deterioration. The risk and levels are normalised using the maximum risk reported on each diagram, so that worst case is as 12 o'clock and is 100%. Working clockwise around the diagram, the risk ranking will drop away of assets have lower levels of urgent work. Hence the results reflect the relative risk ascertained for each of the assets in the diagram, allowing it to be used as a searching tool to find assets of interest (Fig. 2).

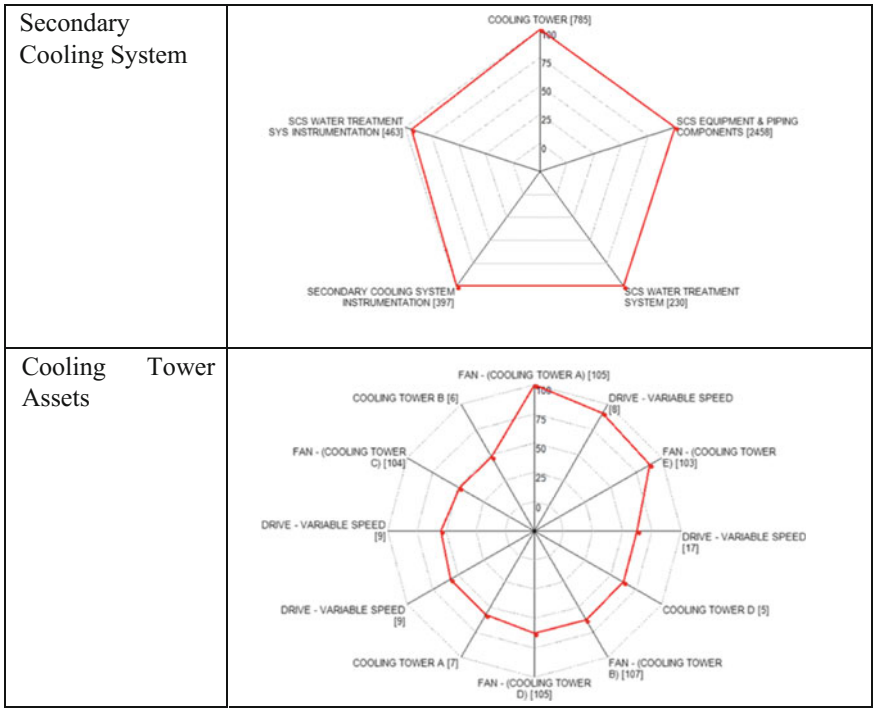


Fig. 2 Risk profiling based on high priority work

The statistic printed against each asset name is the number of corrective maintenance work orders. This allows the reader to not only test for which assets have comparatively high rates of high priority work, but also the numbers of repairs which these assets have required. When the repair rate is high, the asset may be judged to be of additional concern, beyond the ranking organised by the work prioritisation statistic.

Charts of this kind are highly scalable across the thousands of assets which make up a complex industrial facility. Summary reports (as simplified below) covering these assets can be generated from the analytic system to supply the asset management plan of the overall facility with credible risk profiles. Based on these profiles, further verification can lead to the proposal of effective work, which would be prudent to undertake (Table 2).

The reference to Fan A mentioned above was picked up in the top 10 list for the entire facility. Before any other investment is committed, consideration for renewal of these assets including Fan A must be considered. This table was auto-generated from the analytics and not developed manually. Hence the risk analysis is used as a searching tool for key issues of concern across the asset portfolio.

4 Risk Assessment—Asset Condition Data

With respect to the statistical evaluation of measurement data, we note that there are characteristic statistics of a distribution for a common meter applied to many assets of the same type. This is a result of the meter logic presented in Sect. 2, where measurements are obtained using consistent meter schemas applied to many assets of the same kind. This provides a means to reduce the analysis of potentially millions of individual data points into a manageable approach.

It is possible to derive a set of condition levels for each meter based on a deterministic understanding of the failure mechanism. These relate to the condition levels which are shown in the diagram but are completely unrelated to the history of measurements. In many industrial data sets these levels are not established, and there are no registered alert levels to assist the inspection teams. This lack will result in plant reinstated from maintenance which is out of specification (Fig. 3).

Table 2 Sample watch list for assets—top 10 list for entire facility

Description	Number of corrective work orders	Risk level
Pump—centrifugal (PCS C)	161	100
Fan—centrifugal (CVS exhaust B)	117	100
Fan—(cooling tower A)	105	100
Fan—centrifugal	100	100
Filter—self cleaning (SCS)	57	100
Lift—goods	20	100

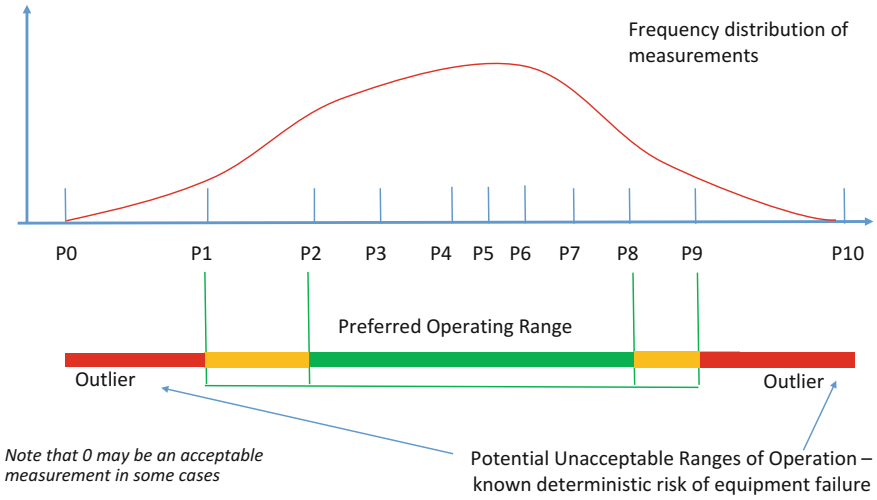


Fig. 3 Determination of alert limits

Considering the full distribution of all data associated with a specific meter, a frequency distribution plot can be compiled as shown by the red trace on the plot above. If we mark of the distribution in deciles, then it is plausible that:

- Readings in the P0–P1 and P9–P10 range are outliers which are worthwhile being followed up but may not represent an actual material risk in the assets. There is a concern that measurements in these ranges may be anomalies such as incorrect readings;
- Readings in the P2–P8 range are in the preferred operating range of the asset class for that kind of meter; and
- Readings in the P1–P2 and P8–P9 level are worthwhile warning limits since the measurements are starting to go outside the stochastically determined boundaries of preferred operating range.

The stochastic limits, as compared to the deterministic limits developed on physical considerations, represent reasonable early warning signs when an asset is operating outside the preferred operating range and can warn of:

- Deteriorated asset drifting out of the preferred range;
- Measurements which have accuracy issues and need to be validated; or
- An asset which is consistently operating outside the normal range for these assets and which is therefore more highly stressed than its counterparts.

Measurement trend data can be assigned a risk statistic based on the following:

$$Risk = \frac{\Delta(X_{max} - X_i)}{X_{max} - X_{min}} \times \frac{\Delta(Date_{Last} - Date_i)}{Date_{Last} - Date_{First}} \quad (1)$$

This is a qualitative measure which reflects that measurements approaching the alert limits are of concern but also considers how recently the measurement was taken. If the measurement is quite old but the asset is still in service or more recent measurements are not of such concern, then the second multiple takes this issue into account. Note that this statistic is a searching parameter, not intended to reflect an absolute measure of risk such as likelihood of failure, but a comparator intended to search out assets of concern amongst thousands of readings.

This is shown graphically on the chart below (Fig. 4)

$$Risk = \frac{\sum R1 \times R2}{Count(Points)} \quad (2)$$

R1 and R2 are normalized to lie between 1 and 10 where the formulation sets 10 to be the highest possible value and 1 the lowest. Hence if the data point is close to P10 then R1 approaches 10 and if it is current, then R2 will approach 10. This score will be inverted when the condition is monitoring is associated with data in the lower bound (i.e. in the P0 to P2 range). Risks are typically normalised between 1 and 100, with a risk score of 100 being the most severe risk. This allows a searching algorithm to search a sea of condition data (i.e. millions of data points) and identify the bad actor equipment which should then be followed up.

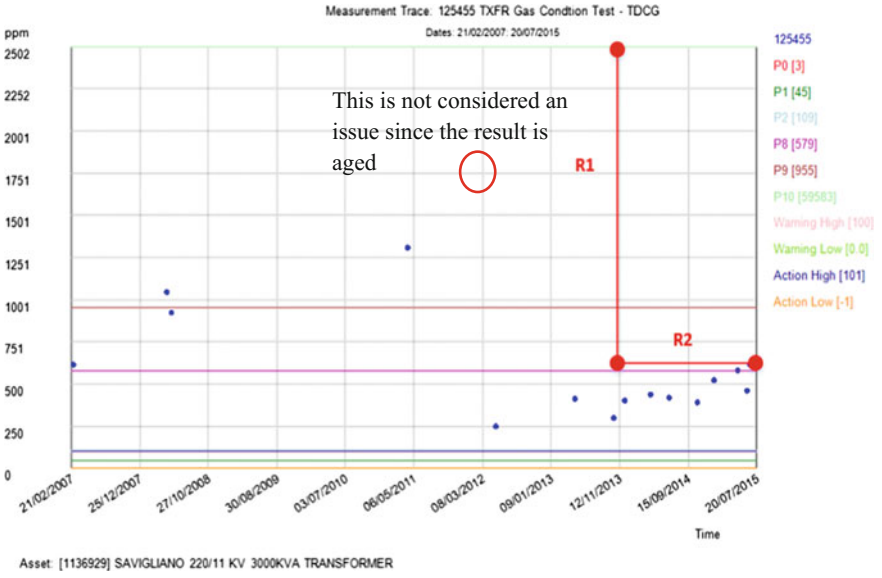


Fig. 4 Qualitative risk statistic for the data trend

The risk-based reporting methods to search of assets of concern which were introduced in Sect. 3 for work transactions can be applied to this approach to identify assets of concerns based on asset condition. An example is shown in Fig. 5.

The searching routine allowed a potential issue to be detected with the monitoring of Brushgear on a specific unit (Fig. 6).

The following observations may be made. The original routine monitoring strategy was discontinued, the results were above the P9 range established by the older measurements. Given that this is Brushgear and a long measurement is preferable since the failure condition is wear of the brush, this means that the surveys were picking up lengths which were optimal.

One measurement was down at the P2 level which was a trigger for replacement. It is noted that measurements resumed after the replacement and validated the trend in the wear rate. Based on these results the risk is small and the relaxation of routine brush length measurements appears validated. However, the period between measurements in the new regime was too long and risk was being incurred as the late measurement approached P2.

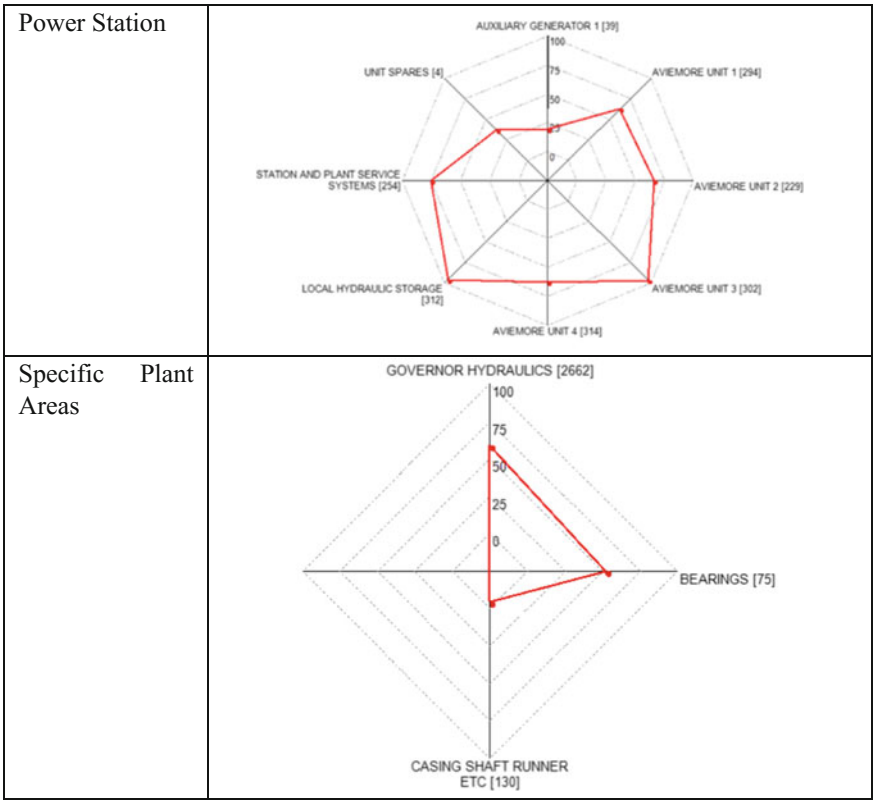


Fig. 5 Risk profiling based on condition risk assessment

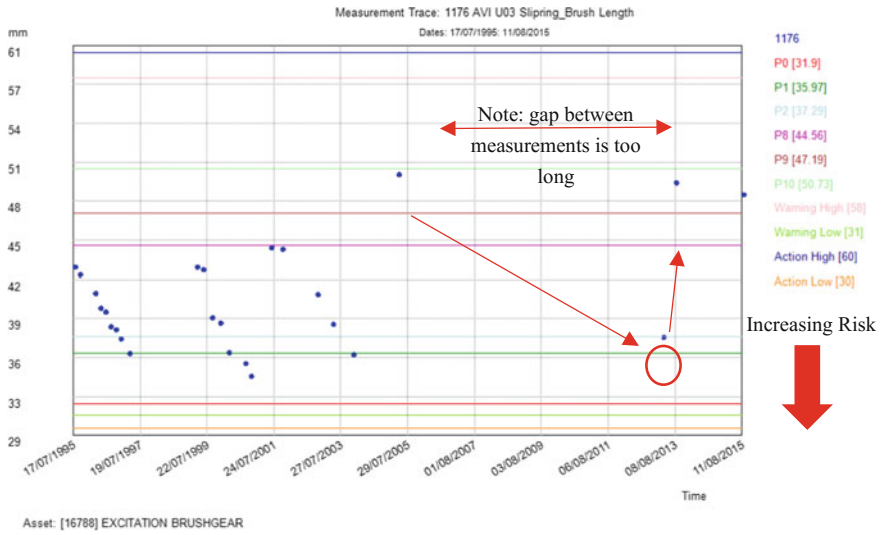


Fig. 6 Brush length measurement data

Statistical analysis can back fill measurement alert levels when these are not set deterministically (which is a major exercise). Millions of individual measurement points, representing a data sea, can be mined using logical algorithms which consider likely deterioration behaviour. An extensive portfolio of assets can be assessed swiftly and individual risks can be brought forward and then analysed. Credible recommendations for either asset renewal or modification of maintenance strategy can be supported by results from the analytics.

5 Conclusion

Organisations need to use the asset information data which their work delivery teams routinely capture. The premise of this work is to accept the quality of data captured in normal work delivery rather than seek high levels of codification and quality control before it is used. Robust searching techniques based on simple concepts such as introduced in this paper can be applied to search out the critical few measurements or transactions of concern, identifying assets with potential issues. Priority should be given to use reports which are insightful and can be used to optimise investment.

As the communicability of the reports improves and hence people have a greater likelihood of basing their decisions on them, the data environment will improve because there is an improved awareness that taking the measurements or recording the transactions in the first place supports something valuable. Improved data quality will then support an even greater capability in analysis and reporting.

In the first instance, optimisation is sought by identifying the critical few assets which will lead to operational downtime, non-compliance, high cost or other consequences. The analysis had to inform what is the issue with these assets: all work is urgent, high corrective maintenance work loads, lack of knowledge of condition or unacceptable deterioration?

This paper has presented three considerations an organisation must consider: the efficient management of large data sets which enables creative algorithm design and scenario testing, robust mathematics which will work with industrial quality data sets (as compared to highly conditioned data) and effective reporting which allowed drill down and clarity for the executive.

References

1. Brissaud F, Lanternier B, Charpentier D (2011) Modelling failure rates according to time and influencing factors. Int J Reliab Safety 5(2):95–109
2. Cao X, Jiang P, Zhou G (2008) Facility health maintenance through SVR-driven degradation prediction. Int J Mat Prod Technol 33(1/2):185–193
3. Hodkiewicz MR (2013) Relevant asset health information for executives. Int J Strateg Eng Asset Manag 1(4):425–438
4. 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(1):100–126
5. Mkandawire BO, Ijumba NM, Whitehead H (2011) Application of maintenance tools and strategies in integrated risk management of critical physical assets. Int J Agile Syst Manag 4 (3):261–279
6. Platfoot R (2014) Practical analysis for maintenance teams using computerised maintenance management system work history. Aust J Multi-Disciplinary Eng 11:91–103
7. Solozhentsev E, Karasev V (2015) Risk management technologies in structural complex systems. Int J Risk Assess Manag 18(3/4):307–318
8. Yigitbasioglu OM, Velcu O (2012) A review of dashboards in performance management: implications for design and research. Int J Account Inf Syst 13:41–59

Vibration Analysis of Machine Tool Spindle Units

Ali Rastegari

Abstract Machine tools cannot produce accurate parts if performance degradation due to wear in their subsystems (e.g., spindle units) is not identified and controlled. Appropriate maintenance actions delay possible deterioration and minimize machining system stoppage time that leads to lower productivity and higher production cost. Measuring and monitoring machine tool condition has become increasingly important because of the introduction of agile production and increased requirements for product accuracy. Condition Based Maintenance (CBM) techniques, such as vibration monitoring, are becoming a very attractive method for companies operating high-value machines and components. One of the most common problems of rotating equipment, such as machine tool spindle units, is the condition of bearings. Vibration analysis can diagnose bearing damage by measuring the overall vibration of a spindle or, more precisely, by high-frequency techniques such as enveloping. This paper focuses on the use of vibration analysis to monitor and analyse the condition of machine tool spindle units. The aim of the paper is to present the important factors in a vibration analysis of machine tool spindle units. The method is a case study at an automotive manufacturing company in Sweden. CBM, using vibration monitoring, is implemented on different types of machine tools, including turning machines, machining centres, milling machines and grinding machines. The results of the implementation, as well as a vibration analysis of a spindle unit and its cost effectiveness, are presented in the paper. The results indicate that detecting faults or damages by vibration monitoring of complex structures, such as spindle units, is challenging because there are different sources of frequencies from spindle bearings, gearboxes, gear meshes, etc. However, with help of advanced vibration analysis, such as high-frequency measuring methods, it is possible to detect bearing damage in spindle units from a very early stage of the damage.

A. Rastegari (✉)
Volvo Group Trucks Operation, Köping, Sweden
e-mail: ali.rastegari@volvo.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_51

1 Introduction

To be competitive, it is possible to reduce fabrication downtime by applying CBM techniques such as vibration analysis. Measuring and monitoring machine tool accuracy and capability have become increasingly important because of increasingly stringent accuracy requirements for industrial products and the products' functional and legislative requirements [3]. According to ISO 230 [2], machine tool vibration must be controlled to mitigate the types of vibration that produce undesirable effects such as "*unacceptable cutting performance with regard to surface finish and accuracy, premature wear or damage of machine components, reduced tool life, unacceptable noise level, physiological harm to operators*" (p. vi). The increased capabilities of manufacturing in measuring and monitoring will provide fewer machine failures, smaller spare parts inventories, and reduced production and maintenance costs [1].

The spindle system is one of the critical subsystems of a machine tool, which supplies the necessary power to the cutting process [1]. The spindle is a high-precision component comprised of several parts, including the rotor shaft, bearings, and the clamping system. In the majority of spindles, high-precision bearings with built-in preloads are used to enhance the stiffness of the bearing arrangement and increase running accuracy. The spindle bearings operate without load, or under very light load, and at high speeds. In such cases, the preload in high-precision bearings serves to guarantee a minimum load on the bearings, prevents bearing damage resulting from sliding movements and prolongs service life [12]. However, because the lifespan of spindle bearings is unknown for different machines, and each machine is in a different work environment, it is difficult to plan for spindle renovation. According to Randal [7], one of the principal techniques for obtaining information on internal conditions is vibration analysis. A machine in standard condition has a certain vibration signature; the growth of a fault changes that signature in a way that can be linked to the fault. Vibration analysis can diagnose spindle bearing damage and prevent the undesirable consequences of that damage.

As extensions to the previous studies by Rastegari et al. [11], further studies were performed on the vibration analysis of machine tool spindle units. The case studies were performed in a large automotive manufacturing company in Sweden. The company's product is gearboxes, with a production volume of 135,000 pieces per year. The company operates with approximately 800 medium-size CNC machine tools. The CBM, using vibration monitoring technique was implemented on different types of machine tools, including turning machines, machining centres, milling machines and grinding machines. The aim of this paper is to present the important factors in a vibration analysis of machine tool spindle units and the results of the analysis.

2 Rolling Element Bearing Vibration Analysis

One of the most common concerns of rotating equipment such as machine tool spindles is the bearing condition. Bearing failure can cause major damage to shafts, rotors, and housings. The majority of bearings fail before the natural fatigue limit of the bearing steel has been reached [6]. The most common causes of bearing damage are inappropriate lubrication, fatigue due to normal and parasitic loads, poor installation, and contaminations [4]. Bearing condition monitoring provides information about the condition of bearing lubrication, possible bearing damage, and the need for special maintenance or bearing replacement. Vibration analysis is a common bearing condition monitoring technique. In rotating equipment, vibration analysis can diagnose failures by measuring overall machine vibration or, more precisely, frequency analysis [13]. Vibrations are measured with an accelerometer. The vibration velocity parameter is measured as the broadband vibration magnitude in mm/s rms (root mean square). The vibration acceleration parameter is measured in mm/s^2 rms. Vibration measurement can be performed either by portable analysers or online continuous monitoring systems.

It is necessary to analyse spectrum and time waveforms when there is any suspicion of a fault condition. The spectrum is a summary of the vibration within the machine. Fast Fourier Transform (FFT) uses the time waveform to calculate how much of each frequency is present and displays that as a line in the spectrum. The time waveform is a record of what occurred from moment to moment as the shaft turns, the gears mesh, and the rolling elements roll around the bearing. Each minute change that results from impacts, rubs, scrapes, rattles, surges, etc. is recorded in the time waveform and then summarized in the spectrum [5]. For example, if there is damage on the inner race of a bearing there will be an impact each time the ball or roller comes into contact with the damaged area. It will be seen in the spectrum as harmonics of a frequency that is not a multiple of the shaft turning speed (non-synchronous harmonics) with sidebands. Impacts are also observable in the time waveform [5].

Ball pass frequency, inner race (BPFI) is the rate at which a ball or roller passes a point on the inner race of the bearing [4]. If there is damage on the inner race, a periodic pulse of vibration is observed at this rate. Ball pass frequency, outer race (BPFO) is the rate at which a ball or roller passes a point on the outer race of the bearing [4]. If there is damage on the outer race, a periodic pulse of vibration is observed at this rate. FTF is the fundamental train frequency, and BSF/RSF is the ball/roller spin frequency [8].

Rolling bearing damage in the early stages can be detected by using high-frequency measuring methods for the analysis, such as enveloping or demodulation, shock pulse method (SPM) and PeakVue. Enveloping and demodulation are two names referring to the same technique. These techniques use the high-frequency vibrations that are observed as bearing damage and make them available as low-frequency vibrations that can be analysed [4]. The SPM monitors and analyses the high-frequency compression (shock) waves created by a rotating

bearing [6]. Damage to the bearing surface causes a significant increase in shock pulse strength that can be seen by SPM. Shock pulses are measured with a transducer with a natural frequency of 32 (kHz) [14]. The PeakVue method uses a signal processing technique to detect signs of bearing wear using a technique similar to acceleration enveloping. It samples the vibration at a very high rate (102.4 kHz) to detect any short duration stress waves that occur in the earliest stage of bearing fault [4]. Similar to the enveloping technique, the PeakVue method involves a filter setting suitable for the application based on machine speed and type.

The bearing wear or failure process can be described in four stages [15]. At the first signs of minor bearing damage, the vibration amplitude will be very low. The vibration generated will be very high frequency—possibly over 10 kHz. Vibration spectrum analysis and time waveform techniques will not detect the fault. Only spectrum from the high frequencies will reveal a fault. High-frequency techniques such as enveloping, SPM, and PeakVue may detect the fault in stage one, but only if the filter is set correctly and the accelerometer or shock pulse sensor is properly mounted. As the bearing fault develops, high-frequency techniques will be more effective. However, it is questionable if a linear velocity spectrum will indicate the fault. The fault is more likely to be seen in the acceleration spectrum. The time waveform in units of acceleration will demonstrate signs of the defect, especially when applied to slow-speed machines [5]. At this stage, measurement should be planned more often. When the bearing fault reaches stage three, the damage is more severe and will be visible if the bearing is removed. The velocity spectrum can be used to detect the fault in addition to the time waveform (in velocity or acceleration units) and high-frequency techniques [5]. At this stage, replacement should be planned soon. Bearing damage at this stage in a spindle unit can affect the quality of the work piece, or it can cause damage to the other machine components such as the cutting tool. When the bearing fault reaches stage four, the bearing has major damage and should be replaced as soon as possible. As the condition deteriorates, high-frequency techniques become less effective. Vibration levels will increase, and the velocity spectrum will indicate the fault. The non-synchronous harmonics and sidebands will disappear in the spectrum, and the spectrum will be very noisy. As vibration becomes noisier, the waveform will become noisier and less effective [5].

3 Findings

Part of the results of the studies, including a vibration analysis of machine tool spindle units and its cost effectiveness, are presented in this section.

3.1 Implementation of Vibration Analysis on Machine Tool Spindle Units

In previous studies performed by Rastegari, the implementation process and strategy of vibration measurement on spindle housings are presented [9, 11]. Some of the important factors of the vibration measurement of spindle units are summarized in Table 1 [11, 16].

The following items are relevant:

- If the machine is in the warrantee period, the machine manufacturer should be informed of any change in the machine for measurement.
- The sensors should be calibrated after five years of use to make sure they function properly. Consider using close contact to install sensors in a wet condition.
- Techniques such as the bump test can be used to find the resonance areas to avoid measuring machine vibrations in the resonance areas. It is best to measure at the machine's operational speed to avoid resonance frequencies. Background vibrations should be removed while measuring. Background vibrations can emanate from nearby machines.
- The frequency range of measurements varies for different spindle speeds, bearing types, numbers of gear teeth and different measuring methods for the analysis. Vibration velocity parameters can be measured within a frequency range of 100 times the operational speed frequency (e.g., a spindle with an operational speed of 3000 rpm and 50 Hz can be measured in the frequency

Table 1 Implementation factors of vibration measurement of spindle units

Implementation factor	Explanation
The vibration sensor type	Vibration accelerometer with sensitivity of 100 mV/g for spindles with operating speed between 60 and 30,000 rpm
The sensor installation	In radial and axial directions at the front and back ends of the spindle, as close to the bearing as possible with use of threaded fittings, chemical bonding wax or a magnetic base
The machine loading condition for measurement	No-load conditions (no cutting, grinding, or milling) either with a tool clamped or no tool in the machine
The machine rotational speed for measurement	The machine's operation speed, to avoid resonance frequencies
The measurement frequency range	Varies for different spindle speeds, bearing types, number of gear teeth and using different measuring methods for the analysis
Automatic or manual measurement	Using handheld vibration instruments or on-line vibration measuring units (analysers)
The alert and alarm limits	The alert value is set to 50% higher than the baseline value and the alarm value to 150% higher than the baseline value
The measurement strategy and measurement interval	Different measurement time interval depending on types and criticality of the machine

range of 10–5000 Hz). However, high frequency techniques, such as enveloping, use filter settings to detect bearing damages; i.e., a frequency range of 0.5–10 and 5–40 kHz are used in the enveloping technique.

- In online measurement, a measuring unit/analyser should be installed on the machine, and the machine should be programmed to activate the measuring unit to start and stop the measurement at a specific time interval. The machine should be programmed to run the measurement according to the required factors such as constant spindle running speed and no-load condition.
- Different types of machines with different criticalities should be measured at different time intervals, i.e., slow-moving machines only measured once a month. Very fast machines and the most important machines should be measured more often or be provided with online vibration monitoring.
- All measurements should be documented and a trend analysis should be performed. In the case of unacceptable vibration levels, the cause must be found immediately by using spectrum analysis and time waveform analysis with different methods such as high-frequency methods. The types of the rolling bearings should be known in order to ease finding bearing damage.

3.2 Vibration Analysis of a Machine Tool Spindle Unit with Bearing Damage

By performing vibration analysis on a hard-turning machine's spindle with integral drive, damage was found in a N1018 bearing at the non-drive end of the spindle. As indicated in Fig. 1, an accelerometer was installed in the axial direction on the spindle housing close to the bearing location.

The spindle has an operational speed of 1100 rpm and 18.33 Hz that was measured in the frequency range of 5–2500 Hz. Figure 2 shows the spectrum analysis of the vibration velocity and acceleration parameters of the spindle. In the spectra, non-synchronous harmonics and shaft-rate sidebands are seen. The higher peaks in the spectrum correspond with Ball Pass Frequency Inner Race (BPFI), which equals 253.08 Hz and 13.80 orders for the N1018 bearing. Since periodic pulses of vibration are observed at this rate, there is damage on the inner race because the impact harmonics of the BPFI frequency are seen in the spectrum.

Since a spall travels around the bearing once per revolution, the impacts will not be equal in amplitude. Therefore, in an inner race defect, there are shaft rate sidebands in the spectrum, as indicated in Fig. 3.

By observing the damage in the velocity spectrum, the possible actions are to plan for spindle replacement and, in the meantime, monitor more frequently.

The damage was clearly observed using high-frequency techniques, including enveloping, SPM and PeakVue. As shown in Figs. 4 and 5, the non-synchronous harmonics in the spectra clearly correspond with the BPFI.

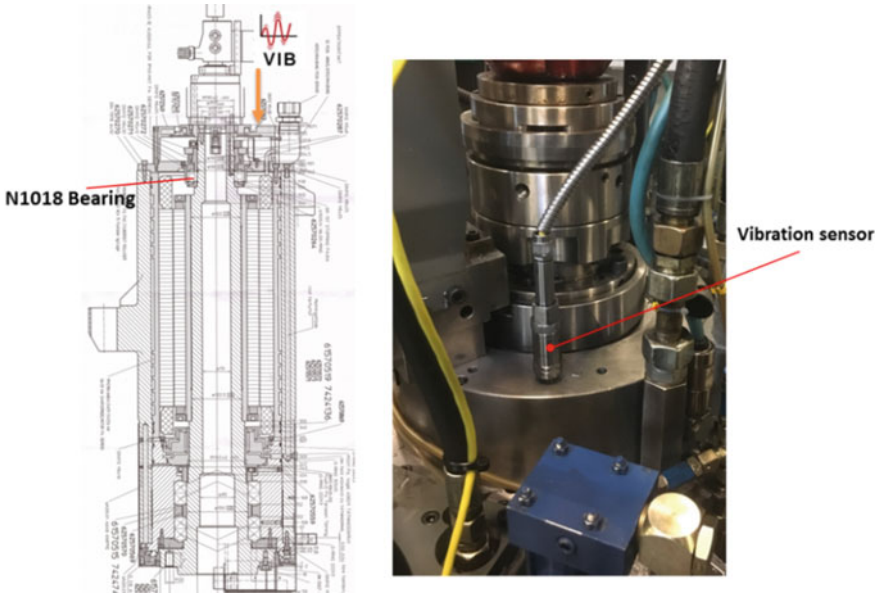


Fig. 1 Hard-turning machine spindle

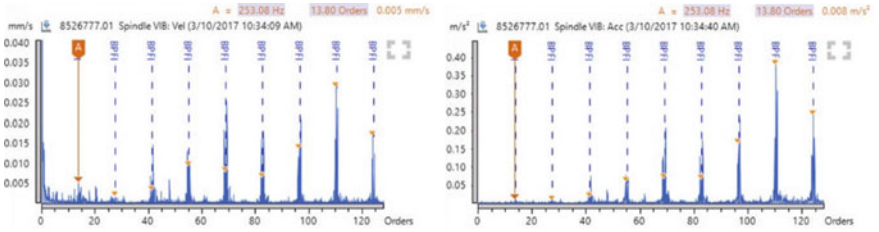


Fig. 2 Spectrum analysis of the spindle with bearing damage in inner race using vibration velocity unit (on the left) and vibration acceleration unit (on the right); frequency range of 5–2500 Hz

As indicated in Fig. 6, the time waveform also shows inner race damage with a clear impact. The typical pattern in the time domain is a result of the damage on the inner race entering and leaving the loaded zone. The pattern in the time domain is repeated for every revolution (B2-B1); the distance between the small peaks equals BPFI (A2-A1). The carrier frequency is BPFI, and it is modulated by the shaft turning speed.

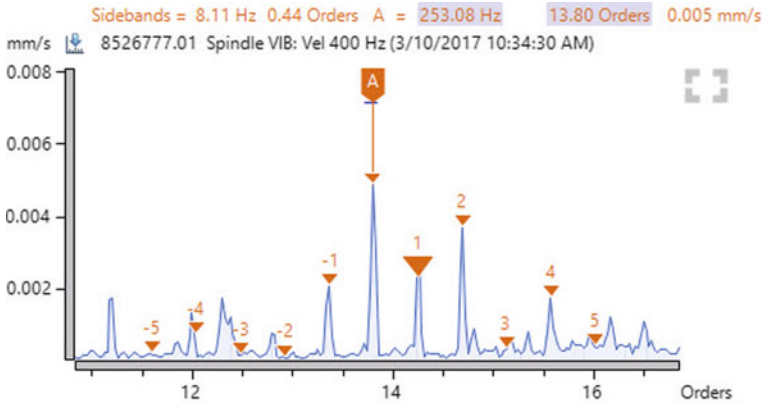


Fig. 3 BPFI with sidebands

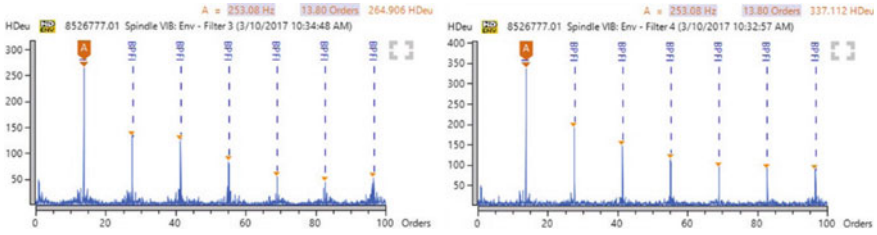


Fig. 4 Spectrum analysis of the spindle with bearing damage in the inner race using enveloping method; frequency range of 0.5–10 kHz (on the left) and 5–40 kHz (on the right)

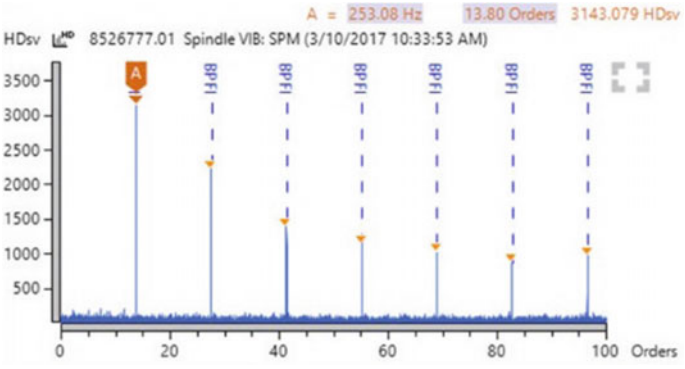


Fig. 5 Spectrum analysis of the spindle with bearing damage in the inner race using SPM (on the left) and PeakVue (on the right)

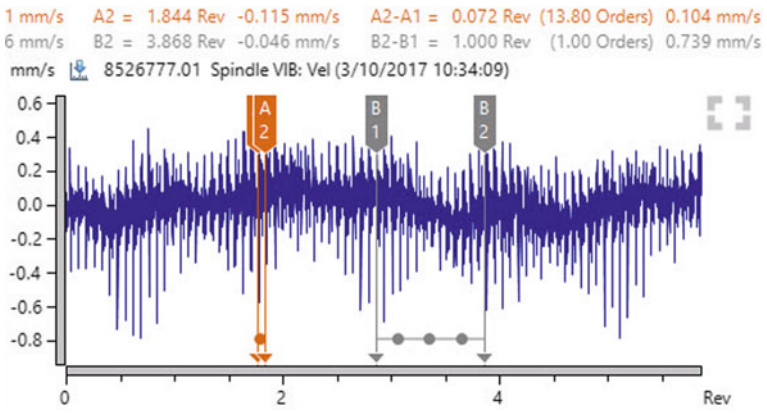


Fig. 6 Time waveform analysis of the spindle with bearing damage in the inner race

3.3 Bearing Damage and Failure Analysis

By considering these analyses, the spindle bearing fault was successfully diagnosed, and the spindle has been replaced. In addition, the bearing was removed from the spindle for further analysis. Figure 7 shows the damage on the inner race of the bearing in the rolling element contact area.

The early diagnosis was false brinelling damage due to vibrations during standstill. This occurs when vibrations from nearby machines travel through the bearing of the stationary machine and damage it. However, the machine does not have long enough stops for this type of damage. Therefore, further analysis shows that the root cause of the damage was that the spindle had the same start and stop

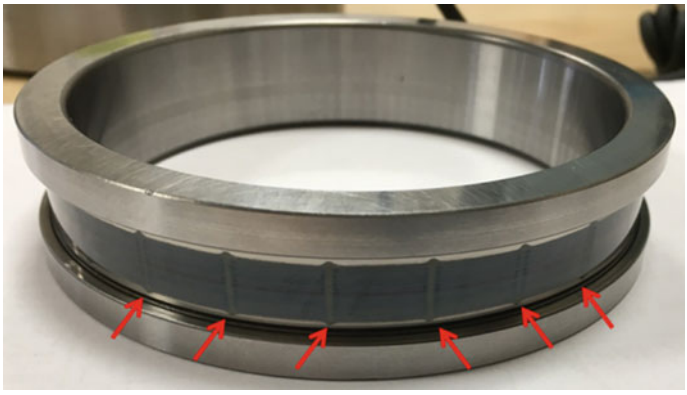


Fig. 7 Wear on the inner race of the cylindrical roller bearing

geometry position. This means that the spindle has always been stopping in the same position when changing the work piece. The bearing has also been stopping in the same place. This is the root cause for wear on the inner race in the rolling element contact areas. The recommended action was to re-program the machine to have different stopping positions for the spindle.

3.4 Cost Effectiveness of Vibration Analysis of Machine Tool Spindle Units

The initial reason for troubleshooting the machine was a quality issue, which caused an extra 10 s for each work piece. Therefore, other actions were taken to fix the problem with the machine, such as changing the ball-screws and pulleys at a cost of 150 KSEK. These actions required 50 h of production time with the loss of at least 50 KSEK for production. However, the main issue with the machine was not solved. Vibration analysis was then performed, and the bearing damage was found in the spindle. The spindle was replaced and the root cause of the bearing damage was eliminated. Therefore, the probability of the same problem occurring in the spindle is quite low, and the company can avoid a 350 KSEK cost for spindle renovation in the near future. As a result, the cost avoidance with vibration analysis is estimated as at least 550 KSEK in this case. According to Rastegari and Bengtsson [10], there are other indirect costs that are not considered in this estimation such as quality loss. Production time is reduced by 10 s per work piece and tool life is increased; these are not estimated in the cost avoidance.

The cost effectiveness of vibration analysis of spindle units in general is considered in the following objectives; reduced risk of safety issues, reduced risk of scrap or low-quality products, increased tool life, reduced costs for troubleshooting, reduced costs of spare spindles in the warehouse, using the maximum life length of the spindle unit instead of replacing it by planned maintenance when it is not needed, and reduced downtime due to acute breakdowns. Vibration analysis can also be very cost effective on machines that “cannot/do not have to” have unplanned stops in the production line due to high production volume. Machines should be measured and analysed before being entered into service. In the case of identified deviations, the machine users can demand actions by the machine manufacturers, such as free machine renovation, free warrants/upgrades for some years and prolonged guarantees.

4 Conclusions

The aim of this paper is to present the important factors in a vibration analysis of machine tool spindle units and the results of the analysis. Detecting faults or damages by vibration monitoring of complex structures, such as spindle units, is challenging because there are different sources of frequencies from spindle bearings, gearboxes, gear meshes, etc. However, with help of advanced vibration analysis, such as high-frequency measuring methods, it is possible to detect bearing damage in spindle units from a very early stage of the damage. This results in shorter maintenance cycles and lower costs because of lower downtime, lower catastrophic failures, lower secondary damage and reduced parts inventory. By increasing competence in this area, it is expected that machine condition monitoring systems using vibration analysis will enable companies to increase productivity, maintain quality and improve technical availability.

References

1. Abele E, Altintas Y, Brecher C (2010) Machine tool spindle units. CIRP Ann Manuf Technol 59(2):781–802
2. ISO/TR 230-8:2009 (2009) Test code for machine tools—part 8: vibrations
3. Martin KF (1994) A review by discussion of condition monitoring and fault diagnosis in machine tools. Int J Mach Tools Manuf 34(4):527–551
4. Mobius Institute (2016) Vibration analysis training course book
5. Mobius Institute (2017) Vibration analysis definitions. Available at: http://mobiusinstitute.com/
6. Morando L (1996) Technology overview: shock pulse method. In: Proceedings of a joint conference, Mobile, Alabama
7. Randall RB (2011) Vibration-based condition monitoring: industrial, aerospace and automotive applications. Wiley
8. Randall RB, Antoni J (2011) Rolling element bearing diagnostics—a tutorial. Mech Syst Signal Process 25(2):485–520
9. Rastegari A, Bengtsson M (2014) Implementation of condition based maintenance in manufacturing industry. In: IEEE international conference on prognostics and health management, Washington, USA
10. Rastegari A, Bengtsson M (2015) Cost effectiveness of condition based maintenance in manufacturing industry. In: IEEE 61st annual reliability and maintainability symposium, Florida, USA
11. Rastegari A, Archenti A, Mobin M (2017) Condition based maintenance of machine tools: vibration monitoring of spindle units. In: IEEE 63nd annual reliability and maintainability symposium, Florida, USA
12. SKF General (2003) Principles of bearing selection and application, Catalogue
13. Starr A (2000) A structured approach to the selection of condition based maintenance. In: 5th international conference on factory, pp 131–138

14. Sundström T (2010) An introduction to the SPM HD method. An SPM Instrument White Paper, SPM Instrument AB
15. Sundström T (2013) The shock pulse method and the four failure stages of rolling element bearings. Technical report, SPM Instrument AB
16. Swedish Standard 2014. 728000-1:2014: machine tool spindles—evaluation of machine tool spindle vibrations by measurements on spindle housing—part 1: spindles with rolling element bearings and integral drives operating at speeds between 600 min-1 and 30 000 min-1

Unsteady Rotor Blade Forces of 3D Transonic Flow Through Steam Turbine Last Stage and Exhaust Hood with Vibrating Blades

Romuald Rzadkowski, Vitally Gnesin, Luba Kolodyazhnaya and Ryszard Szczepanik

Abstract Presented below are numerical calculations of the 3D transonic flow of an ideal gas through a steam turbine last stage with the exhaust hood, taking into account blade oscillations. The approach is based on a solution of the coupled aerodynamic-structure problem for 3D flow through a turbine stage in which fluid and dynamic equations are integrated simultaneously. This provides the correct formulation of the coupled problem because the blade oscillations and loads are a part of solution. An ideal gas flow through stator, rotor blades and the exhaust hood is described by unsteady Euler conservation equations, which are integrated using the explicit monotonous finite-volume difference scheme of Godunov-Kolgan and a moving hybrid H-H grid. The structure analysis uses the modal approach and 3D finite element model of a rotor blade. The pressure distribution behind the rotor blades were non-uniform on account of the exhaust hood. In the nominal condition a self-excited rotor blade vibration was discovered.

1 Introduction

Several high vibration amplitude problems have been reported regarding the slender last stage blading of commercial LP steam turbines. To solve this problem, unsteady forces acting on rotor blades have to be calculated. Unsteady prediction

R. Rzadkowski (✉)
Department of Dynamics of Machines, Institute of Fluid Flow Machinery, Polish Academy of Sciences, Gdansk, Poland
e-mail: z3@imp.gda.pl

V. Gnesin · L. Kolodyazhnaya
Department of Aerohydromechanics, Institute for Problems in Machinery, Ukrainian Academy of Sciences, Kharkiv, Ukraine
e-mail: gnesin@ukr.net

R. Rzadkowski · R. Szczepanik
Air Force Institute of Technology, Warsaw, Poland
e-mail: ryszard.szczepanik@itwl.pl

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_52

models for 3D non-viscous and viscous flutter and unsteady forces acting on rotor blades have been discussed in literature over the last ten years [1–7, 9]. So far most of the calculations of unsteady rotor blade forces have concerned the stator as well as rotating and vibrating rotor blades, with uniform or non-uniform pressure distribution behind the rotor blades, but not the exhaust hood. In this paper 3D Euler flows were used to calculate unsteady non-viscous flow through the last stage of a LP steam turbine with rotating and vibrating rotor blades and an exhaust hood. The regime selected for these calculations was one where high vibration amplitudes could occur.

2 Aeroelastic Model

The 3D transonic flow of inviscid non-heat conductive gas through an axial turbine stage was considered in the physical domain, including the nozzle cascade (NC) and the rotor wheel (RW), rotating with constant angular velocity and the exhaust hood (EH). In practice the number of stator and rotor blades differ. Taking into account the flow unperiodicity from blade to blade (in the pitchwise direction) it is convenient to choose a calculated domain including all the stator and rotor blades, the inlet region, the axial clearance and the exhaust hood (see Figs. 1, 2, 3 and 4).

The spatial transonic flow, generally including strong discontinuities in the form of shock waves and wakes behind blade trailing edges, was written in the relative Cartesian coordinate system, rotating with constant angular velocity ω, according to full non-stationary Euler equations, in the form of integral conservation laws of mass, impulse and energy. The calculated domain, including all blades on the whole annulus as well as inlet and the exhaust hood domains, consisted of subdomains (NC and RW, see Fig. 4) and (RW and EH) with a common parts. In turns, each of

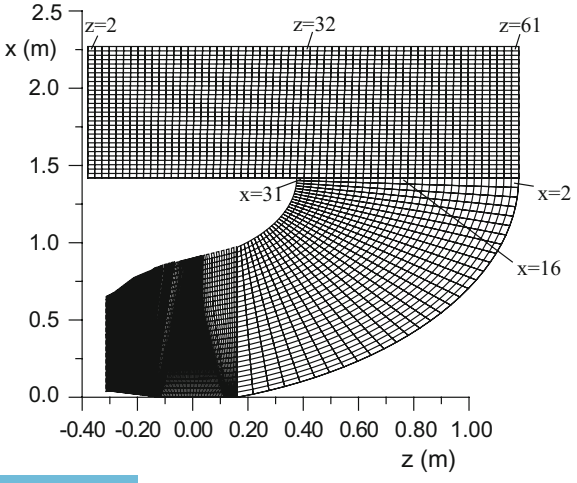


Fig. 1 CFD mesh of stator, rotor blades and exhaust hood

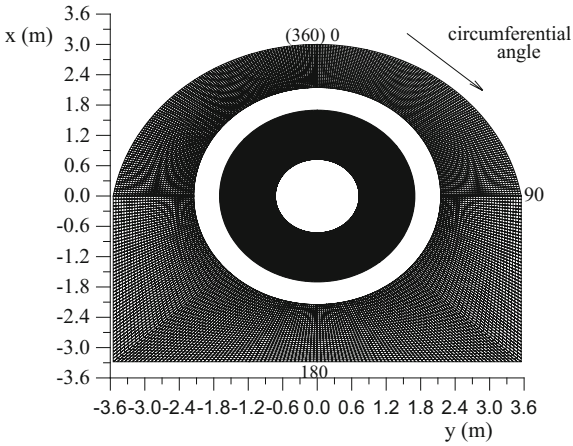


Fig. 2 CFD mesh of exhaust hood

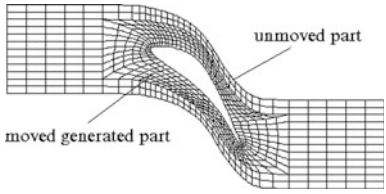


Fig. 3 Unsteady rotor blade grid generation

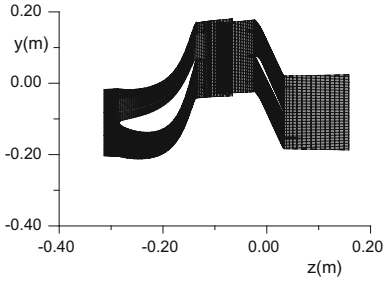


Fig. 4 CFD mesh of stator and rotor in peripheral layer

passages were discretized using H-type grid for stator domain and hybrid H-H grid for rotor domain [1].

The external rotor blade H-grid remains immobile during the calculation, while the inner H-grid is rebuilt in each iteration by a given algorithm. Therefore, the external points remain fixed, but internal points (points on the blade surface) move according to the blade motion (see Fig. 3). The gasodynamic parameters on the lateral sides are defined using the Riemann problem of an arbitrary discontinuity on the moving interfaces between two adjacent cells and by using of the iteration process [1]. In the general case, when axial velocity is subsonic, at the inlet boundary initial values for total pressure, total temperature and flow angles are

used, while at the outlet boundary in the exhaust hood only the static pressure has to be imposed. Non-reflecting boundary conditions must be used, i.e., incoming waves (three at inlet, one at the outlet) have to be suppressed, which is accomplished by setting their time derivative to zero. On the blade's surface, because the grid moves with the blade, the normal relative velocity is set to zero.

The dynamical model of the oscillating rotor blade in linearized formulation is governed by the matrix equation:

$$M\ddot{u} + C\dot{u} + Ku = F \tag{1}$$

where: M, C, K are the mass, mechanical damping and stiffness of the blade respectively u is the blade displacement; F is the unsteady aerodynamic forces vector, which is a function of blade displacement.

The first step of the modal approach consists of solving the problem of the natural mode shapes and eigenvalues without damping and in a vacuum. Then the displacement of each blade can be written as a linear combination of the first N modes shapes with the modal coefficients depending on time:

$$u = Uq = \sum_{i=1}^{N} U_i q_i \tag{2}$$

where: U_i is the displacement vector corresponding to i-th mode shape; q_i(t) is the modal coefficient of i-th mode. Taking into account Eq. 2 and the orthogonality property of the mode shapes Eq. 1 can be written in form of:

$$I\ddot{q} + H\dot{q} + \Omega q = \lambda(t) \tag{3}$$

where: I = diag (1, 1, …, 1), H = diag ($2h_1$, $2h_2$, …, $2h_n$), Ω = diag (ω_1^2, ω_2^2, …, ω_n^2) are diagonal matrices; ω_i is i-th natural blade frequency; $\lambda(t)$ is the modal forces vector corresponding to the mode shapes, $h_i = \omega_i \xi_i$, where ξ_i is the i-th modal damping coefficient [1]. Thus the dynamic problem (Eq. 3) reduces to the set of independent differential equations relatively to modal coefficients of natural modes:

$$\ddot{q}_i + 2h\dot{q}_i + \omega_i^2 q_i = \lambda_i \tag{4}$$

The equations of motion (Eq. 4) can be solved using any standard integration method.

The modal forces λ_i are calculated at each iteration with the use of the instantaneous pressure field in the following way:

$$\lambda_i = \frac{\iint_\sigma p\bar{U}_i \cdot \bar{n}^\circ d\sigma}{\iiint_v \rho \bar{U}_i^2 dv} \tag{5}$$

where: p is the pressure along the blade surface.

3 Numerical Results

The numerical calculations presented below were carried out for the last stage of an LP steam turbine. It included 48 stator blades and 52 rotor blades. The gasodynamic parameters at the inlet stage were P_0 = 12,800 Pa, T_o = 323. K. Three regimes were applied at the outlet region. Regime 1 with uniform pressure P_2 = 10,400 Pa behind the rotor blades. Regime 2 with non-uniform pressure distribution behind vibrating rotor blades calculated from a CFD model of the stage with exhaust hood, not vibrating but rotating rotor blades and static pressure P_2 = 10,400 Pa in the outlet [8]. And regime 3 with static pressure P_2 = 10,400 Pa in the exhaust hood outlet where the rotor blades are rotating and vibrating. The last stage with exhaust hood was analysed using an in-house code. The calculation domain included all blade-to-blade passages in the last stage and exhaust hood zone. In total 17 million finite volumes were used for 48 stator blades, 52 rotor blades and the exhaust hood. This model took into account blade vibrations. Thus the unsteady forces and blade motion were calculated.

Figure 5 shows the main components of the unsteady axial force acting on the rotor blades in the peripheral layer (regime 3). The low excitation harmonics are not multiples of 50 Hz, but occur at 54 Hz (0.0065 A/A_o), 63 Hz (0.0114), 72 Hz (0.038), 81 Hz (0.03) and 153 Hz (0.044). The harmonic amplitude of the stator travelling frequency, 50 Hz × 48 = 2400 Hz (0.055), was smaller than that of the low frequency components. The low excitation harmonics for regime 2 occur at 84 Hz (0.052 A/A_o), 108 Hz (0.049), 156 Hz (0.07) and 168 Hz (0.077). The harmonic amplitude of the stator travelling frequency, 50 Hz × 48 = 2400 Hz is 0.055. The average values of aerodynamic loads under uniform pressure behind rotor blades (regime 1) were different from those under non-uniform pressure (regime 2 and 3) (Fig. 6), but the high excitation frequency amplitudes were smaller (0.03) in the case of unsteady axial forces. The main low frequency components for the uniform pressure distribution (regime 1) were 68 Hz (0.016), 88 Hz (0.006), 92 Hz (0.01) and 136 Hz (0.027). A similar dependence was found for unsteady circumferential forces, where the maximal low frequency amplitude for (peripheral layer) non-uniform pressure distribution (regime 3) was 0.0063 A/A_o (A_o = 14.2 N) at 63 Hz, 0.019 (72 Hz), 0.014 (81 Hz) and 0.009 (153 Hz), for regime 2 0.119 A/A_o (A_o = 18.73 N) at 84 Hz, whereas for uniform pressure distribution (regime 1) it was 0.0084 A/A_o (A_o = 18.21 N) at 64 Hz, 0.034 (68 Hz), 0.0080 (88 Hz), 0.0085 (92 Hz), 0.017 (136 Hz). Unsteady components of axial and circumferential forces for the exhaust hood and uniform pressure distributions are different.

The blade oscillations caused by unsteady aerodynamic loads were presented in the form of rotor blade peripheral section displacements. Figure 7 presents the peripheral circumferential bending oscillations and the amplitude-frequency characteristics of rotor blades with non-uniform pressure distribution caused by the exhaust hood (regime 3). The main frequency component of the axial bending amplitude was 0.57 A/A_o for A_o = 6.23 mm at 80 Hz. The main frequency

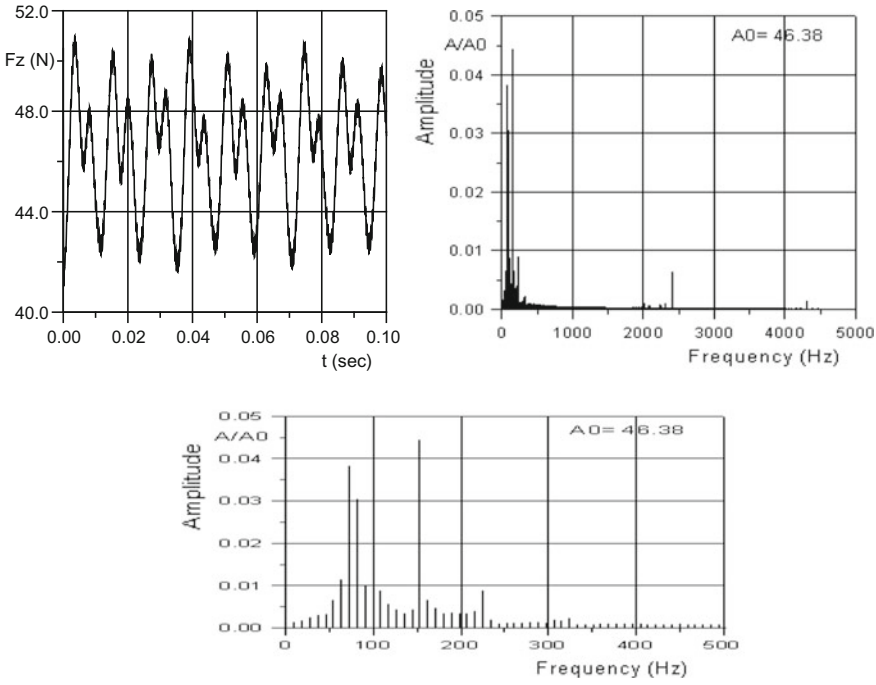


Fig. 5 Amplitude-frequency spectrum of peripheral axial force with non-uniform pressure distribution behind the rotor blades (regime 3)

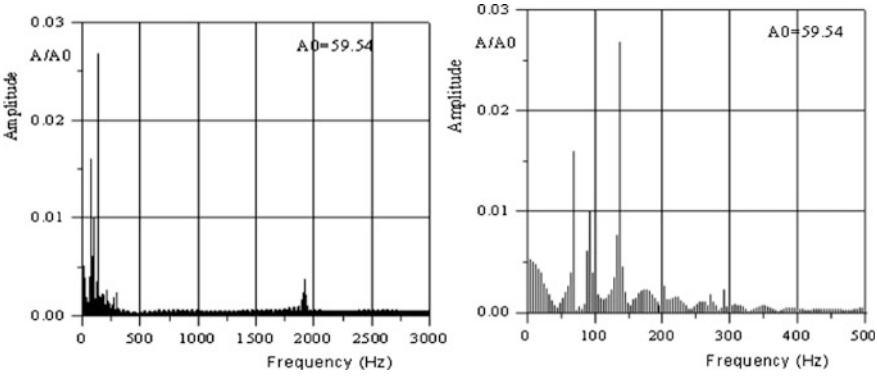


Fig. 6 Amplitude-frequency spectrum of peripheral axial force with uniform pressure distribution behind the rotor blades (regime 1)

component of the circumferential bending amplitude was 0.66 A/A_o for A_o = 5.38 mm at 80 Hz. In the case of non-symmetrical pressure distribution (regime 2, [6]) the axial bending amplitude was 0.65 A/A_o for A_o = 7.33 mm at

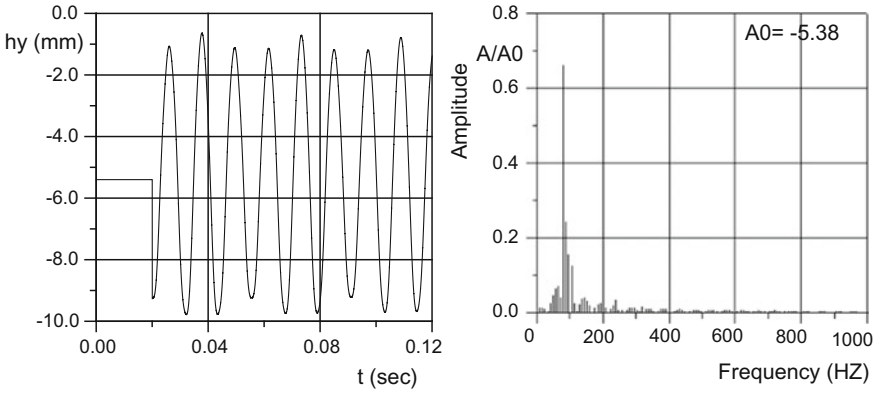


Fig. 7 Amplitude-frequency spectrum of peripheral circumferential bending with the exhaust hood (regime 3)

84 Hz. The main frequency component of the circumferential bending amplitude was 0.78 A/A_o for A_o = 5.02 mm at 84 Hz.

In regime 1, the axial bending amplitude was 0.54 A/A_o for A_o = 7.13 mm at 68 Hz, and the circumferential bending amplitude 0.86 A/A_o for A_o = 4.26 mm at 68 Hz.

As a result of the fluid-structure interaction, the first mode was no longer bending but bending-torsion. The main low frequency torsion components (regime 3) were: 0.44 A/A_o, A_o = 2.28 deg. at 80 Hz and 0.34 at 160 Hz (Fig. 8), but for non-uniform pressure distribution (regime 2, [6]) 0.33 A/A_o, A_o = 3.1 deg. at 84 Hz, 0.27 A/A_o at 108 Hz, 0.189 at 156 Hz and 0.18 at 168 Hz. In regime 1, 0.29 A/A_o, A_o = 3.17 deg. at 68 Hz, 0.12 at 92 Hz and 0.25 at 136 Hz.

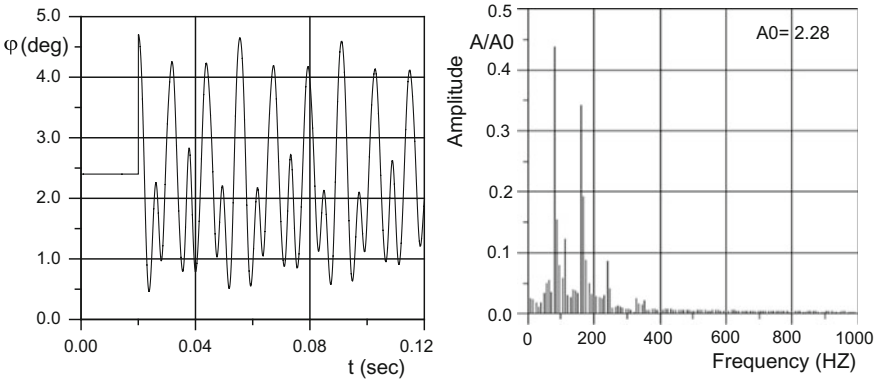


Fig. 8 Torsion angle and amplitude-frequency spectrum of blade in peripheral layer (regime 3)

The vibration blade frequencies and modes for uniform pressure distribution (regime 1), non-uniform pressure distribution (regime 2) and those with an exhaust hood (regime 3) were different. The main frequency of blade vibration for regime 3 was 80 Hz for regime 2, 84 Hz for regime 1, 68 Hz. The model with an exhaust hood gives the most reliable frequency. The values of vibration frequencies are very important with respect to blade resonance.

4 Conclusions

A partially-integrated method based on a solution of the coupled aerodynamic-structure problem was used to calculate the unsteady 3D flow through an LP steam turbine last stage and the exhaust hood, taking into account rotor blade oscillations. This paper has investigated the interaction of uniform and non-uniform pressure distribution behind the rotor blades and blade oscillations. The results show that the amplitude-frequency spectrum includes harmonics with frequencies which are not multiples of the rotation frequency and are actually close to the natural frequencies of the blades. This is a new finding. Conventional calculations of the last stage with exhaust hood concern blades that rotate but do not vibrate and therefore estimate low frequency components as multiples of 50 Hz. This is not enough to analyse blade stress. The first blade vibration mode is no longer bending but bending-torsion. In the case of non-uniform pressure distribution self-excitation occurred. But the general conclusion from the current study is that last stage rotor blades vibrate in the bending-torsion mode at frequencies lower than their first natural bending frequency. The results for uniform pressure distribution (regime 1), non-uniform pressure distribution calculated from CFD analysis with rotating but not vibrating blades without an exhaust hood (regime 2) and ones with an exhaust hood (regime 3) were different. This shows the significance of the exhaust hood in the calculation of unsteady forces acting on blades and blade vibration amplitudes.

Acknowledgements This research has been financed from Polish Government funds for the years 2012–2017 as a development project PBS1/B4/5/2012. All numerical calculations were made at the Academic Computer Centre TASK (Gdansk, Poland).

References

1. Gnesin V, Kolodyazhnaya L, Rzadkowski R (2004) A numerical model of stator-rotor interaction in a turbine stage with oscillating blades. J Fluids Struct 19(8):1141–1153
2. Huang XQ (2006) Influence of upstream stator on rotor flutter stability in a low pressure steam turbine stage. Proc Inst Mech Eng Part A J Power Energy 220(1):25–35
3. Petrie-Repar P, Fuhrer C, Grübel M, Vogt D (2015) Two-dimensional steam turbine flutter test case. In: ISUAAAT'2015, the 14th international symposium on unsteady aerodynamics, aeroacoustics and aeroelasticity of turbomachines, Stockholm, Sweden, 8–11 Sept 2015

4. Petrie-Repar P, Makhnov V, Shabrov N, Smirnov E (2014) Analysis of a long shrouded steam turbine. In: Proceedings of ASME turbo expo 2014, GT2014-26874
5. Rządkowski R, Gnesin V (2007) A 3D inviscid self-excited vibration of the last stage turbine blade row. J Fluids Struct 23:858–873
6. Rzadkowski R, Gnesin V, Kolodyazhnaya L (2018) Unsteady rotor forces of the last stage of LP steam turbine with non-uniform pressure behind vibrating rotor blades. J Vibr Eng Technol
7. Rzadkowski R, Kubitz L, Gnesin V, Kolodyazhnaya L (2018) Analysis of rotor blade flutter in last stage of LP steam turbine. J Vibr Eng Technol
8. Rzadkowski R, Surwilo J, Kubitz L, Lampart P, Szymaniak M (2016) Unsteady forces in last stage LP steam turbine rotor blades with exhaust hood. In: Proceedings of ASME turbo expo 2016, GT2016-57610
9. Sanvito M, Pesatori E, Bachschmidt N, Chatterton S (2012) Analysis of LP steam turbine blade vibration: experimental results and numerical simulations. In: 10th international conference on vibrations in rotating machinery, London, IMechE, 11–13 Sept 2012, pp 189–197

Advanced NDT Contributing Performance Evaluation of Civil Structures

Tomoki Shiotani, Takahiro Nishida, Hisafumi Asaue, Katsufumi Hashimoto, Shigeru Kayano, Yasushi Tanaka, Takuya Maeshima and Yoshikazu Kobayashi

Abstract Through the life cycle of civil infrastructures, quality assessments shall be implemented during construction, in-service, before/after repair and so forth; however, there are no decisive techniques to evaluate the inside of structures non-destructively. The authors have developed an advanced measurement method using a tomographic approach referred to as AE tomography without artificial excitations, leading to one-side access evaluation of large infrastructures. With this advanced technology, internal damage or defects can be visualized as a distribution of elastic wave parameters such as velocities so that damage identification consisting of locations and damage degree would be possible. In the paper, qualities of infrastructures in the consecutive phases of the life cycle are assessed with these tomographic approaches and the estimated damage is quantified with actual damage observation.

T. Shiotani (✉) · T. Nishida · H. Asaue · K. Hashimoto · S. Kayano
Kyoto University, Katsura Campus, Nishikyo, Kyoto, Japan
e-mail: shiotani.tomoki.2v@kyoto-u.ac.jp

T. Nishida
e-mail: nishida.takahiro.6e@kyoto-u.ac.jp

H. Asaue
e-mail: nishida.takahiro.6e@kyoto-u.ac.jp

K. Hashimoto
e-mail: hashimoto.katsufumi.8a@kyoto-u.ac.jp

S. Kayano
e-mail: kayano.shigeru.64m@st.kyoto-u.ac.jp

Y. Tanaka
The University of Tokyo, Tokyo, Japan
e-mail: yasuxi@iis.u-tokyo.ac.jp

T. Maeshima · Y. Kobayashi
Nihon University, Tokyo, Japan
e-mail: maeshima@civil.ce.nihon-u.ac.jp

Y. Kobayashi
e-mail: kobayashi.yoshikazu@nihon-u.ac.jp

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_53

1 Introduction

Severe deterioration in ageing infrastructures is one of the major concerns around the world, in particular fatigue damage in concrete slabs of highway bridges due to increasing heavy traffic. In order to maintain these infrastructures properly, preventive/proactive maintenance is desirable for prognosis, instead of reactive maintenance conducted after remarkable degradation. For a proper detection method of the early damage visualization, AE tomography is under investigation and developed as an innovative nondestructive testing (NDT). By evaluating the 3D velocity distribution inside an RC slab, damaging or deteriorating areas are to be identified. In the present study, evaluation of the fatigue damage progress introduced by wheel-loading is carried out to identify the damage locations quantitatively by means of the 3D AE Tomography. Furthermore, the velocity data obtained by 3D-AET was converted to the fracture parameter of the elasto-plastic and fracture model of concrete and the further live load deflections of reinforced concrete slab was estimated by the multi-scale analysis for performance assessment of the RC slab. In addition, the tomographic approaches are applied for actual bridge decks, specifically as for the steel plate bonded RC decks. An acceleration monitoring system is used for AE tomography, followed by the verification of tomography results with epoxy injection methods.

2 Two of Tomography Methods as Advanced NDT

2.1 Elastic Wave Tomography and AE Tomography

The preventive maintenances of concrete slabs in service are in urgent demand for establishing the prognosis in the concrete slab. In this respect, an advanced non-destructive testing (NDT) such as elastic-wave tomography and AE tomography are under investigation and development by authors [1–3, 7, 9]. Both tomography techniques can estimate internal deterioration of concrete by elastic wave parameters such as velocity distribution. In general, elastic-wave tomography requires information of the input of the excitation, i.e., generation time and location of excitation, for computing results. By evaluating resultant 3D velocity distribution inside the concrete, damaged or deteriorated areas are identified. In the present study, a comparative study during internal progress of fatigue damage introduced by a wheel loading apparatus is performed to identify the damaged areas quantitatively by means of the elastic wave tomography. On the other hand, AE tomography, as one of advanced tomographic approaches, can estimate both excitation sources and elastic wave parameters such as velocity, as proposed by Shiotani [8] and coded by Kobayashi and Shiotani [3]. An advantage of the AE tomography is that it is possible to diagnose structures of which only one-side access is allowed i.e., concrete bridge decks, linings of tunnels or huge structures.

In addition, secondary AE activity which is generated inside the objects also provides great benefit. Existent damages readily generate this secondary AE activity and, considering this activity as the excitation of tomography, elastic waves' parameters of the area of interest can be evaluated.

2.2 Methodology of the Tomography

In tomography, the region of analysis is divided into finite-element meshes. The slowness (reciprocal of the velocity) is calculated and assigned to each mesh, and finally the contour map of the velocity distribution is obtained. AE tomography can simultaneously calculate both of source locations and velocity distributions with the following procedure. The source location technique is based on the ray trace algorithm [6]. The arrival times t_{ij} from a receiver i are calculated. Since the first arrival time T_i at receiver i is observed, the arrival times E_{ij} at either nodal point or relay point j are derived,

$$E_{ij} = T_i - t_{ij} \tag{1}$$

The step is applied for all receivers, and then variance of the E_{ij} is computed from,

$$\sigma_j = \frac{\sum_i \left(E_{ij} - m_j\right)^2}{N} \tag{2}$$

where the coefficient m_j is defined as,

$$m_j = \frac{\sum_i E_{ij}}{N} \tag{3}$$

Here N is the number of receivers. The deviation is applied to estimate the accuracy of calculation. According to the definition of the deviation, σ_j is equal to zero and m_j becomes the emission time, in the case the distribution of the slowness is identical to that of the deviation. Although σ_j is not equal to zero even in normal cases because of discretization errors, the deviation, σ_j could become minimal and m_j shall correspond to a plausible time of emission by iterative procedure. In the conventional tomography, the time of signal excitation and arrival times are necessary for the analysis. In contrast, the time of signal excitation are readily estimated from the arrival times and location of AE source at once. This tomography is referred to as AE tomography. Some AE sensors can record arrival time of elastic wave when the wave is generated by an AE activity or a hammering. After each arrival time is determined, the propagation velocity through the path of the elastic wave is calculated using both the distance from the excitation point to the receiver point and T_{obs} (observed propagation time) which is obtained by Eq. 1.

$$T_{obs} = T_o - T_s \tag{4}$$

where T_s is the excitation time and T_o is the arrival time. On the other hand, in the algorithm of the tomography, the inverse of velocity, which is specifically referred to as the slowness, is given as an initial parameter into each element as shown in Fig. 1.

Then, T_{cal} (theoretical propagation time) obtained by a finite element model is the total of the propagation time calculated by the slowness and the distance the length crossing each element. ΔT is defined by observed propagation time (T_{obs}) and theoretical propagation time (T_{cal}) in Eq. 6.

$$T_i' = \int_0^L Sdl \tag{5}$$

$$\Delta T = T_{obs} - T_{cal} \tag{6}$$

Then the slowness in each element is revised in order to reduce ΔT. The slowness correction amount and the revised slowness are obtained by Eqs. 7 and 8 respectively.

$$\begin{bmatrix} \Delta s_1 \\ \Delta s_2 \\ \vdots \\ \Delta s_j \end{bmatrix} = \begin{bmatrix} \sum_i \frac{\Delta T_i^* l_{i1}}{L_i} / \sum_i l_{i1} \\ \sum_i \frac{\Delta T_i^* l_{i2}}{L_i} / \sum_i l_{i2} \\ \vdots \\ \sum_i \frac{\Delta T_i^* l_{ij}}{L_i} / \sum_i l_{ij} \end{bmatrix} \tag{7}$$

$$s_j' = s_j + \Delta s_j \tag{8}$$

where L_i is the total distance of the wave in the i element. The iterative calculation of Eqs. 7 and 8 enables to obtain the accurate slowness and finally the velocity in each element, corresponding to the observed propagation time of multiple waves over the structure, resulting in forming the tomogram of the elastic wave velocity over the target area [2]. Through these steps, velocity distributions were determined in the structure.

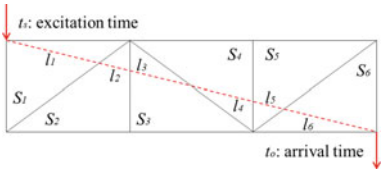


Fig. 1 Slowness for calculation of propagation time

3 Evaluation of Fatigue Damage by Experimental Test

3.1 Wheel Loading Test with Elastic Wave Tomography

In order to simulate cracks actually observed in an RC slab due to fatigue, repeated loading with a steel wheel was conducted as shown in Fig. 2. Dimensions of the RC slab specimen are 3000 mm in length, 2000 mm in width and 160 mm in thickness. The steel wheel with 400 mm width runs at the center of span. In the experiment, the cyclic motion of ±500 mm and the repeated rate 8.97 rpm were applied. Here, step-wise cyclic loadings are conducted based on the loading program as shown in Fig. 3. The initial wheel load was set at 98 kN which coincides with the allowable wheel load specified in the Japanese specification for highway bridges. The wheel load was subsequently increased by 33% from the previous load cycle at 100, 200 and 250 thousand cycles. In this way, the test was accelerated i.e., the magnitude of the load was increased step by step and as a result their traveling cycles were converted to equivalent cycles with regard to 98 kN load level by using Miner's law [5]. The fatigue limit of the specimen was estimated as 235 thousand cycles. The compressive strength of concrete was 20.7 N/mm^2. The specimen is simply supported in the longitudinal direction, and behaves elastically in the lateral direction. As both sides of top and bottom can be accessed, the elastic-wave tomography with artificial excitations was implemented. Specifically, elastic-wave excitations are implemented at designated locations illustrated as in Fig. 4. As the figure, 32 AE sensors of 60 kHz resonance are placed at 10 locations on the top, 18 locations on the bottom and 2 each on the two sides. The artificial excitations are driven by a steel-ball hammer of 35 mm diameter at designated 18 locations on the top and 22 on the bottom surface. The wheel loadings were applied in the area between green lines in the longitudinal direction as in Fig. 4.

Fig. 2 Wheel loading apparatus

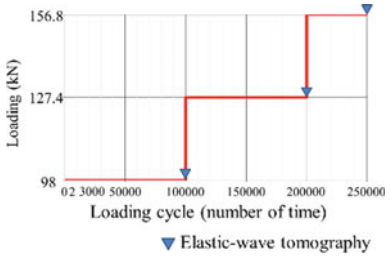


Fig. 3 Loading program

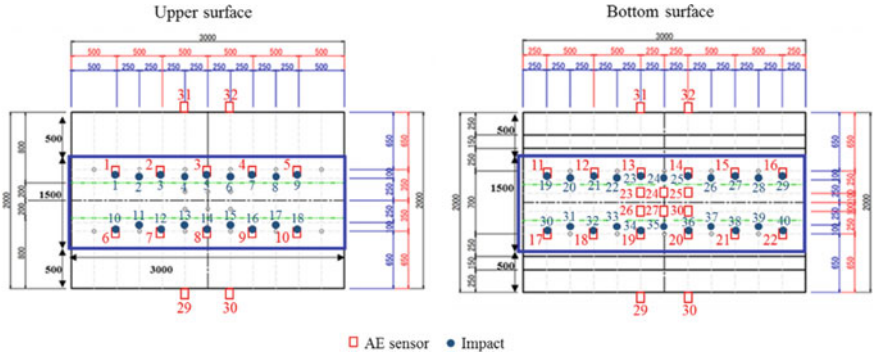


Fig. 4 Locations of AE sensors and tapping points

The analysis of the 3D tomography was conducted in the region of 3 m × 1 m of the slab, which consists of 25 nodes in the axial direction, 13 nodes in the lateral direction and 5 nodes in the thickness direction. Contour maps on velocity distributions are shown in Fig. 5. The situations are almost similar after 100×10^3 and 200×10^3 cycles except for the area at 2.3 m distance in the axial direction after 200×10^3 cycles, where the zone of the velocities lower than 2700 m/s is observed. After 250×10^3 cycles, which exhibits cycles beyond the fatigue limit, the zones of the velocities lower than 2500 m/s widely spread over the concrete slab. It is found that the velocity distributions show minor variations up to 100×10^3 cycles, while generation of low-velocity zones is initiated after 200×10^3 cycles, leading to the fatigue limit before 250×10^3 cycles.

3.2 Data Assimilation of the Multi-scale Analysis (MSA)

The internal nonvisible cracking is expected to be generated by solving the fictitious set-up.

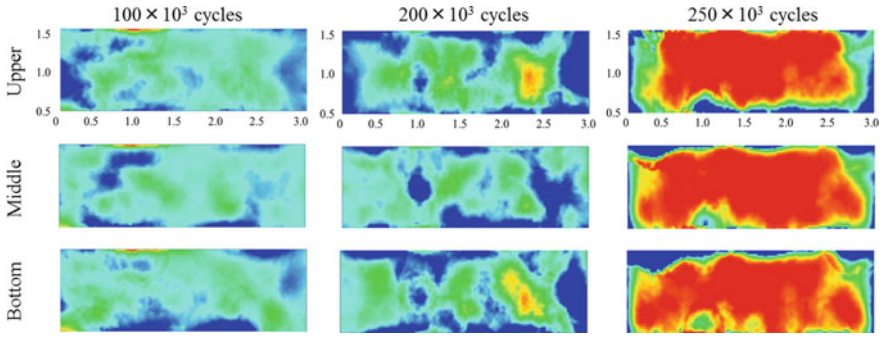


Fig. 5 Velocity distributions in RC slab during wheel loading

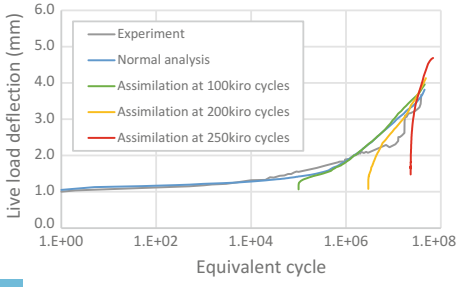
$$E_0 K_c = f(v) \tag{9}$$

where, E_0 is initial compressive stiffness of concrete, K_c is fracture parameter and v is elastic wave velocity. Elastic wave tomography recognizes cracks as the locations where elastic wave velocity is small. But the directions of cracks are not explicitly estimated because isotropy is assumed in the backward analysis. Non-uniformly forced initial damage creates the most stabilized solution with regard to the fracture parameter K_c in the multi-scale modelling [4]. The square root function is found to be suitable to transfer inspected data of the velocity v to the set-up fracture parameter K_c.

$$K_c = (v/v_0)^{0.5} \tag{10}$$

Figure 6 shows the live load deflections versus equivalent cycles of 98 kN load estimated by multi-scale analysis using the velocity data obtained by the 3D tomography at 100, 200 and 250 thousand cycles, respectively. As shown in the assimilation of Fig. 6, the live load deflections are consistently converging to the empirical one as a reference, implying that the assimilation analysis could predict the substantial live load deflection compatibly with any velocity of load cycles. Also, the strain distribution of multi-scale analysis is equivalent to that of referential analysis as shown in Fig. 7.

Fig. 6 Live load deflection



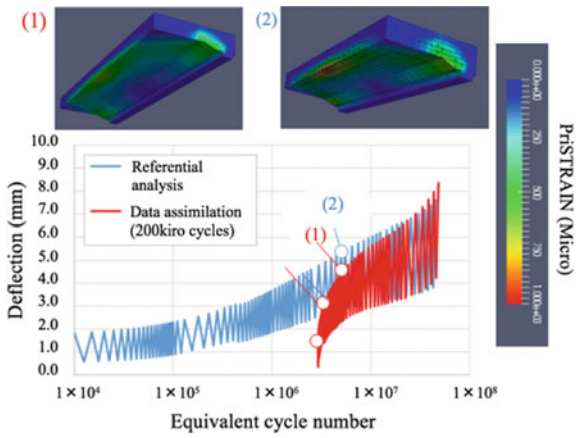


Fig. 7 Strain distribution after 5.42 million cycles loading in referential analysis and data assimilation of multi-scale analysis

The damage evolution and progress in the RC slab during wheel loading test could be detected by means of 3D elastic wave tomography and the future live load deflections, corresponding to the remaining service life, were successfully estimated for a reinforced concrete slab using velocity data obtained by 3D elastic wave tomography assimilating to the multi-scale analysis.

4 Evaluation of Fatigue Damage by AE Tomography Using Acceleration Measurement System in Steel Plate-Bonded RC Slabs

In conventional AE measurements, sensors and measurement system are used at relatively high frequency bands, requiring high cost with many AE sensors aiming at micro damage generating AE sources of high frequency. Presently however, the damages which need to be clarified for the first are relatively macroscopic damages belonging to an acceleration range of at most a couple of tens of kHz. Thus, the AE tomography procedure using an acceleration measurement system is developed for the macroscopic damages. Besides, as the attenuation rate of the acceleration waves are far smaller than of AE waves, a reasonable monitoring system with a small number of sensors, leading to lower cost and wider measuring areas than of AE system, can be possible in the acceleration monitoring system. Steel plate-bonded RC slabs is the measurement target in this study. The steel plate is bonded by epoxy adhesive with temporal installation of steel anchor bolts. Many of this type of RC slabs exist on the highway in Osaka, Japan.

As some areas of interfaces between the RC and steel plates showed delamination, 32 acceleration sensors, having effective frequency range from 3 to 20 kHz with sampling rate of 200 kHz, were set on the anchors installed deeply enough in the concrete (up to 5 cm depth in the RC slab) to detect elastic waves. The sensor

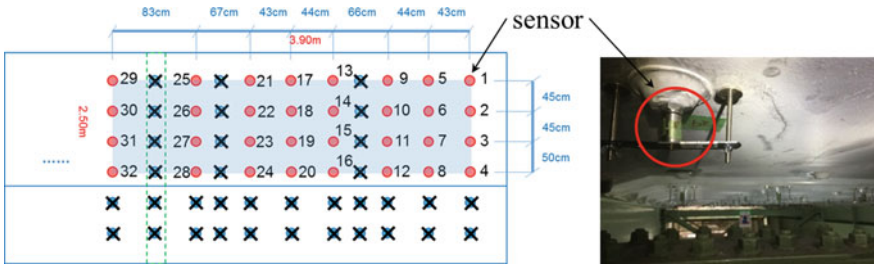


Fig. 8 Sensor arrangement (left) and installation of sensor to an anchor (right)

arrangement is shown in Fig. 8. A result of the AE tomography using 68 AE sources during 10 min measurement is shown in Fig. 9. 2D AE tomography was applied in this case. To observe the damage, epoxy injections were performed at the locations shown by cross symbols in Fig. 9. Specifically in the injection method, 5 mm dia. boring was performed to inject coloured epoxy agent, followed by 8 mm dia. boing to install a camera. The results of the hole observation were divided into three types such as lateral cracks, deterioration and sound area as shown in Fig. 9. It is certain that lateral cracks are found on the low velocity zone, sound situations are observed in the high velocity zone, and in the areas between high and low velocities, deterioration was confirmed. Accordingly, the results provided by the AE tomography using the acceleration monitoring system showed good agreement with actual damage situation, suggesting much wider measurement with less number of sensors would be possible using the acceleration monitoring system.

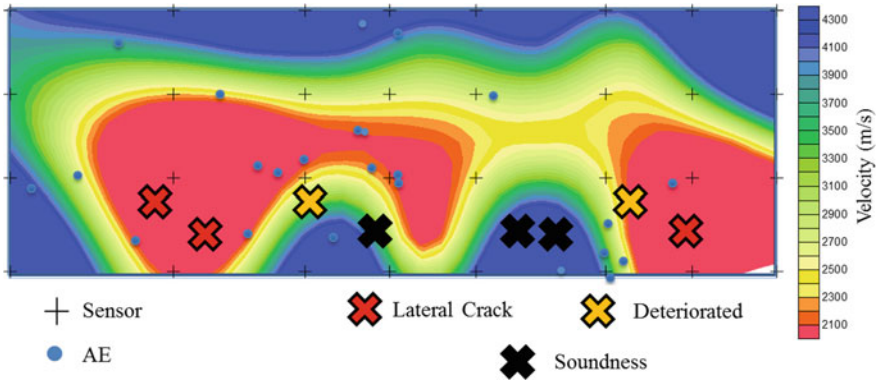


Fig. 9 Two D velocity distribution and conditions expected by epoxy injection

5 Conclusions

Regarding the application of AE tomography for detecting the internal fatigue damage of the in situ RC slab, the following was clarified. In the wheel loading test, it was confirmed that the internal fatigue deterioration occurred from a certain location as a low velocity distribution of 2500 m/s or less, and then spread out as a whole. In addition, data assimilation of multi-scale analysis and velocity distribution can evaluate remaining life in any stage. As for the steel plate-bonded RC slabs, the acceleration monitoring system was employed for AE tomography through detection by anchors. The results provided by the AE tomography using the acceleration monitoring system showed good agreement with the actual damage situation, suggesting much wider measurement with less number of sensors would be possible by the acceleration monitoring system. In general the ultrasonic range i.e., over 20 kHz would be more sensitive to micro-cracks than of the acceleration range; however in this study the detail relation between employed frequency and the scale of damage/crack could not be clarified. This would be introduced by a follow-up paper together with a proposal of reasonable sensor arrangement. Thus, although the system has still room for improvement, advanced NDT which can implement the effective maintenance and management under limited budget could be possible by the suggested tomographic approaches.

Acknowledgements A part of this study was financially supported by Council for Science, Technology and Innovation, "Cross-ministerial Strategic In-novation Promotion Program (SIP), Infrastructure Maintenance, Renovation, and Management" granted by JST.

References

1. Kobayashi Y, Shiotani T, Shiojiri H (2006) Damage identification using seismic travel time tomography on the basis of evolutional wave velocity distribution model. In: Proceedings of structural faults and repair 2006 (CD-ROM)
2. Kobayashi Y (2012) Three-dimensional seismic tomography with tetrahedral element on isoparametric mapping. Int J Struct Eng 3(1/2):37–47
3. Kobayashi Y, Shiotani T (2016) Computerized AE tomography, innovative AE and NDT techniques for on-site measurement of concrete and masonry structures. Springer, pp 47–68
4. Maekawa K, Pimanmas A, Okamura H (2003) Nonlinear mechanics of reinforced concrete. Spon Press, London
5. Matsui S (1996) Life time prediction of bridge. J JSSE Japan 30(6):432–440
6. Schubert F (2004) Basic principles of acoustic emission tomography. J Acoust Emission 22:147–158
7. Shiotani T, Aggelis DG, Makishima O (2009) Global monitoring of large concrete structures using acoustic emission and ultrasonic techniques: case study. J Bridge Eng 14(3):188–192

8. Shiotani T, Okude N, Momoki S, Kobayashi Y (2011) Proposal of AE tomography for assessment of infrastructures. In: Proceedings of 18th AE conference, JSNDI, pp 39–42 (in Japanese)
9. Shiotani T, Osawa S, Momoki S, Ohtsu H (2014) Visualization of damage in RC bridge deck for bullet trains with AE tomography. In: Advances in acoustic emission technology, Springer, pp 357–368

Coordination Between Maintenance and Production by Means of Auction Mechanisms for Increased Efficiency of Production Systems

Günther Schuh, Michael Kurz, Philipp Jussen and Florian Defèr

Abstract In order to cope with the challenges of an increased demand for flexibility, quality and availability of production, maintenance measures provide a major competiveness factor for manufacturing companies. Yet, interdependencies between maintenance and production activities as well as differing target systems within the functional units of an enterprise, especially production and maintenance, raise needs for extended coordination efforts. This paper aims to develop an innovative approach for the coordination between maintenance and production activities for industrial production companies. To achieve this, the novel coordination mechanism is used. It helps to achieve maximised operational availability—for a maximised output of the production system at optimal costs. Based on the developed model, the present paper identifies findings regarding the impact of different maintenance strategies on the medium-term economic efficiency of the production system.

1 Introduction

Maintenance activities are all technical and associated administrative actions intended to maintain or restore a specific condition of an item or product, in order for it to be used for its required function [15]. The relevance of maintenance tasks for producing companies in high-wage countries such as Germany has been fre-

G. Schuh (✉) · M. Kurz (✉) · P. Jussen (✉) · F. Defèr (✉)
FIR - Institute for Industrial Management at RWTH Aachen University,
Campus-Boulevard 55, 52074 Aachen, Germany
e-mail: guenther.schuh@fir.rwth-aachen.de

M. Kurz
e-mail: michael.kurz@fir.rwth-aachen.de

P. Jussen
e-mail: Philipp.jussen@fir.rwth-aachen.de

F. Defèr
e-mail: florian.defer@fir.rwth-aachen.de

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_54

quently outlined and intensively discussed in economic science and business practice. Depending on the industry sector, 15–40% of the production costs account for production-related maintenance activities, making it a significant factor influencing the competiveness and productivity of manufacturing companies. An annual sales volume of approximately €220 billion in the maintenance sector further underlines the high relevance of maintenance services from a macro-economic perspective [1]. Due to rising production complexity, volatility of order inflows, a high variance of products with simultaneously decreasing planning horizons a flexible and well performing maintenance management becomes rapidly more important to remain competitive and highly productive of industries in this era of high technology [4, 15, 20].

The maintenance management aims to maximise production availability and at the same time minimise the risk of machine failures and maintenance/operating costs [1]. Despite providing a key strategic success factor, maintenance management is still limited to the planning and processing of necessary or routine maintenance actions in many producing companies [16]. Beyond maintenance, every functional unit within an enterprise derives its actions from partial and deferring target cascades. Conflicting or even opposing target systems within the functional units of an enterprise, especially production and maintenance as the most interrelated ones, cause economical inefficiencies and raise needs for extended coordination efforts [2]. The production target system is orientated towards maximised output and fulfilment of predefined delivery targets, whereas the maintenance target systems aims toward securing needs-based machine capacities [14].

In industrial practice, maintenance actions are often postponed in favour of higher prioritised production targets. Usually, companies have difficulties measuring the value of a maintenance activity and comparing it with valuable production activities [19]. A lack of understanding for the importance of inter-functional target systems stresses the need for effective and efficient coordination mechanisms.

2 Objectives

In previous research the focus was either on the coordination between the opposing goals of maintenance and production by dynamic stochastic multi-criteria optimization models [7, 9] or in the auctioning as an important market tool for different applications e.g. supply chains [5, 13].

The objective of this paper is to develop a market-driven coordination mechanism, enabling producing companies to overcome local optima in the scheduling of maintenance and production activities by trading machine capacities. Therefore, the current condition of all installed machine parts and components as well as the planned machine utilisation shall be taken into consideration. Based on this information, each maintenance activity has to be assigned with a value by using trading mechanism, see Fig. 1.

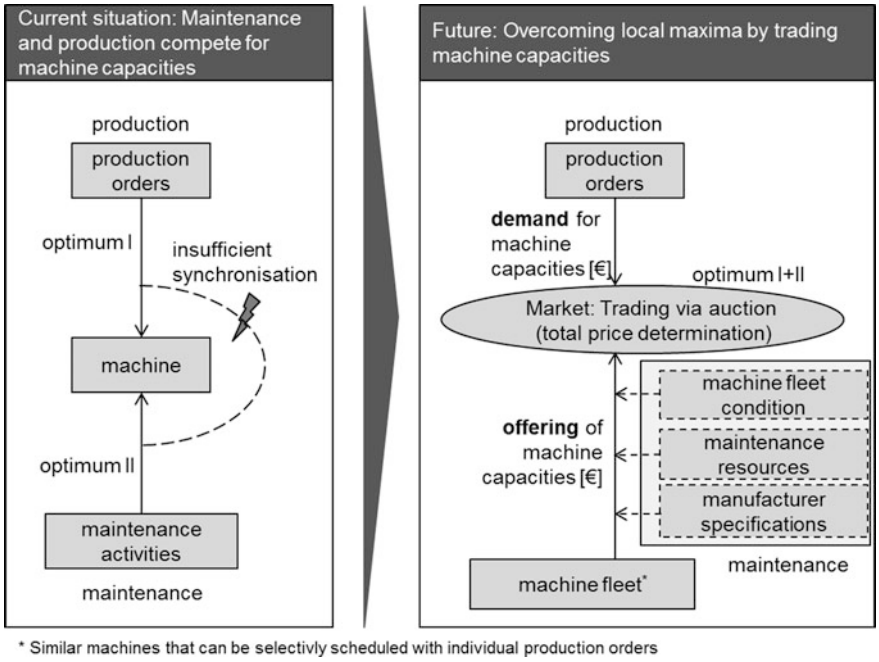


Fig. 1 Target picture—overcoming local maxima by trading machine capacity [10]

As a result, the best possible operation planning will be systematically determined to maximise operational availability subject to a maximised output of the production system at minimal costs. Based on the initial situation and objective target, the research question to be answered can be described as follows:

How should a model for maintenance and production coordination by means of auction mechanism be designed?

3 Theoretical Background

Interdependencies can be identified as the root cause for coordination needs [12]. A commonly used definition of interdependencies was established by Malone and Crowston in the late 1980s and early 1990s. According to them, interdependencies can be distinguished into a sequential dependency between activities/tasks (*prerequisite interdependence*), the simultaneous creation of a common object causing dependency (*simultaneity interdependence*) or the dependencies arising from simultaneous access on resources by different activities/tasks, causing assignment and occupancy conflicts (*shared resource interdependence*) [11].

The descriptive model provides the framework for the generic structure of the coordination mechanism, see Fig. 2.

The partial *coordination area model* includes all involved parties that cause coordination needs. They are described with their individual goals such as activities and resources. The *coordination area* consists of **machines**, **maintenance** and **production orders**. Each model component itself is associated with characteristics and respective variables. In this paper, **machines** for manufacturing constitute as limiting resources. Both the production and maintenance processes need to have access to machines in order to be able to execute their activities and thereby fulfil their goal sets.

Considerable characteristics for the coordination are *machine-operating conditions*, *required service cycles*, *machine wear curves* and the *machine specific wear boundary values.* **Maintenance** contains all activities that can be derived from the four basic principles of maintenance. These maintenance task types (characteristic)

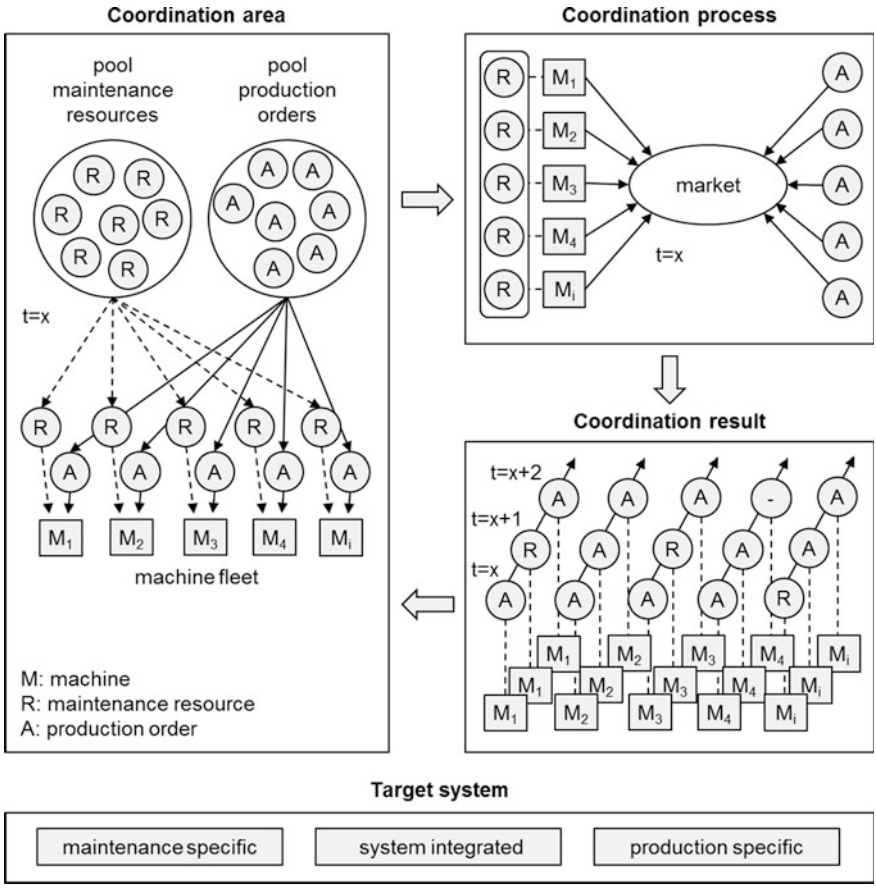


Fig. 2 Framework for the generic structure of the coordination mechanism [10]

serve as model variables and include *servicing*, *inspection*, *repair* and *improvement*. Other considered characteristics are *employee roles*, *employee qualifications*, *personnel and material costs*. Value drivers in **production** processes are machines, more specifically the process of machine fabrication and manufacturing. This stage of the value-added chain can be characterised by the *level of demand* coming from the following value-added chain stage, the *demand curve*, the related *variation and volatility* and possibly *contractual penalty* in case of poor performance.

The *coordination process* describes the second partial model. It contains procedures and applicable codes for handling coordination needs. The underlying trade mechanism for the coordination process is the "market". The definition of market results depends on (1) the traded goods, equal goods are traded on the same market; (2) how the market operates; (3) how the pricing of goods is executed. The model characteristics are differentiated into *market-specific characteristics*, *order-specific characteristics and pricing-specific characteristics*.

The process output is described in the *coordination result*, which defines a solution to meet the coordination need. On the one hand, the determined solution solves the coordination task for time t = x and on the other hand immediately effects the *coordination area* at the time t = x + 1. Due to the periodic changing frame conditions of the *coordination area*, the model is to be understood as an iterative, continuous process.

4 Methodical Approach

To gain a sufficiently valid understanding of functional correlations/interdependencies and conduct further in-depth research, the descriptive model is transferred into a *system-dynamic simulation* model. The method of *system-dynamic modelling* allows holistic analysis of dynamic, nonlinear and complex systems over time [6, 18]. In the concrete application of this case, the model allows to identify and analyse effects of varying influential factors (e.g. employee count, initial machine conditions) on the target system (e.g. economic efficiency of the production system, utilization rate). Shown in Fig. 3 is the dynamic hypothesis in terms of a cause and effect relationship. The model for the coordination between maintenance and production by an auctioning trading mechanism details the broad framework given in Fig. 2. It provides complexity-reduced information about effects and revision cycles in the simulation model.

The validation of the simulation model is necessary to verify whether the model describes the real-life system with sufficient accuracy or not [3, 17]. The validation shall be conducted on a theoretical basis and address the *model's structure*, *parameters* and *system behaviour*. Because the model is intended to be applicable in industrial practice to enable enterprises for a better understanding and optimisation of coordination between maintenance and production. It is essential not only to ensure the scientific model accuracy, but also the usability in industrial practice. The *industrial validation* shall be achieved by defining und studying use cases.

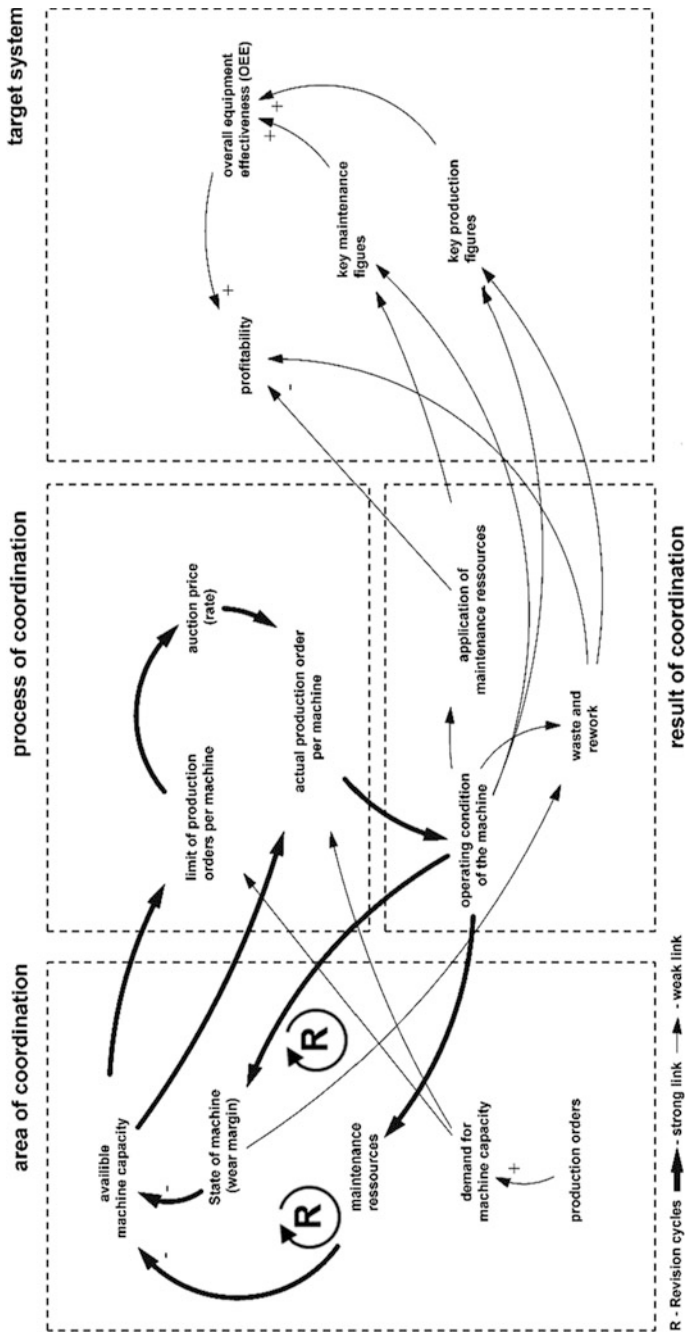


Fig. 3 Dynamic hypothesis model for the coordination between maintenance and production by an auctioning trading mechanism [10]

Case studies are understood as a method that refers to a very limited variety/amount of highly comparable cases. It aims to gain understanding of those cases, thus theories can be derived, modified or validated [8]. A common instrument for the conduct of case studies for a case-specific validation in the industrial practice are qualitative partially structured interviews with experts. For this model, expert interviews were conducted with manufacturing companies from steel, automotive supplier and drivetrain technology branches.

5 Results

In order to carry out specific analysis with this model, one option is to examine the cause-effect relationships between headcount and qualification level of maintenance employees on the one hand, and the main performance indicator (*medium-term economic efficiency of the production system*) on the other hand. For this purpose, it is necessary to define specific scenarios. Each scenario stands for a use case with an individual set of parameters (combination of input variables). Following the selection of scenarios, simulation runs are set up. The following three scenarios are considered: (i) *baseline scenario*, (ii) *cost reduction scenario*, (iii) *investment scenario*. The baseline scenario has to be seen as the current state of the organisation. The cost-reduction scenario and the investment scenario are two contradictory options for action. The characteristics of the three scenarios are described in Table 1.

The scenarios are interlinked with the most important industrial performance indicator for production and maintenance organisations, which was identified in interviews with industrial partners: *medium-term economic efficiency of the production system*, see Fig. 4. The medium-term economic efficiency of the production system describes the ratio between revenue through produced goods (system output) and maintenance costs (labour costs, material costs, external service costs).

The simulation model has to be considered as a *strategic tool* to derive recommendations for action. It aims to strategically align maintenance activities with the optimum operating point. The following results describe effects and implications of the three scenarios on the medium-term economic efficiency of the production system, see Fig. 5. It can be noticed that both alternative scenarios differ

Table 1 Characteristics of baseline scenario, cost reduction scenario and investment scenario

Influential factors/ scenarios	Baseline scenario	Cost reduction scenario (%)	Investment scenario (%)
Count service technicians	Baseline	−20	+20
Count technical specialists	Baseline	−20	+20
Employee qualification	Baseline	−15	+15

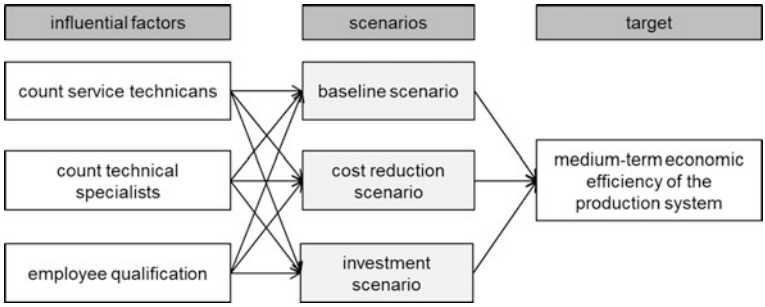


Fig. 4 Company-specific influential factors, scenarios and key performance indicator for the system-dynamic model [10]

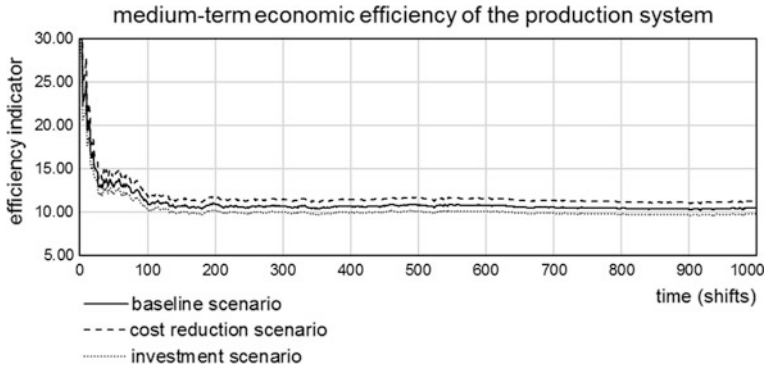


Fig. 5 Impact of scenarios on medium-term economic efficiency [10]

from the baseline scenario for this specific set of parameters. On the one hand, the medium-term economic efficiency for the investment scenario is lower than for the baseline scenario (−6.67% on average). On the other hand, the cost reduction scenario increases the baseline scenario efficiency indicator by 7.56% on average. Based on these finding, the derived recommendation for action for industrial practice should be to follow the strategic approach of the cost reduction scenario.

6 Conclusion

This paper points out the challenges for industrial production companies in the coordination between maintenance and production activities. It describes the theory of coordination and provides quantitative information and a deep understanding of existing interdependencies and their economic effects for both, maintenance and production. The paper provides a descriptive model of coordination by means of

auction mechanism. To gain a sufficiently valid understanding of functional interdependencies and conduct further in-depth research, the descriptive model was transferred into a *system-dynamic simulation* model. The simulation model, as a strategic tool for industrial practice, supports companies in their decision making. In a next step, further efforts should be made to implement first industrial use-cases to further validate the developed coordination mechanism on the basis of industrial practice.

References

1. Acatech (ed) (2015) Smart maintenance for smart factories—Mit intelligenter Instandhaltung die Industrie 4.0 vorantreiben, Munich
2. Al-Turki UM et al (2014) Integrated maintenance planning in manufacturing systems. Springer, Cham
3. Balci O (1994) Validation, verification, and testing techniques throughout the life cycle of a simulation study. In: Tew J et al (eds) Proceedings of the 1994 winter simulation conference, Orlando, USA, 11–14 Dec 1994. Society for Computer Simulation, p 215
4. Blameuser R, Galonske M, Gehmann S (2015) Gegenwart und Zukunft der technischen Instandhaltung—Die technische Instandhaltung im Zeitalter von Industrie 4.0. BearingPoint GmbH, Frankfurt
5. Fiala P (2016) Supply chain coordination with auctions. J Bus Econ 86:155–171
6. Forrester JW (2009) Some basic concepts in system dynamics. Massachusetts Institute of Technology, Cambridge [Online]. Available at http://www.cc.gatech.edu/classes/AY2013/cs7601_spring/papers/Forrester-SystemDynamics.pdf
7. Gössinger R, Kaluzny M (2013) Release of maintenance jobs in a decentralized multistage production/maintenance system with continuous condition monitoring. J Bus Econ 83(7): 727–758
8. Kaiser R (2014) Qualitative Experteninterviews—Konzeptionelle Grundlagen und praktische Durchführung. Springer, Wiesbaden, p 4
9. Kröning S (2014) Integrierte Produktions—und Instandhaltungsplanung und—steuerung mittels Simulationstechnik. Berichte aus dem IFW; 10/2014. Hrsg.: B. Denkena. TEWISS—Technik und Wissen, Garbsen 2014. Zugl.: Hannover, University, dissertation
10. Kurz M (2018) Koordination zwischen Instandhaltung und Produktion mittels Handelsmechanismus. Apprimus Verlag, Aachen
11. Malone TW, Crowston K (1990) What is coordination theory and how can it help design cooperative work systems. In: Halasz F (ed) Proceedings of the 1990 ACM conference on computer-supported cooperative work, Los Angeles, USA, 7–10 Oct 1990. Hrsg., pp 357–370
12. Malone TW, Crowston K (1994) The interdisciplinary study of coordination. ACM Comput Surv 26:87–119
13. Mason AN (2015) Development of horizontal coordination mechanisms for planning agricultural production. Tempe, University, Dissertation, https://repository.asu.edu/attachments/164084/content/Mason_asu_0010E_15464.pdf. Latest access 25 Sept 2017
14. Mostafa S et al (2015) Lean thinking for a maintenance process. Prod Manuf Res 3:236–272
15. Mushiri T (2015) Machinery maintenance yesterday, today and tomorrow in the manufacturing sector. In: Ao SI et al (eds) Proceeding of the world congress on engineering asset management, London, UK, 1–3 July 2015. Bd. 2. Newswood Limited, Hong Kong, pp 1149–1153

16. Patel M (2015) The future of maintenance. Infosys Ltd., Bangalore [Online]. Available at https://www.infosys.com/industries/aerospace-defense/white-papers/Documents/enabled-predictive-maintenance.pdf
17. Rabe M et al (2008) Verifikation und Validierung für die Simulation in Produktion und Logistik. Springer, Berlin, p 15
18. Shannon RE (1998) Introduction to the art and science of simulation. In: Medeiros DJ et al (eds) Proceedings of the 1998 winter simulation conference, pp 7–14
19. de Souza JB et al (2014) The maintenance function in the context of corporate sustainability: a theoretical-analytical reflexion. In: Grabot B et al (eds) Advances in production management systems. Innovative and knowledge-based production management in a global-local world. IFIP advances in information and communication technology. Springer, Berlin, Heidelberg, pp 222–229
20. Srinivasa Rao M, Naikan VNA (2014) Reliability analysis of repairable systems using system dynamics modeling and simulation. J Ind Eng Int 10

Quantification of Valve Severity in Reciprocating Compressor by Using Acoustic Emission Technique

H. Y. Sim, R. Ramli, A. A. Saifizul and M. F. Soong

Abstract Failures in a reciprocating compressor are often caused by the valves, as they are the major moving components in the reciprocating compressor. The acoustic emission (AE) technique has been employed by condition-based monitoring personnel to detect valve problems prior to failure. However, the extent to which the valve should be replaced, or more specifically, the relationship between AE measurement and valve severity remains unclear. This study attempted to establish a method to quantify the severity of valve problems in a reciprocating compressor by examining the relationship between AE root-mean-square (rms) value and the rotational speed at increasing severity. The study involved the acquisition of AE signals in different simulated valve conditions with increasing severity and the correlation of these signals with valve severity at its corresponding valve events. The results of the study showed that AE signal is linearly related to the 4th power of speed. Besides, the signal showed large decrement at increasing severity. It is believed that the study can help to quantify valve severity level with AE signal in the near future.

1 Introduction

Acoustic emission (AE) is the release of rapid and localised energy in the form of transient elastic waves as a result of deformation or dislocation within or on the surface a material. It is often found in materials under stress, in processes such as

H. Y. Sim · R. Ramli (✉) · M. F. Soong
Advanced Computational and Applied Mechanics (ACAM) Research Group, Department of Mechanical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
e-mail: rahizar@um.edu.my

A. A. Saifizul
Department of Mechanical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
e-mail: saifizul@um.edu.my

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_55

rubbing, cavitation and leakage. Numerous studies showed that the impacts of valves can be detected through AE. Nevertheless, studies relating AE with valve severity are less common. Wang et al. [4, 5] combined AE and simulated valve motion to investigate valve events at various discharge pressures and valve conditions. They concluded that amplitude of burst AE signal and advanced/delayed degrees of valve events can be the indicator for valve severity. However, the thresholds of these data are not specified. This finding can only be implemented if the baseline data of a normal valve is given.

This study intends to quantify the severity of valve leakage at different rotational speeds. The first part of the study examines the correlation between AE rms value and the rotational speed at different valve conditions and time segments. Segregation of one complete revolution of AE signals into different time segments was introduced by past researchers [3] to relate the AE rms value to its corresponding valve movement. As valve impact velocity increases linearly with rotational speed [6], this study can deduce behaviour of AE signal with increasing valve impact velocity. The second part of the study investigates the effect of valve severity on AE rms value at different speeds and time segments. This enables to deduce the severity of a leaking valve by observing the AE rms value.

2 Theoretical Background

As the piston moves downwards from the top-dead-centre (TDC), the lower pressure created in the cylinder causes fluid to flow into the cylinder through the suction valve. This sudden introduction of mass into the cylinder generates monopole radiation [2]. The sound power generated by this monopole source is defined as [1]:

$$P_s \approx \rho v^2 d^3 (v/c)(v/d) \tag{1}$$

where P_s is the sound power (W), ρ is the fluid density (kg/m^3), v is the averaged velocity fluctuation (m/s), d is the size of the region of pressure fluctuation, d^3 is the small volume element (m^3), and c is the sound velocity in the medium (m/s).

Equation 1 can be simplified further as:

$$P_s \approx (\rho A/c)v^4 \tag{2}$$

where A is the flow area of valve (m^2). In this study, as the Mach number is small ($\ll$0.3), fluid density ρ can be assumed to be constant. Besides, as the variation of sound velocity in the medium c and flow area of valve A under different valve conditions are relatively small, it can be deduced that the sound power P_s is proportional to the fourth power of velocity fluctuation v^4, shown as:

$$P_s \propto v^4 \tag{3}$$

The AE signal measured in the study is in the form of u acoustic velocity (m/s). The acoustic velocity and pressure are related by:

$$p(x,t) = Zu(x,t) \tag{4}$$

where $p(x, t)$ is the acoustic pressure (Pa), Z is the specific acoustic impedance (kg/m^2 s^1), and $u(x, t)$ is the acoustic velocity (m/s). Since u and p are directly proportional, and the square of acoustic pressure p is directly proportional to sound power P_s, it can be deduced that:

$$U^2 \propto v^4 \tag{5}$$

where U is the AE rms value and v is the velocity fluctuation of fluid.

Fluctuation of fluid velocity is originated from the piston movement. Thus the fluid velocity is directly related to the piston velocity.

Figure 1 shows the schematic diagram of a reciprocating compressor's cylinder.

The piston displacement y from TDC can be derived as:

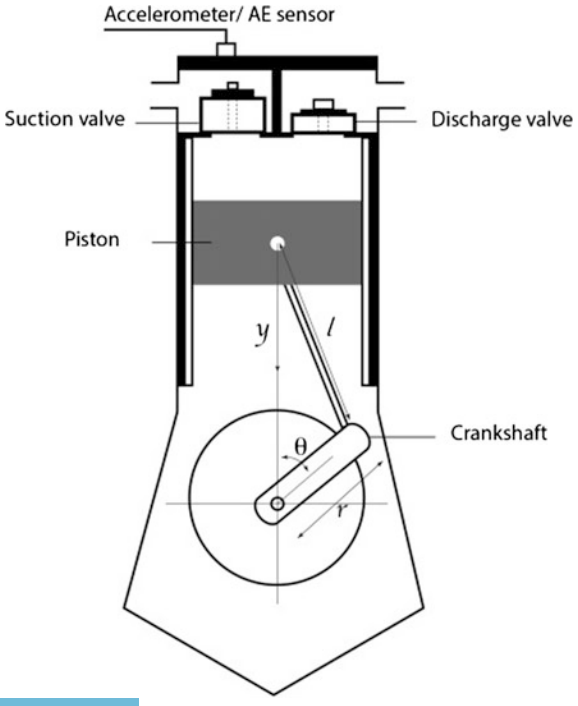


Fig. 1 Schematic diagram of a reciprocating compressor's cylinder

$$y(t) = r(1 - \cos\theta) + l\left[1 - \sqrt{1 - (r/l)^2 \sin^2\theta}\right] \tag{6}$$

where $y(t)$ is the piston displacement (m), r is the crank radius (m), l is the length of connecting rod (m), $\theta(t)$ is the crank angle (°).

By differentiating Eq. 6 with respect to time:

$$\dot{y}(t) = r\omega \sin\theta\left[-1 + \frac{r}{l}\cos\theta / \sqrt{1 - (r/l)^2 \sin^2\theta}\right] \tag{7}$$

where $\dot{y}(t)$ is the average piston velocity (m/s) and $\omega(t)$ is the average angular velocity (rad/s). At the same crank angle, it can be said that the average piston velocity is directly proportional to the average rotational speed of compressor. Thus, the theoretical relationship between AE rms value and rotational speed can be deduced as:

$$U^2 \propto \omega^4 \tag{8}$$

where U^2 is the square of AE rms value.

This theoretical relationship forms the basis of the first part of this study, where the empirical relationship between AE signal and rotational speed is to be justified.

3 Methodology

3.1 Experiment Set-Up

The experimental set-up of this study consists of a single stage, air-cooled, two-cylinder reciprocating air compressor connected to a variable frequency drive (VFD), as displayed in Fig. 2. The rotational speed of test compressor is varied from 450 to 800 rpm by adjusting the input frequency of motor through VFD.

The test subject in this study is the suction valve plate. A wide band AE sensor (100–900 kHz) is mounted on the suction valve cover of test compressor. The AE signal will be acquired at 1 MHz simultaneously with the pulses generated from a laser tachometer directed to the flywheel, where the high pulse represents the piston position at the TDC. By referring to the pulse from the tachometer, valve events observed from AE signal can be analysed at their corresponding crank angle. This study is conducted based on the assumption that the rotational speed of the compressor is constant for each measurement.

To simulate leakage in the suction valve, the valve plate is ground with a dent of 6 mm diameter. The severity of valve leakage can be emulated by increasing the number of dents on the valve plate. A single leak valve will have one dent while two dents can be seen in a double leak valve, as displayed in Fig. 3.

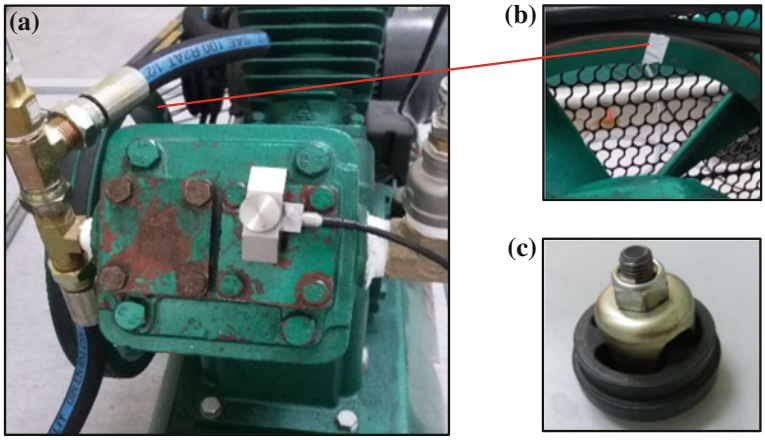


Fig. 2 Experimental test rig with **a** AE sensor mounted on valve cover **b** reflective tape for tachometer **c** suction valve

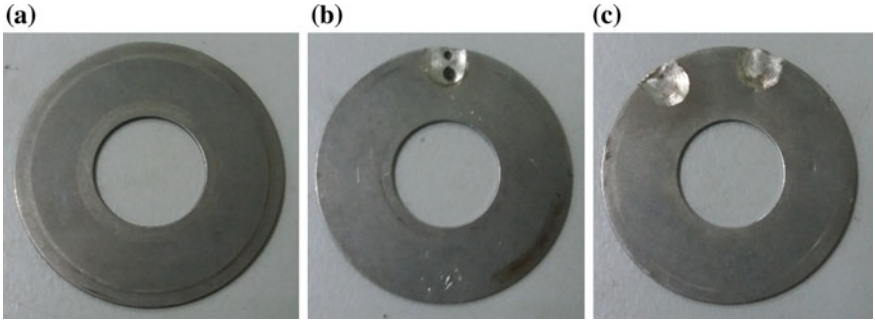


Fig. 3 Suction valve plate at **a** normal **b** single leak **c** double leak condition

3.2 Data Preparation

The experiment is conducted at suction valve temperature of approximately 40–50 °C and atmospheric pressure. To study the effect of valve leakage on AE signal, the rms value is computed according to physical valve events, namely the discharge valve closing (T1), suction valve opening (T2), suction valve closing (T3), discharge valve opening (T4), and post-discharge valve opening (T5). It can be seen from Fig. 4 that T1, T2, T3, T4, and T5 correspond to crank angle from 0–30°, 30–120°, 150–240°, 270–330°, and 330–360° respectively. A total of 30 samples are acquired for each set of data under different valve conditions, time segments and rotational speeds.

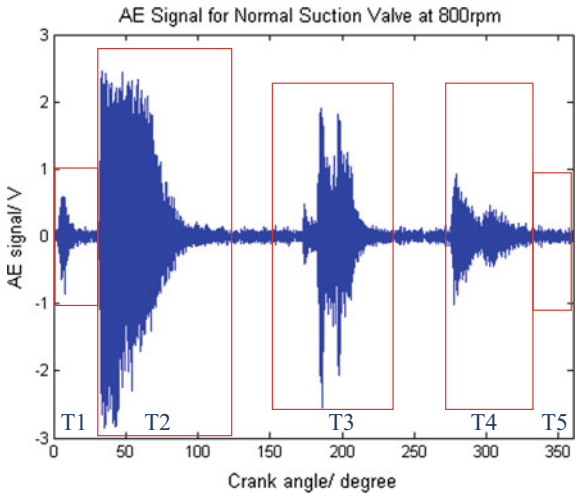


Fig. 4 Segregation of AE signal into different time segments

4 Results and Discussions

4.1 Effect of Rotational Speed on AE Signal

The square values of AE rms obtained between 450 and 800 rpm are compared at normal, single leak and double leak valve conditions for each time segment. The plots of T1 to T3 are displayed in Figs. 5, 6 and 7 respectively. A linear trend line can be observed in most plots, except T1 under double leak condition and T2 and T3 under normal condition. For T1 under double leak condition and T2 under normal condition, the trend line showed its non-linearity clearly at 800 rpm, as shown by the sudden drop of data in T1 double leak and gradual increase of data in T2 normal condition. These are caused by the limitation of the air compressor

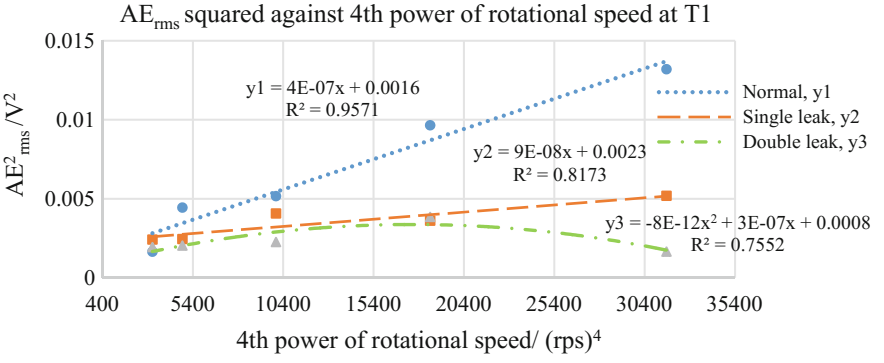


Fig. 5 Square of AE rms at T1 under different speeds and conditions

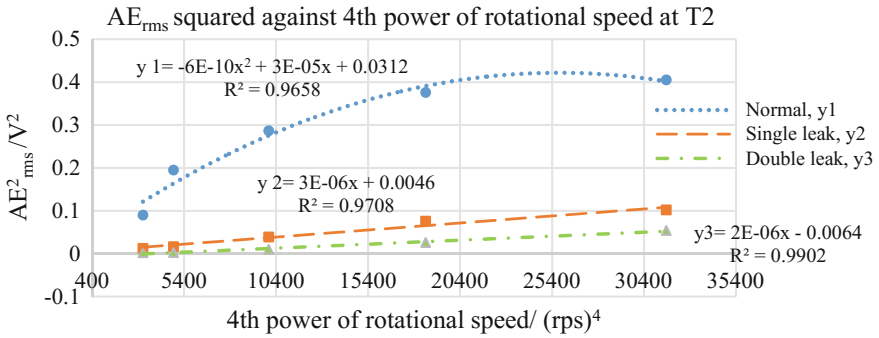


Fig. 6 Square of AE rms at T2 under different speeds and conditions

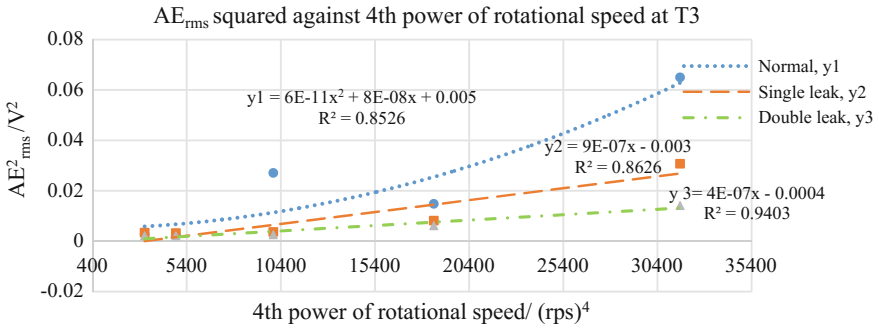


Fig. 7 Square of AE rms at T3 under different speeds and conditions

design as the test compressor can only afford a maximum achievable suction flow rate. The less proportional increment of suction flow rate measured when the compressor speed approaches 800 rpm further supports this deduction. Still, more measurements need to be taken to confirm the deduction. As T3 is the duration for the suction valve closing, a majority of the sounds recorded are originated from valve impacts, especially under normal valve condition. Therefore, a non-linear relationship can be seen for T3 under normal condition. Meanwhile, since there are fluids leaking through the single leak and double leak valve at T3, both conditions displayed a linear trend line with R-squared value slightly lesser than that in T2.

Most of the results of this study showed that AE rms squared is linearly related to the 4th power of compressor speed below 800 rpm, especially at T2. The results deviate slightly from the theoretical deduction of direct proportionality between the 2 variables, partly because of the assumptions made during the theoretical derivation, namely the direct relationship between fluid velocity and piston velocity. In fact, the fluid and piston velocity may not be directly proportional. There is always a distribution of fluid velocity across the inlet manifold due to the shear stress developed at the surface of intake manifold. Moreover, the change of

medium (from gas to solid) during sound propagation is neglected during the theoretical derivation.

The linear relationship of results showed that AE rms value is affected by rotational speed. Thus, one cannot conclude valve failures by just referring to the AE rms value without considering compressor load/speed.

4.2 *Effect of Valve Severity on AE Signal*

Comparisons of time synchronous averaged AE signals at increasing valve fault severities are displayed in Figs. 8 and 9. Under rotational speed of 450 rpm, impact at T3 is less obvious for single leak condition. For double leak valve, impacts at both T2 and T3 are imperceptible. This may indicate improper sealing of suction valve plate during the compression cycle. As severity increases, fewer impacts and

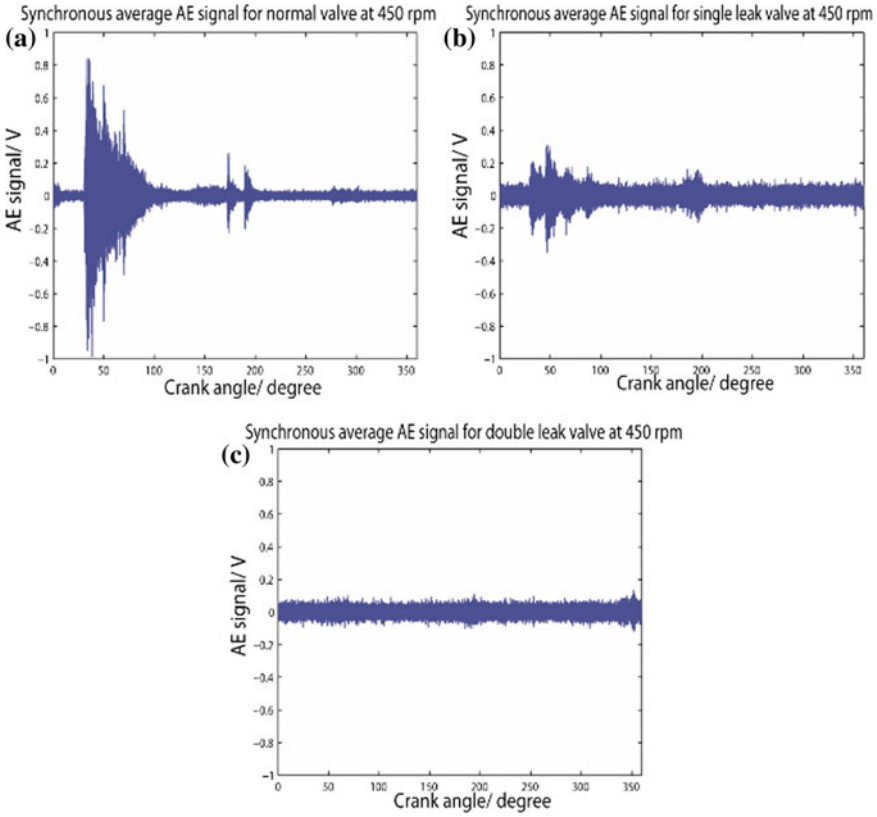


Fig. 8 Comparison of synchronous averaged AE signal at **a** normal **b** single leak **c** double leak valve condition under 450 rpm

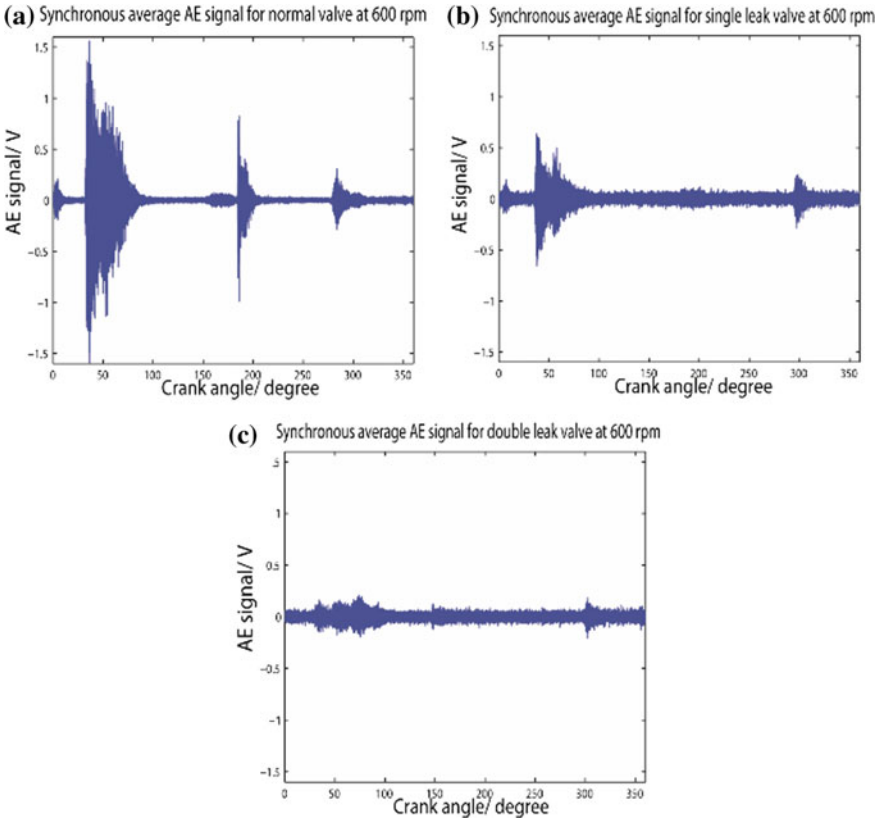


Fig. 9 Comparison of synchronous averaged AE signal at **a** normal **b** single leak **c** double leak valve condition under 600 rpm

more leaks are produced, thus generating higher noise. The random noises buried the valve impacts and appeared as continuous AE signal in Fig. 8c.

As speed increases to 600 rpm, the impact at T4 becomes more obvious compared to 450 rpm. This impact is originated from the opening event of neighbouring discharge valve. Besides, the impact at T3 is absent from Fig. 9b, c, while more noise dominates the plot. Similar plots can be seen for 500, 700 and 800 rpm.

It is hard to discern the valve condition by just visually inspecting the shape of the signal. Therefore, AE rms squared is tabulated according to speed and time segments, as shown in Table 1. It can be seen that AE rms squared decreases with increasing severity at T1, except under 450 and 700 rpm. The decrement of AE value is more drastic at T2, where it decreases almost exponentially from normal to double leak condition. As this segment indicates the suction valve opening event, it can be deduced that the opening impact decreases tremendously with increasing severity. Meanwhile, the post-discharge valve opening represented by T5 can be an indicator for valve condition, where the normal valve has values ranged between

Table 1 Comparison of AE rms squared at T1, T2, and T5 under different speeds and valve conditions

450 rpm								
T1			T2			T5		
Normal	S leak	D leak	Normal	S leak	D leak	Normal	S leak	D leak
0.0016	0.0024	0.0019	0.0902	0.0125	0.0022	0.0005	0.0027	0.0034
500 rpm								
T1			T2			T5		
Normal	S leak	D leak	Normal	S leak	D leak	Normal	S leak	D leak
0.0044	0.0024	0.0020	0.1951	0.0163	0.0026	0.0006	0.0025	0.0027
600 rpm								
T1			T2			T5		
Normal	S leak	D leak	Normal	S leak	D leak	Normal	S leak	D leak
0.0052	0.0041	0.0023	0.2865	0.0389	0.0099	0.0012	0.0029	0.0035
700 rpm								
T1			T2			T5		
Normal	S leak	D leak	Normal	S leak	D leak	Normal	S leak	D leak
0.0097	0.0036	0.0038	0.3758	0.0761	0.0262	0.0005	0.0031	0.0032
800 rpm								
T1			T2			T5		
Normal	S leak	D leak	Normal	S leak	D leak	Normal	S leak	D leak
0.0132	0.0052	0.0016	0.4054	0.1022	0.0542	0.0008	0.0031	0.0016

0.0005 and 0.0012, while the leaked valves have values ranged from 0.0016 and 0.0035. The higher AE value under leaked condition may indicate the escape of fluid through the suction valve plate during compression stroke as a result of failed sealing. Results from T3 and T4 are omitted as the AE values are not consistent with increasing severity.

Table 2 shows the comparison of AE rms squared ratio between T5 and T2 at various compressor speeds and valve conditions. Almost all ratios increase with increasing severity, except those at 800 rpm. Besides, for single leak and double leak condition, the ratio decreases with increasing speed, as the AE rms squared value at T2 increases with speed. Thus, the ratio can also be the indicator for valve leakage.

Table 2 Comparison of AE rms squared ratio between T5 and T2

Valve condition	Compressor speed (rpm)				
	450	500	600	700	800
Normal	0.0059	0.0029	0.0041	0.0014	0.0021
Single leak	0.2168	0.1541	0.0737	0.0412	0.0304
Double leak	1.5504	1.0326	0.3485	0.1203	0.0292

5 Conclusions

It is a common belief that variation in AE signals may indicate a change in valve condition. This study shows that a change in operating conditions especially rotational speed can give rise to a sudden change in AE signal, and the relationship between AE rms squared and 4th power of speed is linear at T2. Besides, valve leakage condition can be quantified by referring the AE rms squared value to the physical movement of valve. It is found that the AE rms squared value decreases almost exponentially with increasing severity at T2, especially under lower speed. The results show that AE rms squared value at T2 and the ratio of AE rms squared between T5 and T2 are the most convincing valve condition indicator. For future study, we can compare the results between AE and vibration signal for effective health monitoring. It is hoped that a generic mathematical model of valve condition can be constructed based on AE and vibration techniques, together with other operating variables such as speed, pressure and temperature of reciprocating compressor.

Acknowledgements The authors would like to thank the Ministry of Science, Technology, and Innovation of Malaysia (Project no. 03-01-03-SF1033) and Institute of Research Management and Monitoring (IPPP) from University of Malaya, Malaysia (Project no. PG233-2014B) for their financial support.

References

1. Morse PM, Ingard KU (1968) Theoretical acoustics. McGraw-Hill
2. Nored MG, Tweten G, Brun K (2011) Compressor station piping noise: noise mechanisms and prediction methods. Interim Report, Gas Machinery Research Council, Southwest Research Institute
3. Sim HY, Ramli R, Saifizul AA, Abdullah MAK (2014) Empirical investigation of acoustic emission signals for valve failure identification by using statistical method. Measurement 58:165–174
4. Wang Y, Gao A, Zheng S, Peng X (2015) Experimental investigation of the fault diagnosis of typical faults in reciprocating compressor valves. Proc Inst Mech Eng Part C J Mech Eng Sci 1–15
5. Wang Y, Xue C, Jia X, Peng X (2015) Fault diagnosis of reciprocating compressor valve with the method integrating acoustic emission signal and simulated valve motion. Mech Syst Signal Process 56–57:197–212
6. Wang Y, Xue C, Feng J, Peng X (2013) Experimental investigation on valve impact velocity and inclining motion of a reciprocating compressor. Appl Therm Eng 61:149–156

Data Quality in Asset Management—Creating and Maintaining a Foundation for Data Analytics

Allen Tam and Iris Kwan

Abstract Data analytics (in particular "Big Data" analytics) is one of the hot topics in recent digital technology innovations. There is an increasing trend and focus on the value data analytics can bring to the business. Internet of Things (IoT) and advancement in technology have resulted in even more data being created and assessed. There are many new tools and software developments that promise "silver bullet" type solutions, with advanced dashboard, reporting, graphic visualisation and predictive capabilities. Little attention is however given to the data quality aspect of the puzzle. Incorrect data can lead to lost opportunities, incorrect decisions, and costly enterprise system implementations. This paper provides an overview of the challenges contributing to master and transaction data quality issues in utility asset management and highlights the potential lost opportunities resulting from a lack of data integrity. The value of text analytics and reviewing of fault data is highlighted. The paper also offers a four step solution to data quality improvement.

1 Introduction

Data analytics (in particular "Big Data" analytics) is a hot topic in recent technology wave. There is an increasing trend and focus on the value data analytics can bring to the business. Internet of Things (IoT) and advancement in technology have made even more data being created and assessed. There are many new tools and software developed that provides "silver bullet" solution, with advanced dashboard, reporting, graphic visualisation and predictive capabilities. However, there is little

A. Tam (✉)
Relken Engineering, Melbourne, Australia
e-mail: allen.tam@outlook.com

I. Kwan
AusNet Services, Melbourne, Australia
e-mail: iris.kwan@ausnetservices.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_56

attention to the data quality aspect of the puzzle. Incorrect data can lead to loss opportunity and incorrect decisions being made.

Data quality is fundamental for building decision models that support evidence-based asset management [2]. It is also noted by Jardine and Tsang [2] that mathematical models themselves with incorrect data do not guarantee accurate insights and therefore may not support right decisions being made. Lin et al. [3] highlighted that data quality problem can result in severe negative consequences for an organisation. The study proposed a data quality framework for asset management aiming to improve data quality via targeted activities. In Sarfi et al.'s [4] study on utilities pointed out that the function of data and the critical data dependencies in various asset management systems.

Different models require different datasets. Previous work by Tam and Gordon [6] examined failure terminology for work orders data. Tam and Price [7] identified the key data requirement for maintenance optimisation models, where the data is captured and the priority ranking of significance for enterprise level optimisation. The data can be both master data and transactional data. The type of models and the optimisation objectives will drive the data quality requirement for specific parameters based on the sensitivity analysis pointed out by Tam and Price [7]. A study by Tam and Price [8] highlighted that data requirement for ensuring good condition-based maintenance (CBM) models outcome. In Dave et al. [1] work, it is highlighted that data granularity and data quality is vital in the modelling process and will have major impact to result quality.

This paper provides an overview of the challenges contributing to master and transaction data quality issues in asset management and highlights the potential loss opportunities due to inaccurate data. The value of text analytics and reviewing of fault data is highlighted. This paper also offers a four-step solution to data quality improvement with learnings from a utility industry.

2 Master and Transactional Data Quality

2.1 Master Data

Master Data are business objects that are shared across the enterprise, relatively static. In Asset Management, some examples of key master data objects are: the Functional Locations, Equipment and Material Master.

The challenges within equipment master data from a brown field site often suffer from missing/incorrect data fields such as Installed Date, Date of Manufacture, Manufacturer, Serial Number and Specification. This information is often only available during installation and commissioning time and if they are not captured during that time, it is often missing or incorrect. Sometimes with equipment that has a nameplate attached, some of this information may be available. However, these nameplates with information can become unreadable over time, especially if the equipment is installed outdoor.

A system configuration is the specific information that describes what the system is made up of. The data describe a system behaviour is configuration data. For example, in electricity network, protection devices have settings which describes how the device should operate during a fault, the data in these settings are configuration data. This data can sometime be found within Master data tables but they should not be confused as master data.

2.2 Transactional Data

Transactional data is data that describes an event. Work orders, notification, projects and faults response dispatch orders, material orders are transactional data.

For Transactional Data especially in the maintenance world, work orders and loss of availability records are the most important data objects for reliability analytics purposes. Work Orders contains data that describe what work has been done to the asset and when, and also the nature of the work (planned or unplanned). Most Computerised Maintenance Management System (CMMS) will have the capability to capture failure modes, effects and remedy codes. These data are often the most important field for reliability analytics purposes (if they are captured properly) as they tells you which asset has failed and when, also tells you how it failed and what causes it to fail.

Loss of Availability data is also a key transactional data. Loss of Availability or downtime system record event data where there is an outage or a stoppage in the production line or power generation. In some operations, operators will record the cause of a stoppage (or loss of availability) and record if any works have been done to ratify the issue.

2.3 Data Quality

A measure of data quality is by assessing the 5 key attributes—the 5Cs [5]. The 5Cs are:

- Clean—no error
- Consistent—no arguments about which version of data is correct
- Conformed—common, shareable for business use. Same datasets are used for decision making
- Current—up-to-date
- Comprehensive—all data needed is available regardless of where data comes from and its level of granularity

Achieving all the 5Cs (Clean, Consistent, Conformed, Current and Comprehensive) of data quality are challenging due to:

- Semi-integrated (or non-integrated) systems and suboptimal interfaces
  - Geospatial Information System (GIS)—e.g. GE Smallworld, ArcGIS
  - Enterprise Asset Management (EAM) and Enterprise Resource Planning (ERP)—e.g. Maximo, SAP
  - Outage management system (OMS) (e.g. PowerOn Fusion)
  - Customer information Systems (CIS)
  - SCADA
- Major systems implementation and changes: EAM and ERP systems changes and/or upgrade—changes of enterprise systems (e.g. from Maximo to SAP or vice versa)—common issues of data mapping and translation between systems
- Business Process Compliance—no or not-adequate business process in place. If process is in-place, business does not follow process
- Increased Volume—the amount of data collection requirement increased

Certain data fields are more important for the reliability analytics perspective; the key fields are listed in Table 1.

Table 1 Data needed for reliability analytics and maintenance optimisation

Data table	Data field	Data type	Description	Where the data is kept
Equipment	Equipment type	Master	To identify the equipment group	EAM
Equipment	Equipment install date	Master	To obtain time-to-failure (TTF)	EAM
Work order	Equipment	Transactional	To identify equipment that has failed	EAM
Work order	Failure date	Transactional	To obtain time-to-failure data	EAM
Work order	Work orders failure mode/ cause	Transactional	To identify equipment failure mode	EAM
Work order	Material cost	Transactional	To identify cost of work	EAM
Work order	Labour cost	Transactional	To identify cost of work	EAM
Measurement	Condition measurement records	Transactional	To identify equipment condition during inspection	EAM
Outage or down time system	Event downtime/ losses	Transactional	To establish losses due to a failure or unplanned outage	OMS
Financial	Equipment cost	Master	To obtain cost of replacement	ERP
Financial	Labour cost	Master	To obtain cost of planned and unplanned labour	ERP

3 Data Quality Impact to Analytics—An Example

This section provides an example of how data quality can impact on the analytics outcome and also how improving data quality changes the analytics outcome. Figure 1 is plot of raw data from a utility work order system showing the Pareto chart of all faults responses in a given year. The top category is "Unknown" and "Deterioration", which is too generic and unusable. If this chart is the only information an asset manager got to make a decision on investment, the first three projects will be:

- Data Analytics—a drill down on "Unknown" and Deterioration" to know which asset deteriorated and are not pick up by inspection and why there are so many unknowns
- Invest in installing Animal cover on poles
- Increase targeted tree cutting program

Figure 2 plotted the results after text analytics. Text analytics are carried out, reviewing the description field entered by the operator during the outages and all other information available with the order (including checking other systems). The analyst then reclassified the causes code after reviewing all information available. The cause—unidentified denotes that there is not data available in the system and N/A refers to event that should not be raised as a fault. If we remove these two categories and replotted the graph, as shown in Fig. 3, the top causes are: Tree, Age, Bird or Animal and Corrosion.

The data is for this analysis is obtained by filed workforce and operators. They form a very important part of data quality. It is important that their technical knowledge and field narratives forming part of the free text fields should be considered during design and implementation of the engineering asset management system change. Further training will also assist in improving data quality.

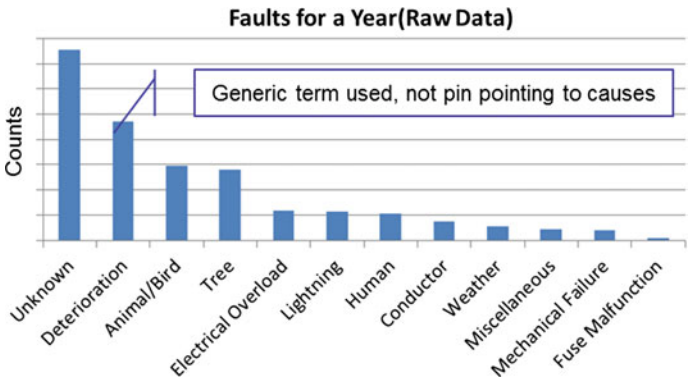


Fig. 1 Raw data from work order system—looking at causes of outage

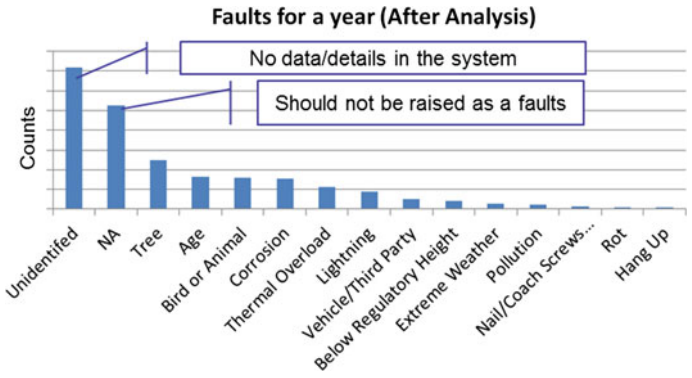


Fig. 2 Plotting the same data after text analytics—recoding the causes

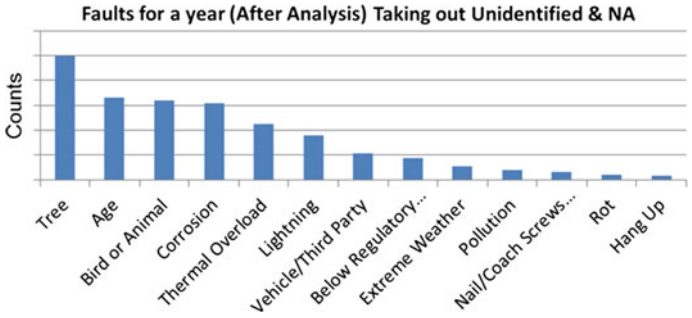


Fig. 3 Plotting the same data after text analytics—recoding the causes removing unidentified and N/A

4 Four Steps Solution

The paper has highlighted the challenges and importance of data quality. This section, the paper offers a Four Steps Solution to remediate and improve asset data quality. Data Remediation is the process to clean-up data of poor quality—data that do not meet the quality 5Cs (data that is incomplete, inconsistent, out-of-date, i.e. "wrong"). The four steps process is: Identify, Fix, Stop and Automate (IFSA).

- Identify what data is wrong/incomplete, how it gets into the system
  - Develop Data Standard
  - Develop and monitor data quality dashboard as per data standard
  - Identify and prioritise key data fields first—monitor data where its quality has major impact to the business

- Fix the data, ensure correct values are in the system
  - Perform analytics to obtains correct values
  - Analytics can involve desktop analysis and looking through old drawings
  - Carried out field visit to capture missing information
- Stop bad quality data from getting into the system
  - Apply data standard at point of data entry
  - Introduce mandatory fields
  - Reduce the number of data entry requirement, has a single source of data entry—like establishing the data management office for master data
  - Develop cascading drop down list in EAM/ERP system to allow the right codes being available for field workforce and operator
- Automate: Reduce Handling and Simplify Process. Automate if possible
  - Eliminate double data entry requirement—if data is needed in different systems, use an interface
  - Clearly define data master system—where if data is to co-exist in 2 systems, define which one is the master and synchronise data from the master to the rest using interface
  - Use as much automation for data entry as possible

    Bar code, QR code or RFID
    Use upload tool for mass data update (for example: SAP LSMW, WinShuttle)
  - Using analytics, develop tools/scripts to identify incorrect or conflicting data values within systems; then use advance analytics such as Bayesian theorem and Text Analytics to automatically correct data.

5 Conclusion

For the asset management practitioners, this paper provided a guide to data quality improvement and 4 steps approach. The paper highlighted the challenges and importance of data quality to creating and maintaining a foundation for data analytics and intelligence. In some data quality challenges, the paper has pointed out that in some instances—such as work order faults codes text analytics, the only way is to relay on experienced analyst. The analyst reads through tens of thousands of the comments free text fields and digging different database; discussion with operators to identify the appropriate classification of the true fault codes. With advancement in technology, it is possible to invest in "machine learning" approach where we can train the computer to pick up patterns on data. Data quality is pivotal to success in any data analytics and intelligence implement and is the most important aspect to deliver good value to the business. This is where the academics

can assist in research effort into simplifying text analytics and data capturing requirement in the field—using analytics, combining Bayesian theorem and test analytics to provide more meaningful data and interpretations on field data. A continuous improvement approach coupling with the 4 steps IFSA model as proposed in this paper is needed to ensure data quality is always one of the top priority in the enterprise performance dashboard.

References

1. Dave A, Oates M, Turner C, Ball P (2015) Factory eco-efficiency modelling: the impact of data granularity on manufacturing and building asset simulation results quality. Int J Energy Sect Manage 9(4):547–564
2. Jardine AK, Tsang AH (2013) Maintenance, replacement and reliability theory and applications, 2nd edn. CRC Press, FL
3. Lin S, Gao J, Koronios A (2008) A data quality framework for engineering asset management. Aust J Mech Eng 5(2):209–219
4. Sarfi RJ, Tao MK, Lyon JB, Simmins JJ (2012) Data quality as it relates to asset management. s.l., PES T&D 2012, May 2012, pp 1–5
5. Sherman R (2015) Business intelligence guidebook from data integration to analytics, 1st edn. Morgan Kaufmann, USA
6. Tam AS, Gordon I (2009) Clarification of failure terminology by examining a generic failure development process. Int J Eng Bus Manag 1(1):33–36
7. Tam AS, Price JW (2008) A generic asset management framework for optimisating maintenance investment decision. Prod Plann Control 19(4):287–300
8. Tsang AH, Yeung W, Jardine AK, Leung BP (2006) Data management for CBM optimisation. J Qual Maintenance Eng 12(1):33–57

Statistical Analysis for Wood Poles Using Sound Wood Measurements Data

Allen Tam, Iris Kwan and Mark Halton

Abstract Wood poles are one of the key structures for electricity distribution network. A key inspection technique used is a drill test to identify sound wood (good quality wood) left in a wood pole. A condition-based maintenance decision—replacement or reinforcement is made when the wood poles sound wood measurement is less than 30 mm. This paper presents the findings of a study of over 1.2 million records of data from about 200 thousands wood poles over a 15 years period. This paper aims to provide a view of wood poles P-F curve using sound wood deterioration data.

1 Introduction

Reliability Centred Maintenance (RCM) process determines the type of maintenance required to be applied to an asset—it answers the question "what type of maintenance action needs to be taken?" but not when to perform [5]. Maintenance optimisation models are therefore required to be developed to assist in determining optimal maintenance, replacements or inspections policy. Some of the popular models used are developed by Jardine and Tsang [5]. It is pointed out by Jardine and Tsang [5] that it is essential to have good data as inputs to maintenance optimization models.

In the RCM2 book [6], a chapter is dedicated to establish the ways to determine best interval for maintenance work. The chapter introduced the concept of P-F interval.

A. Tam (✉)
Relken Engineering, Melbourne, Australia
e-mail: allen.tam@outlook.com

I. Kwan · M. Halton
AusNet Services, Melbourne, Australia
e-mail: iris.kwan@ausnetservices.com.au

M. Halton
e-mail: mark.halton@ausnetservices.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_57

Wood poles are one of the key structures used for electricity distribution network. A study of wood poles in Norway was carried out by Gustavsen and Rolfseng [3]. The study examined probabilistic approach for economic impacts of alternative maintenance strategies.

A non-destructive test is available called Resistograph [2] and [4]. It captures the resistance of the contact force of a drill bit as it travels through the wood pole. A plot is produced showing depths and density change. This allows the users to identify spots where there are decay or holes. The data collected can then be used in computer models for determining remaining strength of the pole. However, early decay cannot be identified by this technique.

The studies by Datla and Pandey [1] had examined inspection data of a random sample of 100,000 poles (selected from 2 million). The inspections carried out were mainly visual and hammer test. Resistograph test and wave propagation is also used to examining internal rotting and remaining shell thickness. The data used for fitting lifetime probability distribution are pass-fail data—the population of substandard poles in a given sample. The data used in Datla and Pandey's [1] study do not provide insight into how deterioration occurs over time.

Ultrasonic inspection of wood poles decay study had been carried out by Tallavo et al. [8]. The study presented methodology using advances ultrasonic wave propagation in cylindrical orthotropic medium, signal process, numerical simulations, statistical analysis, and ultrasonic transducer characterization.

A risk management framework model for age-dependent fragility curves for utilities wood poles was proposed by Shafieezadeh et al. [7]. The risk assessment approach looked at the risk profile of an aging wood pole failure under strong wind and hurricanes.

2 Inspection, Data Collection and Assessment Process

The data used in this paper are collected from a drill test during a schedule 5 yearly inspection. The test is used to identify sound wood (good quality wood) left in a wood pole. A condition-based maintenance decision—replacement or reinforcement is made when the wood poles sound wood measurement is less than 30 mm.

The high-level work process is as follows:

1. An inspection order is generated in the asset management system for a drill test every 5 years.
2. The inspector carry out the inspection—digging and exposing the wood under the ground and perform a drill test at ground level. At the same time a drill test at 1 m from ground is also carried out.
3. The inspector record the measurements in the mobility device.

4. The inspector will then recommend if corrective actions are required.
5. If further actions are required (such as reinforcement or replacement), the pole data will be sent to the assessor to determine if the pole is suitable for reinforcement or replacement is needed. The assessor will also determine the work priority.

If the pole ground soundwood measurement is less than 30 mm, corrective action is required. There are 2 options: reinforcement and replacement. If the pole's 1 m drill test measurement is at least 80 mm and meeting other criteria such as no other defects on the pole, no terminates, the pole can be considered for reinforcement via a process called staking. Staking is a process where metal plates (Fig. 1—cycled in red) are installed to strengthen the wood pole and extend its life.

This generic measure of 30 mm is based on engineering judgement, experience and structural engineering calculation considering of the environmental factors (wind, soil), loading, design types and wood species. Another consideration factored in is the time to replacement/reinforcement work.

Fig. 1 A stake pole

To gain more insights of wood poles deterioration mechanism, work is undertaken to examine the historical sound wood measurement data. This paper is a study of over 1.2 million records of data from about 200 thousands wood poles over a 15 years period. This paper aims to provide a view of wood poles P-F interval using sound wood deterioration data.

3 PF Curve

Potential-to-Failure (P-F) curve is a theoretical view showing how failure starts (P) and deteriorates to the point of functional failure (F) [6]. The P-F Interval is "the interval between the occurrence of a potential failure and its decay into a functional failure. For a wood electricity power pole, one measurement that determines the condition is the sound wood measurement, measured at the ground level. Figure 2 is an illustration of the theoretical P-F curve and P-F interval.

3.1 Data Points and Calculations

Measurement records are sourced from the asset management systems. To calculate the deterioration rate, measurement records for two consecutive inspections for the same pole are needed. The rate is simple the difference between the two measurement divided by the time between inspection (illustrated in Fig. 3) as given in Eq. 1.

$$Rate\ (mm\ per\ year) = \frac{m_{last-1} - m_{last}}{t_{last-1} - t_{last}} \tag{1}$$

where m is the measurement in 'mm' and t is in years.

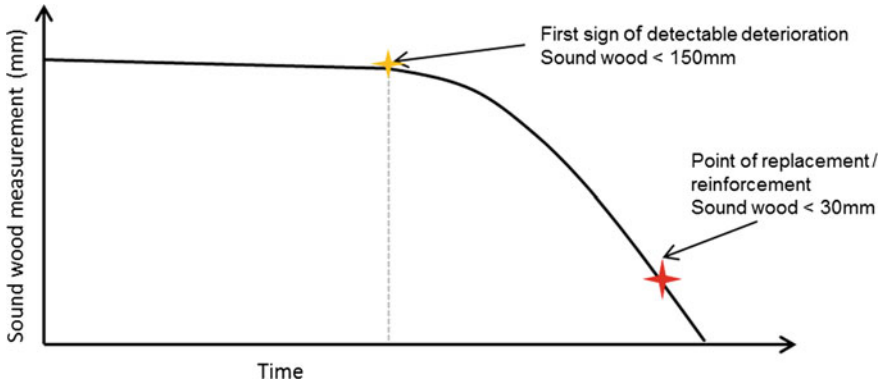


Fig. 2 The PF interval for a wood pole

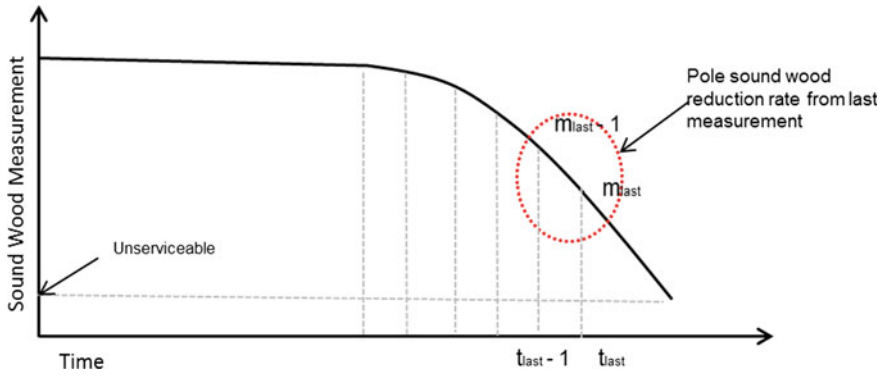


Fig. 3 Soundwood reduction rate

There is approximately 670 thousands data points (as rate mm/year, from 1.2 million data record) used in this statistical analysis. This data points represent approximately 200 thousands wood poles over a 15 years (3 soundwood drill inspections over that period of time).

3.2 Results

This section we present the results of the data analysis. Figure 4 shown the distribution of rate, where majority of the data points has a rate of 0 (i.e. no deterioration between 2 inspections). And when there is deterioration, most deteriorating about 2–4 mm per year.

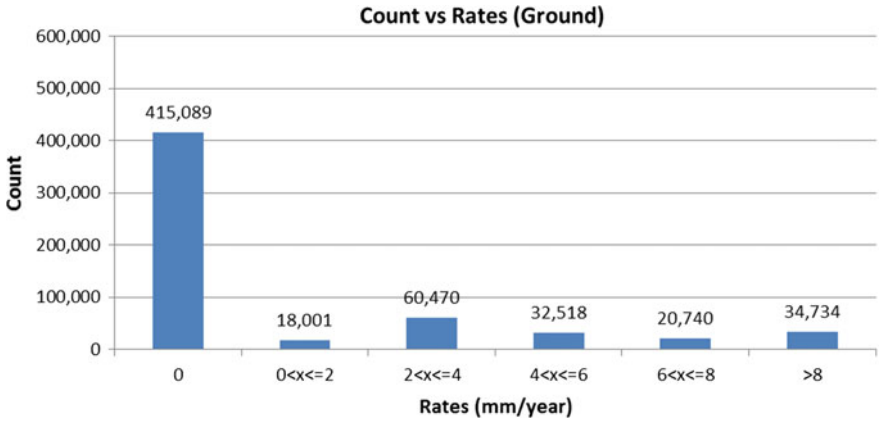


Fig. 4 Data volume in each deterioration rate category

Figure 5 compares soundwood reading at inspection over with pole age and deterioration rate. Figures 6 and 7 provides a detailed view of White Stringybark (WS) and Messmate (MS) poles, which accounts for approximately 40% of wood pole population. It is observed that the PF-Interval (regardless of wood species) is approximately 13 years. For WS poles it is 8 years and for MS poles it is also 8 years. This data is based on average calculation only and has not taken geographic location, loading and design of the poles structure into consideration.

Figure 8 presented the rate against soundwood reading at inspection. It is shown that the highest deterioration rate is close to 5 mm/year when the pole soundwood at measurement is 50 mm.

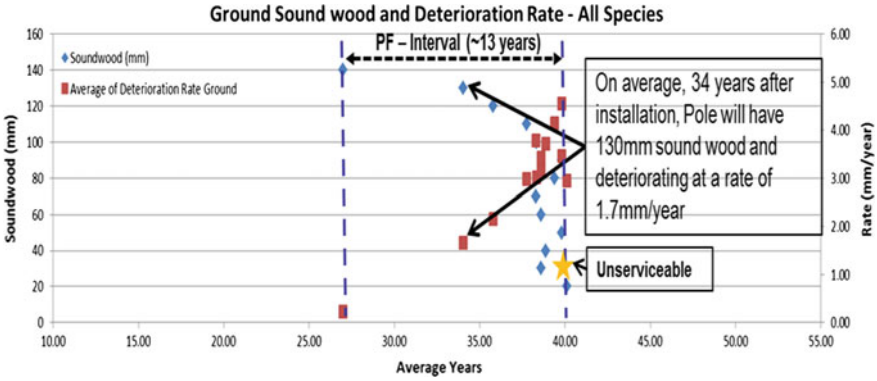


Fig. 5 Deterioration rate over pole age in years

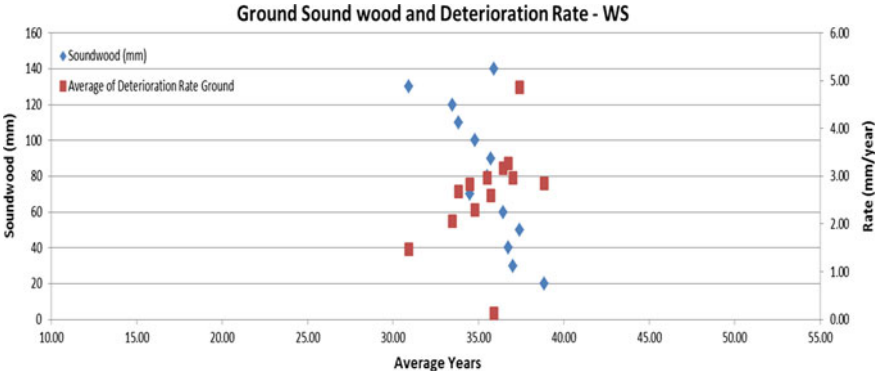


Fig. 6 Deterioration rate over pole age in year—white stringybark (WS)

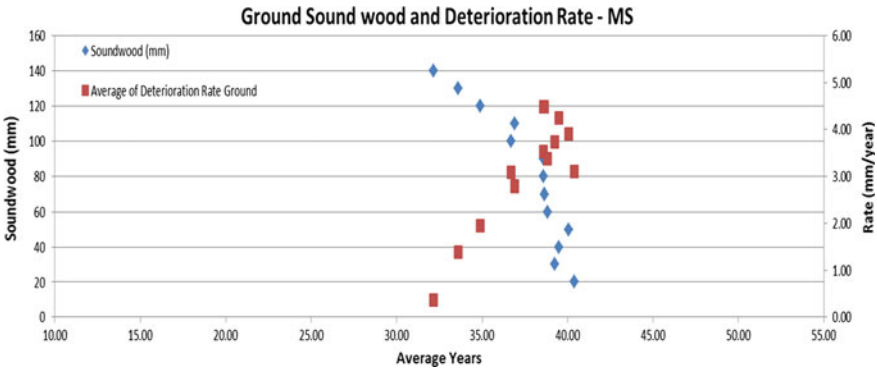


Fig. 7 Deterioration rate over pole age in year—messmate (MS)

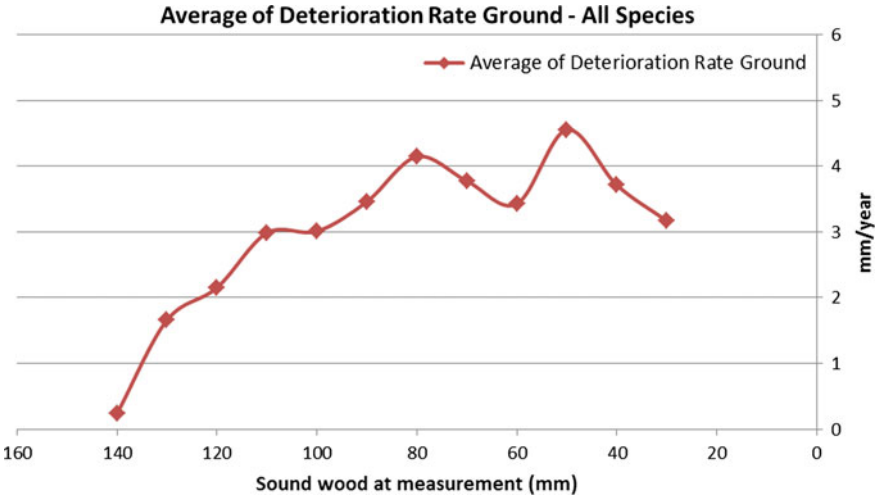


Fig. 8 Average deterioration rate versus soundwood at measurement

4 Conclusion

This paper provides an insight into the PF-interval of wood poles. This information can be used in maintenance planning and scheduling, forecasting replacement volume and also managing risk of pole failure. There is more work required to understand the deterioration mechanism of soundwood losses and also the geographical contribution to wood poles deterioration. It will also be of interest to compare soundwood deterioration rates of each different species and extending this study to all other wood poles in Australia.

References

1. Datla S, Pandey M (2006) Estimation of life expectancy of wood poles in electrical distribution networks. Struct Saf 28:304–319
2. Freeman MH, Ragon WK (2010) A review of wood pole testing equipment compared to visual and excavation techniques used in test and treat programs. In: 2010 southeastern utility pole conference—technical forum, pp 178–188
3. Gustavsen B, Rolfseng L (2005) Asset management of wood pole utility structures. Electr Power Energy Syst 27:614–646
4. IML (2017) Safety with system: IML pole inspection [Online]. Available at http://www.imlusa.com/html/pole_testing.html. Accessed 5 May 2017
5. Jardine AK, Tsang AH (2013) Maintenance, replacement, and reliability, 2nd edn. CRC Press, Boca Raton
6. Moubray J (2001) Reliability-centred maintenance 2, 2nd edn. Butterworth-Heinemann, Oxford
7. Shafieezadeh A, Onewuchi UP, Begovic MM, DesRoches R (2014) Age-dependent fragility models of utility wood poles in power distribution networks against extreme wind hazards. IEEE Trans Power Delivery 29(1):131–139
8. Tallavo F, Cascante G, Pandey MD (2012) A novel methodology for condition assessment of wood poles using ultrasonic testing. NDT&E Int 52:149–156

Seawalls for Coastal Protection and Climate Change Adaptation: A Case Study from the Gold Coast

Rodger Tomlinson and Leslie Angus Jackson

Abstract Best practice for shoreline protection works involves the combination of seawalls with beach nourishment and associated stabilisation works. Seawalls have been constructed at various beaches along the Gold Coast shoreline since the 1920s in an attempt to prevent damage to public and private property. The current design of the Gold Coast seawall dates back to the late 1960s and has proven to be effective against major storm events. In recent years the design of the wall has been reviewed and status of the integrity of the mainly buried wall investigated. In this paper an overview of the history, design and future capacity of the wall to address climate change is presented.

1 Introduction

Beach erosion is a part of a natural cycle for sandy beach environments and generally occurs during periods of storm activity or large swell and can be more severe during large tides and storm surge events. Beach erosion will generally be increased by sea level rise and can also be caused through inappropriate activity that disturbs dune vegetation or through the construction of structures that interrupt natural patterns of sand movement. Gold Coast beaches have endured several seasons of wide-spread significant erosion over the past 90 years in which public and private property and infrastructure have been damaged.

Best practice for shoreline protection works involves the combination of seawalls with beach nourishment and associated stabilisation works: a strategy which has been adopted on the Gold Coast since the 1970s. Basco [1] undertook a design and economic analysis showing the advantages of dune construction and nourish-

R. Tomlinson (✉)
Griffith Centre for Coastal Management, Griffith University, Southport, Australia
e-mail: r.tomlinson@griffith.edu.au

L. A. Jackson
International Coastal Management Pty Ltd., Main Beach, Australia
e-mail: a.jackson@coastalmanagement.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_58

ment including a buried seawall as compared to a seawall only. More recently, Irish et al. [8] and Smallegan et al. [10] have demonstrated the role of buried seawalls in mitigating the impact of Hurricane Sandy, and called for the use of sustainable multi-level protection against natural hazards to be more wide-spread.

Seawalls have been constructed at various beaches along the City's coastline since the 1920s in an attempt to prevent damage to public and private property. The current standard for construction of Gold Coast Seawalls was established in response to major erosion in the 1960s. The design, comprising primary and secondary armour layers backed by a filter layer, reflects standards of the time which have changed little in recent decades [3, 14]. The level of compliance with the seawall standard varies for each seawall depending on when it was constructed and whether it was constructed as emergency protection, or whether it was constructed through a regulatory framework that included certification through the Gold Coast Planning Scheme [4]. There is an alignment along the coast where seawalls are required to be constructed called the A line. The A line was surveyed following erosion events in 1967, 1972 and 1974 and was located as far landward as practical to protect private beachfront homes and public infrastructure.

The present Gold Coast Planning Scheme requirements with respect to coastal management have their genesis in the 1970s and 80s. In the 1980s, climate change and sea level rise were recognised as an issue that was and would continue to affect the Gold Coast beaches. The plan requires seawalls to be constructed by developers as a condition of development of structures upon erosion prone beachfront land. The Planning Scheme also requires that coastal structures include foundations that can resist coastal scour and allowance for wave attack for larger storm events that have breached the seawall. The Seawall standard was established by the Queensland Government's Coordinator General Department in consultation with Dutch experts. In order for the seawall to provide protection against erosion, a continuous wall is desirable along urban sections of the coast. In this paper, the design of the seawall will be discussed, and examined in the context of the seawall being an effective climate change adaptation option.

2 Brief History of the Gold Coast Seawall

The Gold Coast has a well-documented history of storm events resulting in the beaches being prone to erosion and accretion cycles. The occurrence of extreme storms is also cyclical with periods from the 1860–1890s, 1930s and 1950–1970s being particularly stormy. In recent times, the worst recorded sequence of storm events occurred in 1967. However, erosion problems have been recorded on the Gold Coast as early as the turn of the twentieth century in Southport. Development on the beach front was low-key during the first half of the 20th century, but increased from the 1950s with rapid development from the 1970s onward.

As an example of the past use of seawalls, serious storm erosion was recorded in 1920 at the small beachside resort village of Main Beach at Southport—now part of the City of Gold Coast, and reports indicate that a timber log wall was constructed to effectively halt the erosion. This was a typical response using easily obtainable local timber logs that could be easily transported and installed. Such walls were leaky unless suitable gravel was available to place behind them as a filter to retain the sand but allow the elevated water table behind to escape. They also are very reflective and, if not driven deep enough, fail by toe scour. However, they were often a reasonably successful short term measure. Remnants of early timber pile walls are still evident in many areas along the coast.

A few years later, the Pacific Highway linking Brisbane and Sydney was threatened at Narrowneck near Surfers Paradise. Timber groynes proved ineffectual and eventually a very substantial bitumen coated timber log wall was constructed, with a rock toe and gravel behind. This was successful and was reinforced with a boulder wall along the front many decades later when the wall was extended. Up to the late 1990s this wall was still protecting the road, now not a highway but a major local road.

In 1967, seven back-to-back cyclones caused extensive damage to all of the southern Queensland beaches. The effectiveness of earlier seawalls was investigated and the Queensland state government undertook to develop a standard design. A key part of the negotiation over beach protection legislation at the time included provision to ensure certainty that the Gold Coast City Council would not be liable for any negative outcomes from implementing this seawall design. The Council adopted this design and developed a standard installation for the construction of a seawall for all Gold Coast beaches. Seawalls for coastal protection have evolved from timber log walls to dumped rip rap and sand bags for emergency works to engineered rock seawalls [9].

3 Seawall Design and Performance

3.1 Design

The standard Gold Coast wall (Fig. 1) was designed to withstand a low 5% damage value, for between a 1 in 60 year and 1 in 100 year cyclone event. In 1976, the limit of the protection wall was defined within a line called the “A-line” and designed to be continuous and as smooth in plan and curvature as possible.

The A line was to be as far landwards as possible and not be seawards of the 1967 and 1954 erosion scarps. It was preferable that the rear section of the wall be sited landwards of the property line. The plan alignment was to be, as far as possible, parallel to the long-term beach profile in plan [11]. The only significant design change in over 40 years has been the adoption of an alternate design developed by the Beach Protection Authority (BPA) in the 1980s which replaced

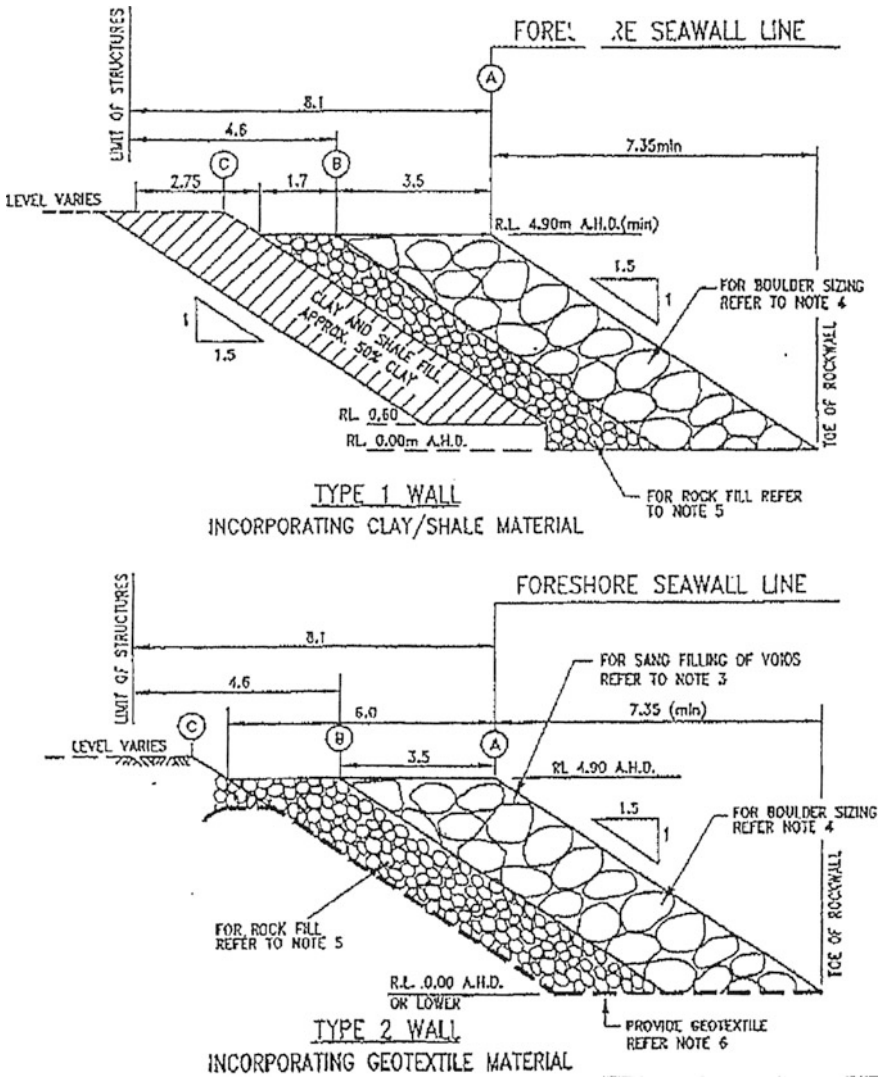


Fig. 1 Seawall designs as incorporated in Policy 7 of Gold Coast Planning Scheme, 2003

the tertiary layer of clay/shale with geotextile. The two adopted designs (Fig. 1) are presently incorporated as a minimum standard requirement in the Gold Coast Planning Scheme. The standard seawalls typically include three layers including an outer face of boulders weighing up to five tonnes each supported by rock fill and a filter layer. There are two types of filter layers, one comprising of clay/shale and an alternative design using geotextile material. Rock sizes can vary between 90 and 360 kg (50% must be over 270 kg). Boulders range between 1.5 and 4 tonne (50% must be over 3 tonne).

All buildings and other structures must be set back 8.1 m (min) from the foreshore seawall A-line to give a 4.6 m (min) flat area clear of all permanent structures. The Gold Coast seawall is designed to have overtopping which leads to the damage of the ground surface behind the wall during an event. The crest level of the seawall varies along the coast depending on the ground levels. Beachfront property owners would not want a seawall that came up higher than the natural ground surface. The minimum height for the top of the Gold Coast seawalls is 4.9 m AHD (15′ before metrification) with some up to 7 m AHD where the sand dunes have been maintained at a higher level.

3.2 Maintenance and Performance

Maintenance of the seawall is performed infrequently and mostly involves the placement of extra sand into rock voids which usually occurs after major rain events. Beach and seawall inspections occur more frequently in the aftermath of a coastal event such as a king tide or storm event, or if construction has occurred in the vicinity of the seawall. If the seawall is not disturbed by events or construction then it is recommended that conditioning reports are obtained every 10–25 years.

Since the original seawall design, there has been considerable change to the coastal landscape of the Gold Coast, including significant increases to beachfront assets and infrastructure and the implementation of extensive nourishment and coastal protection works. These changes have affected the exposure to future climate change of the Gold Coast seawall. As part of the development of a whole-of-coast strategy for coastal protection on the Gold Coast, Queensland, Australia, the structural integrity of the so-called A-Line Seawall has been investigated and a review of the design of the wall has been undertaken [12]. Methods included pot-holing, Ground Penetrating Radar (GPR) and excavation/peel back of the armour layers. The design review has applied present rock seawall design guidelines and standards for the current known extreme conditions and those predicted under climate change. Future wave height and sea level predictions have been applied to assess scour levels and overtopping at the wall.

The seawall investigations helped identify and confirm the condition of the public seawall in priority areas and helped the City in improving the protection of community infrastructure against severe storm erosion, while ensuring that suitable existing walls are not rebuilt unnecessarily.

4 Seawalls as an Adaptation Option

The current status of the seawall was addressed in the Gold Coast Shoreline Management Plan [6] as part of a review of beach management strategies and recommendations to ensure Gold Coast beaches could withstand extreme events

over the next 50 years. The Plan made a high priority recommendation for the completion of the A-Line seawall along the whole of the developed coastline. The seawall was, of course, designed without incorporating current sea level rise projections, and in response to storms in 2009 and 2012 a re-assessment of the design of the wall to accommodate climate change impacts was undertaken [12].

Gold Coast beaches are already highly variable and subject to high energy conditions. While the sandy beach buffer and terminal rock seawall along the A-line presently provides adequate protection in most areas, climate change can be expected to increase the exposure of these seawalls over time and increase the damage experienced during a design event due to increased toe scour and overtopping. Climate change would also result in beach narrowing and the more frequent presence of the seawall in the beachscape. Sandy beaches are essential to the Gold Coast economy and to avoid their loss, it is imperative that the City and State continue to protect and fund the management of the beaches for the long term. Even with active beach management, some increase in the capacity of the rock seawall is expected to be necessary.

As rock seawalls are inherently flexible and easily maintained, they lend themselves to implementation of an adaptation strategy in response to climate change [9]. As such, the upgrade of existing structures can effectively be undertaken over time. Given the uncertainty surrounding climate change, a key component of the adaptation strategy is the inclusion of monitoring of the structure and periodic design review to respond to local changes to seawall vulnerability and actual and projected climate change. No severe erosion events have occurred since the review and thus no monitoring to date.

There are many measures which can be implemented to enhance the capacity of the seawall. Many options are outlined in the Climate Change Adaptation Guidelines in Coastal Management and Planning [5]. These include (but are not limited to):

- Retreat/Allow erosion to occur
- Reduce exposure by increasing upper beach volume (i.e. nourishment and other coastal protection works)
- Accommodate higher level of damage over time and ensure maintenance occurs
- Adapt existing seawall design—larger armour units, higher crest, flexible toe
- Re-construct seawall using best practice design and incorporate elements which allow for future maintenance and design adaptation.

Townsend and Burgess [13] and Headland [7] have addressed the structural integrity issues associated with retrofitting seawalls.

The Gold Coast beaches are one of the City's biggest assets for residents and tourists. While beach nourishment remains the primary long term strategy (with stabilising coastal structures where required), terminal seawalls are still required, to varying extents, to protect infrastructure during severe erosion events and are still a key element of the coastal management strategy for the Gold Coast [2]. As such, increasing the upper beach volume to reduce the exposure of the seawall and

maintain a usable beach width is an essential component of addressing the increased exposure of the Gold Coast seawall associated with climate change. Given the significant lengths of seawall already in place and certified, a strategy of adaptation is generally preferred to reconstruction. The seawall redesign has been based on a 20-year design validity and a 50-year design life. As such, the design considers current 100 year ARI conditions as well as projected impacts from climate change based on average IPCC projections for 2081–2100 (including SLR of 0.5 m). Nearshore design conditions have been determined based on published offshore design conditions, projected impacts of climate variability, assessment of historical beach volumes and cross-shore numerical models. It was recognised that the vulnerability of the seawall was no longer consistent along its length and zones of vulnerability were established. H_s is expected, on current data, to increase from 2.1 m (2014) to 2.7 m (2085) and scour to increase from −1 m AHD to up to −2.7 m AHD in the most exposed areas. These are significant changes but boulder seawalls are flexible structures and are easily maintained during or after an event by placement of additional rock volume on the crest. Its nature lends itself to design adaptation in the form of inclusion of a flexible toe and increased crest levels (Fig. 2).

Design adaptation includes: the use larger primary armour to enhance the stability of armour units; an increase in notional permeability of seawall by modifying secondary armour and filter layers to enhance the stability of armour units by increasing wave dissipation and reducing reflection; incorporating a higher or wider crest to reduce overtopping damage, and founding the toe to anticipated scour depth or incorporate a flexible toe detail to accommodate additional scour.

The capacity of the Gold Coast seawall is such that it is presently able to withstand expected conditions with only minimal (0–5%) damage in most locations. During eroded conditions on the most vulnerable beaches, the present seawall

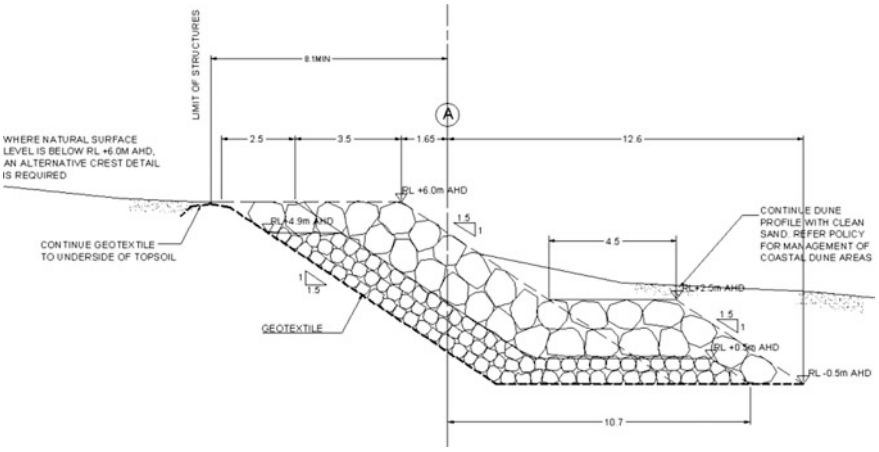


Fig. 2 Design adaptation of Gold Coast Seawall

would presently be expected to accommodate intermediate damage, but not complete failure, during a design event. As such, implementation of adaptation strategies can occur over a reasonable timeframe and the structures can accommodate slight increases in exposure in the interim. The new design is still under review and has not yet been formally adopted.

5 Conclusions

Gold Coast beaches have been exposed to extreme erosion events over the period of recorded history. The over-arching strategy adopted since the 1960s is to maintain a sandy beach buffer and a buried terminal seawall. The existing provides adequate protection in most areas, and status investigations have helped identify and confirm the condition of the public seawall in priority areas and have helped the City in improving the protection of community infrastructure against severe storm erosion, while ensuring that suitable existing walls are not rebuilt unnecessarily. A review of the wall design found that the standard wall designs conformed to present structural performance standards.

However, climate change can be expected to increase the exposure of these seawalls over time and increase the damage experienced during a design event. Even with active beach management and nourishment, some increase in the capacity of the rock seawall is expected to be necessary. As rock seawalls are inherently flexible and easily maintained, they lend themselves to implementation of an adaptation strategy in response to climate change.

References

1. Basco D (1998) The economic analysis of “soft” versus “hard” solutions for shore protection: an example. In: Proceedings 26th international conference on coastal engineering, pp 1449–1460
2. Bowra K, Hunt S, McGrath J, Pistol D (2011) Last line of defence—seawalls. In: Proceedings of the Queensland coastal conference 2011
3. CIRIA (Construction Industry Research and Information Association) (2007) The rock manual. The use of rock in hydraulic engineering, CIRIA, London, UK, Report C683
4. City of Gold Coast (2003) Gold Coast planning scheme policy 7: foreshore rock wall—design and construction, City of Gold Coast, Queensland
5. Engineers Australia (2012) Climate change adaptation guidelines in coastal management and planning. Prepared by the NCCOE of Engineers Australia, EA Books, Crows Nest, NSW
6. Griffith Centre for Coastal Management (2010) Gold Coast shoreline management plan. Griffith Centre for Coastal Management Report No. 90
7. Headland J (2011) Coastal structures and sea level rise: an optimized adaptive management approach. In: 6th international conference on coastal structures, Yokohama, Japan
8. Irish JL, Lynett PJ, Weiss R, Smallegan SM, Cheng W (2013) Buried relic seawall mitigates Hurricane Sandy’s impacts. Coast Eng 80:79–82

9. Mulcahy M, Jackson A, McGrath J, Hunt S, Corbett B, Tomlinson RB, Todd D (2014) Design review for the Gold Coast Seawall for climate change, In: Practical response to climate change conference. Engineers Australia, Melbourne
10. Smallegan SM, Irish JL, Van Donergen AR, Den Bieman JP (2016) Morphological response of a barrier island with a buried seawall during Hurricane Sandy. Coast Eng 110:102–110
11. Smith AW (1990) The evolution of erosion protection on the Gold Coast. Gold Coast City Council Beach Replenishment Program Report No. 137. 13p
12. Tomlinson RB, Jackson LA, Bowra K (2016) Gold Coast Seawall: status update and design review. J Coastal Res SI 75:715–719
13. Townsend, IH, Burgess KA (2004) Methodology for assessing the impact of climate change/ upon coastal defence structures. In: Proceedings 29th international conference on coastal engineering, pp 3953–3965
14. USACE (US Army Corps of Engineers) (2003) Coastal engineering manual, pp. A–32

Simulating the Interrelationships of Carbon Taxation, Electric Power Costs, and Solar PV Installation

Amy J. C. Trappey and Charles V. Trappey

Abstract Photovoltaic (PV) systems in Taiwan are affordable because of the country's export oriented semi-conductor industry. Effectively promoting renewable energy to the public requires convincing policies and regulations, e.g., carbon taxation. If the marginal cost of reducing carbon dioxide emissions is lower than the carbon tax rate, the emission can be reduced effectively with increasing use of clean PV electricity. This research develops System Dynamics (SD) simulation models to evaluate the interrelationships of carbon taxation on PV system installations and the prices of fossil and green power supplies. SD models are built with scenarios to assess the results of different tax rates, emission reductions, electricity costs, and PV capacities. Data are collected from Taiwan Bureau of Energy (The statistics of carbon dioxide emissions from fuel combustion in Taiwan. Ministry of Economic Affairs, Taiwan, 2013, [17]) and international power and energy sectors to serve as a reference for policy makers to set green energy regulations.

Keywords System dynamics · Photovoltaic system · Renewable and sustainable energy · Carbon tax

1 Introduction

The Taiwan government approved a series of policies to change the structure of energy consumption to lower carbon emissions and promote renewable energy industries. The Renewable Energy Development Bill promotes the installation of renewable energies using a carbon tax and carbon trade as regulatory strategies to

A. J. C. Trappey (✉)
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan
e-mail: trappey@ie.nthu.edu.tw

C. V. Trappey
Department of Management Science, National Chaio Tung University, Hsinchu, Taiwan
e-mail: trappey@faculty.nctu.edu.tw

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_59

control carbon emission and promote green energy consumption. In general, European countries advocate a carbon tax while the U.S. supports carbon trade [14]. Carbon trade allows for companies to buy and sell carbon emissions and may become detached from public interests. A carbon tax tends yields government income which can be re-invested publicly supported green causes.

In 2009, the Taiwan Executive Yuan approved the Green Energy Industries Sunrise Program identifying seven green industries that demonstrate the viability of PV energy and LED illumination. There are over 160 companies in the Taiwan PV industry supply chain that complements the globally competitive semiconductor chip fabrication and electronic high-tech sectors [3]. In early 2010, the government drafted an energy bill to levy tax on high CO_2 emissions from fossil fuel emissions. The energy tax bill failed legislative enactment due to concerns of a highly negative economic impact. The effects of energy or carbon taxation require further analysis before new legislation can be drafted.

This research, considering the related issues and factors, uses a System Dynamics (SD) approach to evaluate variations and interrelationships in fossil fuel electricity costs, electricity prices, and carbon taxes. The government income from carbon tax is allocated toward the renewable energy sector (PV systems) under different policy scenarios. The SD models established by this research evaluate the impact of a carbon tax on different economic and environmental aspects. Second, this research provides a decision support tool with quantitative measures and a time line to assist government legislation to support carbon tax rates that effectively promote the PV industry.

2 Literature Review

The Taiwan government approved the Sustainable Energy Policy Program in 2008 which promoted renewable energy policies and set targeted reductions in fossil fuel consumption and carbon emissions. The Executive Yuan approved the program as a sustainable strategy from the perspective of the country's energy, environment, and economic needs [16]. The Renewable Energy Development Act was approved in 2009. Tsou [19] wrote that the act included subsidies for specific renewable industries, established a convention for annual feed-in-tariffs, and secured funds for the development of renewable energy. The National Carbon Reduction Project set project goals including laws, regulations, low-carbon emission energy systems, low-carbon community and society, low-carbon industrial structure, green transportation systems, green landscape and buildings, improved low-carbon technologies, low-carbon public works, energy-saving education, and energy-saving promotion.

The US, Japan, China, UK, Germany and other countries have invested over USD185 billion for clean energy initiatives. Taiwan's R&D and manufacture ability is an advantage to create original design and manufacturing for green products. The Green Energy Industries Sunrise Program in 2009 forecasted that promoting these

industries would significantly contribute and sustain industrial growth. The program places the PV energy industry as the bellwether to execute strategies to improve key techniques and materials, establish the supply chain, build labs, set satisfy international standards, and market new products to satisfy growing domestic and international demand [11].

2.1 Carbon Tax and Carbon Trade

There are typically three policies used for reducing greenhouse gases including executive control, market tools (e.g. carbon tax, carbon trade), and volunteer agreements. Government carbon tax is based on the quantity of carbon which specific products contain and encourage users to save energy or develop clean energy sources. When the cost of reducing the source of greenhouse gas is lower than the tax rate, these harmful emissions are reduced. Carbon trade administration often sets a limit (or cap) and allows companies to auction or allocate the carbon dioxide emission credit to emission sources. Energy taxes are often levied on sources of air pollution and fuels.

The Taiwan energy tax draft is not a standard carbon tax because the rate is not calculated based on the quantity of CO_2 emissions and the tax income is redirected back to the development of a green energy infrastructure [6]. Energy taxes should provide incentives with flexibility. Research by Wen and Li [20] indicates that energy taxes are difficult to control and to verify the achievement of environmental goals.

Carbon tax and energy taxation has been implemented successfully by some countries. This section discusses these policies as references for Taiwan. Wu et al. [21] reported that Australian carbon emission reduction laws increases the export prices of coal, nature gas and other energies and mining sales loss is often transferred to the consumer. Coal from Australia accounts for 45.4% or almost half of Taiwan's total. This research uses the DSGE model to analyse how rising import energy prices negatively affect Taiwan's economy. Rising prices cause inflation, and decrease GDP, but reduce carbon dioxide emissions. Liang's research [7] is based on the Taiwan DGEMT model and evaluates scenarios the influence of energy tax on the price of goods, economic growth, and greenhouse gas emission. Liang [5] evaluates the influence of the Kyoto Protocol on the Taiwan economy. The DGEMT model is used for simulation and scenario analysis. The results indicate gradually increasing energy prices more effectively reduces negative economic outcomes and reduces carbon emissions. The electricity and fuel price should reflect real costs and the government should implement a carbon tax. Nakata and Lamont [12] studied the influence of energy and carbon tax on Japan's power supply system and forecast the changes needed for the energy infrastructure structure before 2040. Their research shows that carbon tax decreases coal usage but increases the use of nature gas. Massetti's research [10] notes that rapid economic growth in China and India has greatly increased global energy demand and

both countries play important roles for reducing global carbon emission. This research simulates carbon tax scenarios to observe the impact on Taiwan's energy policy goals and the means to reach these goals without negatively impacting other countries.

2.2 Photovoltaic (PV) Energy

Global warming has led many countries to reduce set energy demand, reduce carbon emissions, improve power efficiency, and invest in renewable energy sources. Nuclear power is an option, because of low costs and low carbon emissions. However, after the Japan earthquake in 2011, a review of nuclear power safety and nuclear waste disposal has become a growing concern. Therefore, PV and other renewable energies may be a solution that need further improvement through research and development.

Photovoltaic energy uses radiation from the sun to create electricity and the manufacturing of solar cells is similar to the manufacturing of semi-conductors for computer chips. Manufacturing combines an N type semiconductor and a P type semiconductor with a barrier placed between these two type materials. Sunlight passing through the semiconductor materials and across the barrier generates electric generate power. A PV power generator system contains a solar cell, a solar panel, a power invertor, a power regulator, and a stand. PV systems are placed on roofs, the ground, and are integrated into buildings. There are three critical materials needed to manufacture a solar cell. The first one is a-silicon. The manufacturing cost of this type is relatively high and there are some restrictions with shape. The second is poly-silicon. The cost is lower than the first one, but the efficiency is worse than a-silicon cell. The last is amorphous silicon (thin films). This type is cheap, but the efficiency is poor. However, it can be used in many fields and integrated into buildings because it is flexible. The poly-silicon solar cell holds the largest portion of the market share [8].

The PV systems are defined by stand-alone, hybrid, or grid-connected. Most PV systems are grid-connected with stand-alone systems holding 2% of the market share. However, developing countries tend to install stand-alone and small scale PV systems and developed countries (Australia, Switzerland, and Norway) also developing stand-alone PV systems reported by REN21 [13].

The global installed PV system reached 13.8 GW in 2010, and Germany is the largest market. Overall demand has decreased since 2011 because incentive programs (FIT) are decreasing in Europe. Researchers forecast that the United States and Asia have the greatest market potential. China has become the world's largest solar cell manufacturer but its domestic market is weak. Moreover, although the demand of PV system has decreased, Taiwan and China companies are still aggressively expanding the production of solar cells through subsidized programs. The supply of silicon material is unstable which has increased vertical integration in the supply chain.

The Taiwan government promotes solar energy with the PV accumulated installed capacity reaching 41.1 MW in 2011. The proposed industry policies target the development of low cost, high efficiency silicon PV modules, high efficiency thin film PV system, and integrated grid solar cell systems. The total output value for Taiwan's green industries is NTD 41.49 billion with PV and LED illumination holding the largest shares. The PV supply chain focuses on low cost with the output of solar cells from China and Taiwan holding 76% of the global market.

Carrión et al. [2] used environmental decision support systems to locate solar farms. The models used multiple criteria analysis and geographic resource systems, to evaluate environment, terrain, weather and other criteria with weather being the most important determinant of how much power can be generated. Zhang et al. [24] collected Japanese government data from 1996 to 2006 to analyse different criteria that influence the diffusion of PV. Policies (e.g., subsidy programs), domestic income, and environmental awareness have a positive influence whereas high installation costs limit growth. Liu and Hsu [9] discuss the success factors of the PV industry and recommend that good policies, people's awareness, professional installation of PV system, and building a complete supply chain. The authors recommend that Taiwan's government change the energy structure, encourage self-sufficiency energy options, increase domestic demand for innovation, and promote well defined clear green energy targets. Yuan et al. [23] studied global investments in improving energy efficiency, renewable energy, and intelligent grids. Trappey et al. [18] analysed the future trend of PV and wind power system installation costs using a hierarchical learning curve model. The results show that installation costs, total installed capacity, and silicon prices are significant factors driving installation. The installation costs decrease almost 88% of the original costs when the total installed capacity doubles.

3 SD Methodology Highlights

System Dynamics (SD) models the causal relationships and feedback loops in social science, management science, engineering, organizational structures, and industrial economics [22]. SD uses numerical data to build the structures of systems and to simulate and analyse scenarios and their outcomes [4].

The SD process is divided into problem concept analysis and model simulations analysis to define the components, cause, and relationships that define systems. SD modelers define the problem, describe the system, and plot the causal feedback loops. Plans made from the previous steps become models used to simulate different scenarios. After building and simulating SD models, reliability tests (including expert review) are used to determine if the model represents the real conditions and feasibly models the system.

According to the data provided by TaiPower, renewable energy accounts for a small part of power generation while coal and natural gas dominate. If the government levied a carbon tax, it will influence electricity costs and forecasts that electricity demand will increase 3% per year to satisfy economic development. For

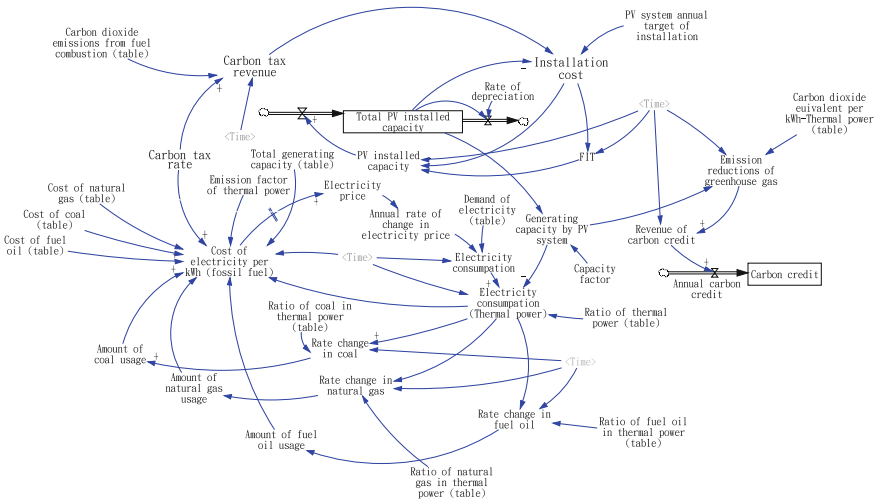


Fig. 1 The SD diagram for evaluating the influence of carbon tax on the PV system installation and electricity prices

the above reasons, we foresee that the demand for fossil fuels will change and the electricity costs will rise with a carbon tax. The TaiPower report shows that the cost of fuel has increased 215% from 1992 to 2012, but the electricity price has only increased about 25% which indicates an imbalance in the system. Considering previous research, causal relationships are defined in Fig. 1. Increasing PV system installation costs will negatively impact the diffusion of PV systems. The installation cost will decrease when the total PV system installed capacity increases, i.e., a positive loop in the system. Taiwan's PV installed capacity increased sharply when the Taiwan government used feed-in tariffs to reduce the risk of renewable energy investment. The PV system replaces electricity generated by fossil fuels and reduces greenhouse gas emissions as shown in the upper-right of Fig. 1. The combined SD model observes factors and causal relationships as a complex system. Different scenarios assess installed capacity and the amount of greenhouse gas emission reductions under varying conditions. This research use CERs to measure these effects. Figure 1 shows the flow diagram for evaluating the carbon tax influences on the PV system and electricity price.

4 The Scenarios and Simulations

This research uses system dynamic software Vensim to construct the quantitative model. The data for parameters and variable such as the installation cost of PV systems, the cost of power generation (per KWh), electricity demand, and carbon dioxide emissions from fuel combustion are collected from TaiPower and government data.

4.1 Verification Reliability and Validity

Forrester and Senge [4] and Balas [1] provide a unique discussion of the verification approaches for the system dynamic model. This research uses the extreme conditions test and the behaviour reproduction test to confirm the reliability and validity of the SD model. The extreme conditions test checks if the system is robust when extreme values are input. In this research, we set the PV installation cost at $0 (extreme condition) and the FIT value is still positive and the annual PV installed capacity doesn't increase without limit indicating normal function. If we set the FIT value to $0, the annual PV installed capacity still increases annually, and the growth rate decreases as expected. When setting the fossil fuel cost per kWh to $0, the electricity price is still positive, and total electricity consumption doesn't become extremely large. Moreover, when the electricity price is extremely high, the electricity consumption is stable.

Four indexes are used for the behaviour reproduction verification. We compare the historical data and the data from simulation to view the degree of consistency. The indexes used are the installation cost, installed capacity, fossil fuel cost per kWh, and electricity price. The results show that the quantitative models, created in this research, match the trends and patterns of the Taiwan market situation.

4.2 The Description of Scenarios

This research uses different carbon tax policies for the different SD scenarios. We refer to the Australia case and the Taiwan Green Energy Industries Sunrise Program to approximate the revenue of carbon tax used to subsidize the PV system. Domestic and international cases show that if government budget surpluses can be used for renewable energy subsidies, then renewable energy sources are more readily distributed across the energy system. For instance, a one hundred thousand NT$ PV roof policy implemented by Kaohsiung City, Taiwan was forecast to increase 10 MW of PV installed capacity per year. Denmark provides many incentives for wind power energy and has become one of the most important wind power generator exporting countries.

Scenario 1: Business as usual (BAU). This scenario is used to simulate without implementing a carbon tax policy and provides a baseline to compare with other scenarios.

Scenario 2: Sweden carbon tax policy. The emission of carbon dioxide has decreased 9% in Sweden from 1990 to 2006 under conditions of stable economic growth. Sweden maintains environmental protection using a carbon tax from 1991. The original tax rate was USD100 per ton of carbon dioxide and increased to USD150 in 2011. In this scenario we assume Taiwan will levy carbon tax from USD100 per ton of carbon dioxide in 2015 and the annual growth is USD2.5.

Scenario 3: Sweden carbon tax policy (adjust to GDP). According to the report form International Monetary Fund (IMF), there is a gap between the Gross Domestic Product (GDP) in Taiwan and Sweden (Taiwan's per capita GDP is USD20,336 while Sweden is USD54,815). Hence, this research decreases the impact on national and social finance by assuming Taiwan will levy a carbon tax from USD37 per ton of carbon dioxide in 2015 with an annual growth is USD2.5.

Scenario 4: Sweden carbon tax policy (adjust to GDP) with high subsidy policy. For this case we allocate 10% of the carbon tax revenue to renewable energy which models the Australia case. However, Australia uses a large proportion of the tax to subsidize industries and lower production costs. Australia numerous pure aluminium ingot aluminium smelters, mines, and pulp mills that consume energy. If the Taiwan government used tax revenue to subsidize particular industries, there would be political challenges. Hence, this scenario assumes an allocation of 30% of the carbon tax revenue to renewable energy.

Scenario 5: Denmark carbon tax policy with high subsidy policy. Demark government established the energy consumption tax from the 1970s. The country was a leader in levying a carbon tax on families and enterprises to stimulate energy saving and lower carbon emissions in 1992. In 2005, their petroleum consumption decreased by nearly 10 million tons. The original carbon tax rate was DKK 100 (about USD18) and decreased to DKK 90 in 2005 [15]. We assume Taiwan will levy carbon tax from USD18 per ton of carbon dioxide in 2015 with an annual reduction of USD0.20. The PV installed capacity fits the Taiwan government goals and assumes an allocation of 30% of carbon tax revenue to renewable energy.

4.3 Comparison and Interpretation of the Simulation Results

The predicted electricity price trends are quite different between scenarios. The electricity price rises sharply in scenario 2 because of the high carbon tax rate. The carbon tax rate is adjusted by GDP in scenario 3 and 4, which moderates the social and financial impact. However, according to official statistics, the electrical and electronic industry and chemical industry consume the most power in Taiwan. Many companies in Science Park are original equipment manufacturers (OEMs) and their gross profit is fairly low. Therefore, high electricity price will severely impact profits. Scenario 5, using a low carbon tax rate, causes the least impact.

The trends of PV system installed capacity are also different between scenarios. The impact of carbon tax on installed capacity is weaker than on electricity price because capacity is largely influenced by installation costs and FIT. One of indexes in the FIT formula is the initial installation cost, so if the installation costs are lower,

the FIT is also lowered. The installation cost is a negative factor for PV installed capacity while FIT is a positive factor. The most important way to earn money is resale of the electricity generated by the PV system. The FIT value is a critical factor for successful investment. Scenario 3 and 4 use the same carbon tax policy but differ in the allocation of carbon tax revenue for renewable energy (i.e., Scenario 4 uses high subsidy). Scenario 5 is the result of the lowest carbon tax rate with relative high subsidy. Scenario 2 has the highest carbon tax rate with a lowest subsidy policy. In summary, the high carbon tax rate will not be the best solution for promoting PV systems in Taiwan (Table 1).

Table 1 Comparison and suggestion of the results of simulation scenarios

Scenario	Introduction	Influence on electricity price	Influence on PV installation	Carbon tax revenue	Suggestion
Scenario 1 (BAU)	Business as usual	–	–	–	–
Scenario 2	Sweden carbon tax policy. Carbon tax from USD100 per ton of carbon dioxide in 2015 and annual growth is USD2.5	The highest, 47% higher than BAU	High, 5.3% higher than BAU	The highest	This scenario has a serious impact on electricity price. Hence, it difficult to implement
Scenario 3	Sweden carbon tax policy (adjust to GDP). Carbon tax from USD37.099 per ton of carbon dioxide in 2015 and annual growth is USD2.5	Medium, 17% higher than BAU	Low, 6% higher than BAU	35–40% revenue of scenario 2	The carbon tax impact on the electricity price is acceptable and the government gains substantial revenue. However, it provides limited assistance for PV system installation
Scenario 4	Sweden carbon tax policy (adjust to GDP), and high subsidy policy	Medium, 17% higher than BAU	The highest, 17.1% higher than BAU	35–40% revenue of scenario 2	The carbon tax impact on the electricity price is acceptable and has the greatest increase in PV system installed

(continued)

Table 1 (continued)

Scenario	Introduction	Influence on electricity price	Influence on PV installation	Carbon tax revenue	Suggestion
					capacity among all scenarios. This appears to be the best option for Taiwan
Scenario 5	Denmark carbon tax policy and high subsidy policy. Carbon tax from USD18.2093 per ton of carbon dioxide in 2015 and annual reduction is USD0.2023	Low, 8% higher than BAU	Medium, 8% higher than BAU	<20% revenue of scenario 2	The carbon tax impact on electricity price is minimum among all scenarios and provides limited assistance for installation of PV system

5 Conclusions

This research builds a SD model to evaluate the effects of carbon tax on electricity price and PV system installed capacity. The contributions of this research are to construct SD models as a decision support tool to refine carbon taxation policies. The assessment is based on different scenarios (referring to other countries' tax policies) to forecast the influences of carbon tax on economic and environmental aspects in Taiwan. The carbon tax directly influences the electricity price. High carbon tax rates increase electricity prices drastically and low rates do not stimulate the installation of renewable energy systems. Hence, the model built in this paper is valuable in assessing the tax framework and subsidy mechanism. Scenario 5 is useful for minimizing the economic shock to consumers while still providing revenue. However, for the PV industry, scenario 4 is the preferred option.

Acknowledgements This research is partially supported by Taiwan Ministry of Science and Technology and Industrial Technology Research Institute.

References

1. Balas Y, Carpenter S (1990) Philosophical root of model validity: two paradigms. Syst Dyn Rev 6(2):148–166
2. Carrión JA, Estrella AE, Dols FA, Toro MZ, Rodríguez M, Ridao AR (2008) Environmental decision-support systems for evaluating the carrying capacity of land areas: optimal site selection for grid-connected photovoltaic power plants. Renew Sustain Energy Rev 12 (9):2358–2380
3. Council for Economic Planning and Development (2011) The research strategy of the Taiwan photovoltaic industry in a globally competitive environment
4. Forrester JW, Senge PM (1980) Tests for building confidence in system dynamics models. In: Forrester JW et al (eds) System dynamics, North-Holland, New York
5. Liang CY (2007) Taiwan sustainable development—energy prices policy. Taiwan Econ Forecast Policy 37(2):1–36
6. Liang CY (2008) Integration of energy tax, carbon tax and carbon trading system. Oil Market Weekly
7. Liang CY (2009) Impaction of energy taxes on energy demand and economy in Taiwan. Taiwan Econ Forecast Policy 44(1):45–78
8. Lin YC (2010) Photovoltaic industry and its market trends. Industrial Technology Research Institute
9. Liu MT, Hsu YC (2012) The urgent needs of forward-looking policies to develop renewable energy in Taiwan—the enlightenment from Germany. J Public Adm 43
10. Massetti E (2011) Carbon tax scenarios for China and India exploring politically feasible mitigation goals. Int Environ Agreements Politics Law Econ 11(3):209–227
11. Ministry of Economic Affair (2009) Green energy industries sunrise program
12. Nakata T, Lamont A (2000) Analysis of the impacts of carbon taxes on energy systems in Japan. Energy Policy 29(2):159–166
13. REN21 (2013) 2012 renewables global status report. Accessed 15 Sept 2013 [Online]. Available at http://www.map.ren21.net/GSR/GSR2012_low.pdf
14. Shaw DG, Hung CM, Lo SF (2010) A comparison of carbon tax and carbon trading and an examination of their compatibility. Monthly J Taipowers Eng 747:59–66
15. Sumner J, Bird L, Smith H (2009) Carbon taxes: a review of experience and policy design considerations. National Renewable Energy Laboratory (NREL) Technical Report
16. Taiwan Bureau of Energy (2008) Sustainable energy policy program
17. Taiwan Bureau of Energy (2013) The statistics of carbon dioxide emissions from fuel combustion in Taiwan. Ministry of Economic Affairs, Taiwan
18. Trappey AJC, Trappey CV, Liu PHY, Lin LC, Ou JJR (2013) A hierarchical cost learning model for developing wind energy infrastructures. Int J Prod Econ 146(2):386–391
19. Tsou CC (2010) A study on the legislation process and execution of the renewable energy development act in Taiwan. National Taiwan University
20. Wen LC, Li YC (2008) Energy tax feasibility analysis in Taiwan. Environmental Protection Administration Executive Yuan, R.O.C., Taiwan
21. Wu YH, Hung CY, Chiu CS (2012) Analysis of the Australian carbon tax. Industrial Technology Research Institute
22. Yang CC, Chang LC, Yeh HC, Chen CH, Yeh CH (2007) System dynamics. Wunan, Taiwan
23. Yuan BJC, Li KP, Kang TH, Shieh JH (2012) Foresight of development of Taiwanese solar photovoltaic industry. Int J Foresight Innov Policy 8(1):1–21
24. Zhang Y, Song J, Hamori S (2011) Impact of subsidy policies on diffusion of photovoltaic power generation. Energy Policy 39(4):1958–1964

Design and Development of a Value-Based Decision Making Process for Asset Intensive Organizations

Manuela Trindade, Nuno Almeida, Matthias Finger and Daniel Ferreira

Abstract Asset intensive organizations have been adopting internationally recognized asset management principles towards the realization of value from their asset base. This normally involves the grasping of opportunities for an enhanced balancing of costs, risks and performance throughout the life cycle of asset portfolios or asset systems. Depending on the asset management maturity level of the organization, this may be done in a more or less explicit and formal manner. This paper proposes a formal and systematized value-based opportunity management process for asset intensive organizations. It presents the designing of the process and its foundations on the background knowledge of internationally recognized management approaches in the infrastructure industry and on the publications of the scientific and technical communities. The expected benefits and application of the process in multi-sector asset intensive organizations, namely infrastructure organizations with varying levels of asset management maturity, are also discussed.

M. Trindade
Infraestruturas de Portugal, Almada, Portugal
e-mail: manuela.trindade@tecnico.ulisboa.pt

N. Almeida (✉)
IST – University of Lisbon, CERIS, Lisbon, Portugal
e-mail: nunomarquesalmeida@tecnico.ulisboa.pt

M. Finger
EPFL – École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
e-mail: matthias.finger@epfl.ch

D. Ferreira
BG Ingenieurs Conseils, Lausanne, Switzerland
e-mail: Daniel.FERREIRA@bg-21.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_60

1 Introduction

Various asset intensive organizations have been contributing to the development of the body of knowledge in the field of asset management [4, 8, 10, 12]. Since 2014, the international series of standards ISO 55000 have also been assisting asset owners, asset managers and service providers to align their asset management practices against common principles, terminology and requirements for their asset management systems.

Asset management is multidisciplinary and engages cross-functional processes, people and technologies within the organization. Thus, the components of an asset management system are applicable to the organization as whole and over its many areas and levels (strategic, tactical and operational). These components may namely apply to specific organizational functions, projects and activities, and interrelate with other areas of concern, such as quality management, and especially risk management, amongst others.

It is relevant to note that the interrelation of ISO 55001-based asset management systems with other ISO-based management approaches is facilitated due to the high-level structure that is progressively being adopted by all ISO management standards (Annex SL of ISO/IEC Directives, Part 1). Figure 1 suggests a framework for managing various areas of concern and integrating risk-based asset management systems in parallel with other management systems through a joint process approach.

Some asset intensive organizations have identified the need to design, develop and implement a set of 19 guiding asset management processes in order to run their asset management systems [1]. There is no fixed grouping for the set of guiding asset management processes, as this depends on the type, size and complexity of the organization and the asset portfolio, or the scope and goals of the asset management system, amongst other issues. For example, asset management and governance processes in large asset intensive organizations may include processes to guide capital project prioritization, cost estimates, business case assessments, risk

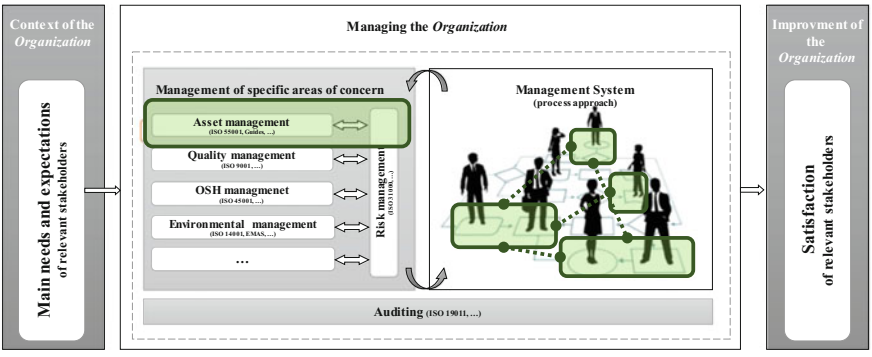


Fig. 1 Process approach of an asset management system within an organization

management, procurement and plan delivery [15]. On the other hand, it is often the case that asset intensive organizations establish, in either a formal or informal manner, short (e.g. 3–5 years) and long-term (e.g. 20 years) asset management plan development processes [11, 23, 24].

ISO 55000 defines asset management as the coordinated activity of an organization to realize value from assets. It thus seems relevant for asset intensive organizations to formally adopt processes that assist and clearly demonstrate the achievement of the goal of asset management, that is, to realize value from assets. This paper intends to provide a contribution in this regard. It presents a methodology and discusses the conceptual background for supporting the design of a value-based opportunity management process for asset intensive organizations, namely infrastructure organizations, which is particularly suitable to be implemented at the strategic governance level.

2 Methodology

Conceptual frameworks can be seen as interlinked conceptual elements that provide a comprehensive understanding of a theorized reality. In this paper a modified version of the methodology adapted by other authors [2, 9, 15] is used to develop a value-based opportunity management process. Firstly, a thorough literature review covering scientific and technical publications and international standards was undertaken. The next step was the systematic and synthetic analysis of the definitions of value, risk and opportunity, as they are used in this paper. This conceptual background sets the foundation for outlining the proposed value-based opportunity management process. The proposed process is to be validated with complementary empirical evidence through its application to a set of real case studies in selected asset intensive organizations.

3 Conceptual Overview

3.1 Value, Opportunity and Risk

The ISO 55000 series of standards lead organizations towards the realization of value from assets. The standard suggests a perception for the concept of value by stating that the goal to be achieved "is the desired balance of cost, risk and performance". But it does not set the ways in which an organization may manage opportunities towards value realization from their asset base.

On the other hand, despite of the publication of various international management standards, such as the ISO 9000, the ISO 55000 and the ISO 31000 series, the concepts of value, opportunity and risk are interchangeably used by different

organizations in the same sector, and sometimes even within different departments of the same organization (see Table 1). In fact, these series of standards establish opposing definitions and uses for the concepts of risk and opportunities. For example, in the latest versions of the quality and asset management standards (ISO 9000 and ISO 55000), opportunity is not the positive side of risk, whereas that has been the case within the context of risk management since the publication of ISO 31000 in 2009. Although the concept of risk is commonly understood as being related to uncertainty, the inconsistency in its definition and use somehow hinders risk-based decision making within asset management [26].

Table 1 Synthetic analysis of the concepts of risk, opportunity and value

Source	Risk	Opportunity	Value
ISO [14]	Risk is commonly understood to have only negative consequences; however, the effects of risk can be either negative or positive	Opportunity is not the positive side of risk. An opportunity is a set of circumstances which makes it possible to do something. Taking or not taking an opportunity then presents different levels of risk	–
ISO 55000	Asset management translates the organization's objectives into asset-related decisions, plans and activities, using a risk based approach Risk—effect of uncertainty on objectives	Necessary to manage, as well as risk, to realize value An asset management system integrated into the broad governance and risk framework of an organisation, can contribute tangible benefits and leverage opportunities	Assets exist to provide value to the organization and its stakeholders
ISO 55002	In planning, an organization should (…) address risks and opportunities for the asset management system and the asset portfolio taking into account how these risks and opportunities can change with time		Value is defined at the organizational level, and not at the asset, asset system or portfolio level. Value must be quantified in relation to the organizational mission in delivering outcomes for its customers and stakeholders

(continued)

Table 1 (continued)

Source	Risk	Opportunity	Value
ISO 31000	Effect of uncertainty on objectives	(…) the management of risk enables an organization to (…) improve the identification of opportunities and threats	Value, can be both financial and non-financial, the same applies to risk, performance and cost. [Risk management creates and protects value /performance is any positive effect on objectives and cost is any negative effect on objectives]
ISO 21500	The risk subject group includes the processes required to identify and manage threats and opportunities [Opportunities selection includes consideration of various factors, such as how benefits can be realized and risks can be managed]	The organizational strategy identifies opportunities (…) selected opportunities are further developed in a business case or other similar document, and can result in one or more projects that provide deliverables. Those deliverables can be used to realize benefits	[Value creation framework]
ISO 21504	Portfolio management processes and systems should be aligned with (…) risk management processes and systems [A method is available to evaluate to what extent the portfolio is in alignment with the tolerated risk exposure]	Opportunities and threats can come from the strategy, customer requests, evolution of offerings, or internal improvements. In some cases, the identification of opportunities and threats may be part of portfolio management. The response to opportunities or threats may lead to one or more new portfolio components or may modify one or more existing portfolio components	Maintain alignment of cumulative portfolio risk with the value created from the successful achievement of strategic objectives Portfolio performance measurements should be used to enable the portfolio manager to ensure that the total investments of all portfolio components are on track to achieve individual and aggregate level benefits and value to strategic objectives

(continued)

Table 1 (continued)

Source	Risk	Opportunity	Value
EN 1325	–	–	Value is the measure which expresses how well an organization, project, or product satisfies stakeholders' needs in relation to the resources consumed Value = Satisfaction of needs/ Consumption of resources
AS 4183	–	–	An attribute of an entity determined by the entity's perceived usefulness, benefits and importance [The terms 'benefits' refers to the advantages gained or enhanced well-being"]
GFMAM [8]	–	The possibility to do something different in the pursuit of realization of value	Satisfaction of stakeholders' needs in relation to the costs incurred and in consideration of the associated risks Value = (Performance − Cost) @ Risk
Parlikad and Srinivasan [19]	–	Any asset management decision should focus on how the decision outcome will affect (positively and/or negatively) the value provided by the asset	Value-based management involves striking the right balance between cost, risks and performance of assets over the life cycle of the infrastructure
Copperleaf [7]	–	Optimizing the value created by the investments an organization could make into sustainment, growth or other opportunities that create value	–
NDA [17]	–	At each stage in an assessment risk should be identified, to reduce potential disadvantages or risk	–

(continued)

Table 1 (continued)

Source	Risk	Opportunity	Value
		failure, whilst at the same time seeking to maximise advantages particularly where there are opportunities to enhance the performance of the option under consideration	
Woodhouse [27]	–	Improvement opportunities with real-time information such as current asset performance, condition data and key performance indicator	"Value" varies greatly in the eyes of the beholder (…) since many (…) interests compete

3.2 *Definitions in This Paper*

In this paper, the following concepts and definitions in Table 1 are followed:

- **Value**: Best result that can be delivered by the assets of an organization (adapted from ISO 55000; ISO 55002; ISO 31000; ISO 21500; [8, 19, 27]).
  Note: Realizing value involves striving to obtain better outputs from the whole of the organizations' assets (EN 12973). This can be achieved by managing risk and opportunity (ISO 55000). Value is not absolute but relative and may be viewed differently by different parties in differing situations (EN 12973). It may be described as the relationship between the satisfaction of need and the resources used in achieving that satisfaction (EN 1325-1).
- **Risk**: Effect of uncertainty (ISO/IEC Directives) on objectives (ISO 31000). The effects of risk can be either negative or positive (ISO 9001).
  Note: In ISO management standards adopting the high-level structure established in ISO/IEC Directives, Part 1—Annex SL (ISO 9001, ISO 55001, etc.), opportunity is not the positive side of risk. However, in ISO 31000, positive risks are expressed as opportunities (and negative risks as threats).
- **Opportunity**: Set of circumstances which makes it possible to do something [14]. Opportunities include options leading to value realization.
  Note: Taking or not taking an opportunity then presents different levels of risk [14]. Opportunities can be assessed based on the best information and percep-

tion existing in the organization, to meet the needs and objectives of the different stakeholders, balanced between performance, cost and risk, leading to the best value (adapted from ISO 9001; ISO 55000; ISO 31000; ISO 21500; ISO 21504; [7, 8, 17, 27]).

4 Design and Development of the Process

4.1 Design Inputs

The inputs and the rationale for designing the model are presented in the two-entry Table 2 (technical publications), Tables 3 (standards) and 4 (scientific publications). The first entry frames the steps of the proposed value-based opportunity management process. These steps align closely to the layout of the well-known ISO 31000 risk management process. The second entry presents the various design inputs of the process and the background knowledge according to the literature review undertaken.

4.2 Process Outline

The layout of the ISO 31000 risk management process was found to be a useful foundation for outlining the proposed management process. Figure 2 represents the framework of this process. It illustrates the fundamentals of asset management established in ISO 55000 (above), the asset management system within an integrated management system (below, right) and the proposed value-based opportunity management process (below, left). This strategy of adopting a well-known standardized layout ensures that the output of the proposed process is robust and comprehensive [2] and that it adheres to principles such as consistency, generality, simplicity, correspondence with existing initiatives and adaptability [3, 13].

4.3 Steps of the Process

The core of the proposed opportunity management process covers three stages, starting with a preparatory stage in which the context for the subsequent stages is established. An assessment stage follows, including the identification, analysis and evaluation of the opportunities for the organizations to realize value from their asset base. Finally, the operationalization of selected opportunities is to be pursued through appropriate implementation. These three core stages and its steps must be

Table 2 Process steps and design inputs (technical publications)

Steps of the process		Technical publications							
		SAVE [20]	Bekefi et al. [6]	Broadleaf [5]	Stewart et al. [21]	Woodhouse [27]	Copperleaf [7]	UIC [25]	GFMAM [8]
Communications and consultation		Workshop/study	Information and communication	Workshop	Workshop	–	–	–	–
Establishing the context	Stakeholders	Pre-workshop activities	–	Information phase: pre-workshop	(Risk identification) identification of stakeholder issues and concerns	–	Educate stakeholders	Stakeholders and context	Stakeholders
	Organizational objectives and AM objectives						Review corporate objectives and risk framework	Organizational directions and asset management objectives	Organizational objectives
Opportunity assessment	Opportunity identification (including options identification)	Workshop (job plan) activities: function analysis phase, creative phase, evaluation phase	Identify risk and opportunities	Information phase: workshop	Risk identification: preliminary risk identification	Identify and prioritise problems/ opportunities	Develop the value framework	Asset strategy	Opportunities
	Opportunity analysis (value drivers, including performance, cost and risk analysis)		Manage risks and opportunities	The functional analysis phase: the ideas, options or creative phase	Risk analysis: FAST dimensioning	Define the problem Identify potential solutions		Asset management enables	Value drivers/ enablers (performance, cost, risk)
	Opportunity evaluation (including value calculation and decision-making)	Workshop (job plan) activities: development phase, presentation phase	Evaluate risks and opportunities through ROI and other methods	The evaluation or judgment phase	Risk response planning: risk response function analysis; risk response brainstorming; risk response evaluation	Evaluate and optimize timing of discrete solution Evaluate and optimize combinations of interventions	Establish a unified value framework	Performance evaluation	Decision making, balancing, value

(continued)

Table 2 (continued)

Steps of the process	Technical publications							
Opportunity implementation	Post workshop: implementation phase	–	The development phase: solution development	Risk response planning: risk response development	Assemble total programme	Executive and stakeholder sign-off	–	–
Monitoring and review	Post workshop: follow up activities	Control activities and monitoring	Post workshop development	Risk monitoring and control	–	Maintain and improve the process	Improvement	–
Recording and reporting	–	–	–	–	–	–	–	–

Table 3 Process steps and design inputs (standards)

Steps of the process		Standards					
		ISO 55002	ISO 9004	ISO 31000	ISO 21500	AS 4183	EN 12973
Communications and consultation		Communication	Communication	Communication and con-sultation	–	Facilitate stakeholder involvement and learning	Select people to be involved in the teams
Establishing the context	stakeholders	Context of the organization: understanding the needs and expectations of stakeholders	Context of the organization: relevant interested parties	Establishing the context	Stakeholders and project organization	Develop scope and identify stakeholders	Define the objectives in relation to the VM policy and programme, set targets for performance and use of resources
	Organizational objectives and AM objectives	Context of the organization: understanding the organization and its context	Context of the organization: internal and external issues	Establishing the context: establishing the internal context	Organizational strategy		
Opportunity assessment	Opportunity identification (including options identification)	Actions to address risks and opportunities for the asset management system	Performance indicators	Risk assessment (identification)	Identify opportunities	Build knowledge, understand success criteria, generate multiples ideas aimed at achieving best value	Identify the methods and the supporting processes needed to achieve the objectives
	Opportunity analysis (value drivers, including performance, cost and risk analysis)		Performance analysis	Risk assessment (analysis)	Select projects	Evaluate ideas against success criteria Develop best value and value for money alternatives, make recommendations, decisions	Measure changes in performance and use of resources
	Opportunity evaluation (including value calculation and decision-making)		Performance evaluation	Risk assessment (evaluation)			Apply the methods supporting processes

(continued)

Table 3 (continued)

Steps of the process	Standards					
	ISO 55002	ISO 9004	ISO 31000	ISO 21500	AS 4183	EN 12973
Opportunity implementation		–	Risk treatment	Projects	Prepare action plan, implement decisions and recommendations	Validate the proposals and implement them
Monitoring and review	Monitoring, measurement, analysis and evaluation management	Improvement, learning and innovation	Monitoring and review	Contribute benefits	Feed-back loop for continuous improvement	Monitor and measure outcomes, compare achievements with targets
Recording and reporting	–	–	Recording and reporting		Document key steps and decision taken	Feedback for continuous improvement

Table 4 Process steps and design inputs (scientific publications)

Steps of the process		Scientific publications			
		Haddadi et al. [9]	Parlikad et al. [19]	Olawumi et al. [18]	Kashiwagi [16]
Communications and consultation		–	–	Select team members	–
Establishing the context	Stakeholders	Identification of the most important stakeholders and their needs	Value mapping: identify stakeholders and their requirements	Information phase	Selection
	Organizational objectives and AM objectives	Identification of the owner's strategies	Establishing the context: set the objectives and scope		
Opportunity assessment	Opportunity identification (including options identification)	Idea creation	Identify the value map	Function analysis phase: look for new opportunities	Pre-planning quality control
	Opportunity analysis (value drivers, including performance, cost and risk analysis)	Idea evaluation	Assess and optimize value	Creative phase	
	Opportunity evaluation (including value calculation and decision-making)			Evaluation phase (project selection) Development phase Presentation phase	
Opportunity implementation		Implementation		Implementation phase	Management by risk minimization
Monitoring and review		Measurement	–	Reporting and follow-up phase	
Recording and reporting		–	–		

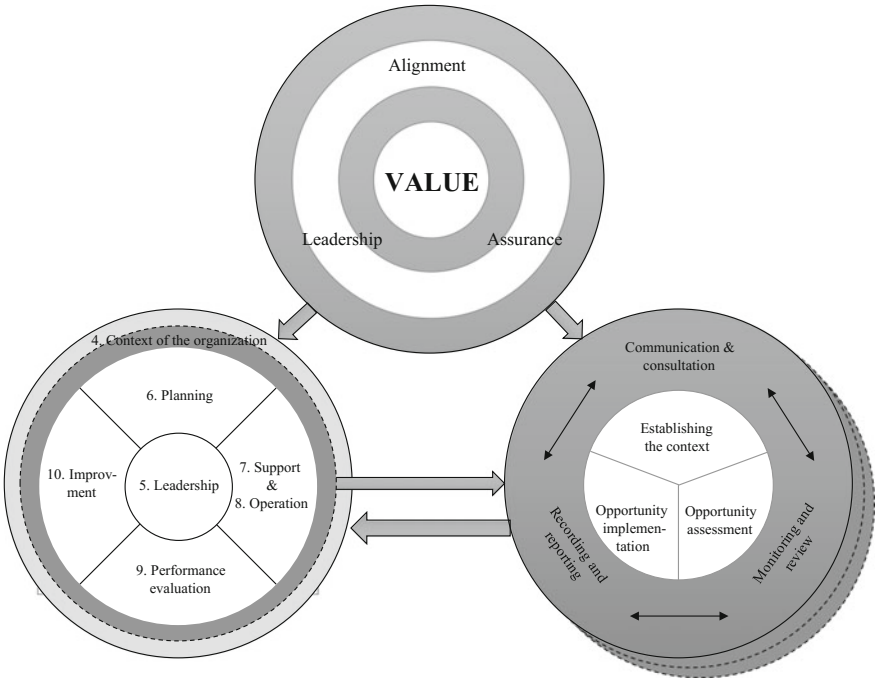


Fig. 2 Value-based opportunity management process framework (adapted from ISO 31000)

supported by appropriate communication and consultation and by monitoring and review as well. The relevance and a summarized description of the various elements of the proposed process is presented below.

- **Communications and consultation**: The engagement of both internal and external stakeholders is important as they make judgements about value based on their perceptions. These perceptions are relevant and can have a significant impact on the asset management related decisions. Thus, appropriate methods of stakeholder communication and consultation should be in place in all steps of the proposed value-based opportunity management process, namely communication and consultation plans (ISO 31000). Appropriate communication and consultation methods can help the identification and education of relevant stakeholders on value-based decision-making [7] and lead asset intensive organizations towards improvements (ISO 55002) and sustained success (ISO 9004).
- **Establishing the context**: The preparatory stage of the proposed process ensures that the objectives to be achieved are understood and aligned with the overall purpose of the organization, and that internal and external factors that can affect the subsequent stages of the process are considered (ISO 31000). This stage covers issues such as the definition of objectives, scope, problem statement and time frame for the assessment [19], or the identification of

stakeholders and their needs or the organization corporate strategies and business and asset management objectives ([7–9]; ISO 21500). The scope of asset management and of the asset management system is also to be taken into consideration in this preparatory stage (ISO 55002; [25]).

- **Opportunity identification**: The systematic identification of opportunities to realize value from assets involves an awareness of the various competing options ([8]; ISO 21500), within the context established for assessing these opportunities and the existing value chain, prior to their detailed analysis and evaluation. Some authors suggest that such an awareness can be attained through the mapping of value-drivers and associated metrics [19] or through the identification of ideas leading to value creation [9]. The use of indicators can also be used in this regard (ISO 9004), especially if formal actions to address risks and opportunities are planned within the framework of an asset management system (ISO 31000; ISO 55002).
- **Opportunity analysis (including risk analysis, cost analysis e performance analysis)**: The results of a complete analysis shall indicate the impacts and measurable benefits in case the opportunities are implemented (ISO 21500; ISO 31000), namely in terms of performance, cost and risk ([8]; ISO 9004; ISO 31000). This provides the basis for informed decisions and optimized asset management [9, 19] and the subsequent evaluation of the relative importance of competing opportunities against the goals of the organization and the established asset management strategy.
- **Opportunity evaluation (including value calculation)**: Evaluating the relative importance of competing opportunities involves the calculation of the value of each individual opportunity or group of opportunities, given expected scenarios with varying time frames identification and education, and the decision of implementing, or not, each opportunity ([9]; ISO 21500; ISO 31000). The evaluation process shall be undertaken against the established context (ISO 31000; ISO 9004), be evidence-based (ISO 55002), and may include a sensitivity analysis [19].
- **Implementation of opportunity**: Evaluation results support informed decision-making by the responsible management (ISO 21500) and the planning and implementation of feasible projects [9] that could transform some or all of the relevant opportunities into realized benefits (ISO 31000). The deliverables of these projects are expected to demonstrate that value is being realized from the asset base in line with the organizational strategy and goals.
- **Monitoring and review**: The core stages of the proposed opportunity management process should involve checking or surveillance with ongoing oversight by top management and those with delegated authority, to ensure that the context of the process is being properly established, that opportunity identification and assessment are being undertaken and that value-adding projects are selected and implemented according to plan. All phases of the proposed process should be well documented and the lessons learned with successes, or otherwise, should be recorded and incorporated into the opportunity management framework, namely towards improving the asset management activities [9] or the asset management system (ISO 55002; ISO 9004).

- **Recording and reporting**: The process and its outcomes should be documented and reported through appropriate mechanisms for feedback of results (ISO 31000; EN 12973, 2000), namely the key activities and the reasoning of the value-based decisions making process. The documented information shall be auditable and available for later use (AS 4183, 2007). Recording and reporting supports the investigation of new opportunities and risk management towards value creation and protection (ISO 31000; [16, 18]).

4.4 Potential Application of the Proposed Process

The proposed process can be used by decision makers is asset intensive organizations to make informed choices, prioritize actions and distinguish among alternative options. It is intended to explicitly demonstrate the best benefit (value) achieved by competing option balancing performance (short and long term), cost (capex/opex, short and long term) and risk (financial and non-financial).

This systematic, structured and timely approach aims at contributing to higher efficiency in asset intensive organizations. Figure 3 illustrates the application of the proposed process to a simplified case study. It presents a sample of the record of three steps of the process (opportunity identification, opportunity analysis and opportunity evaluation).

5 Final Remarks

Recent studies [22] show that asset intensive organizations can optimize asset portfolio investments and attain benefits using value quantification techniques. These techniques are indeed of great importance when competing investment options are involved in the face of multiple alternative scenarios with various timeframes. That is the case in multi-sector asset intensive organizations. However, value-related concepts and techniques are used interchangeably by different sectors, and even by organizations within the same sector. For example, there is still some inconsistency in the definition and use of risk within asset management [26]. This paper deals with the combined use of such concepts and techniques within asset management activities, namely those developed towards the optimized balancing of cost, risk and performance. Namely, it provides an overview of the concepts of value, opportunity and risk, towards solving some inconsistencies such as the one mentioned above. It also proposes a value-based opportunity management process that can be used by asset intensive organizations with varying levels of asset management maturity.

The proposed downstream process can be used by organizations in any stage of their asset management journey. This process can be incorporated into a fully

Opportunities identification				Opportunities analysis																				Opportunities evaluation							
Opportunities	Options	Scheduling		Value drivers	Value metrics	Performance		Cost				Risk			Impacts on Strategic Objectives				Impacts on Strategic Stakeholders				Value Short Term			Value Long Term			VALUE	Decision Making
		Project inception	Project duration			Short Term	Long Term	CAPEX		OPEX		Financial	Non Financial		Mo	AM	In	R/P	Sh	Co	Em	So	Satisfaction of Needs	Use of Resources	VALUE - ST	Satisfaction of Needs	Use of Resources	VALUE_LT		
								Short Term	Long Term	Short Term	Long Term	(+/-)	(+)	(-)																	
Investment #1	1a - Simply Repair	T < 2 Years	12 months	Safety Reliability Financial	Accidents Asset Condition Budget			0	0	$$	$$$	($)	5	10									≈							
	1b - Upgrade	2 ≤ T< 5 Years	18 months					$$$	0	$	$	($$)	10	5																
	1c - Replace	5 ≤ T< 10 Years	24 months					0	$$$$	$$	$	($$$)	20	2																
	...	...	...			...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Investment #2	2a - Do nothing	T < 2 Years	-	Environmental Reputation Customer satisfaction	Env. Benefits Reputation risk Service quality			0	0	0	0	$$$	1	20									≈			-	-	-		
	2b - Improve	T < 2 Years	6 months					0	0	$$	0	($)	10	10												-	-	-		
	...	...	...			...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Investment #3	3a - Maintain	T < 2 Years	12 months	Availability Reliability Financial	Level of service Asset Condition Budget			0	0	$$$	$$	($$)	10	10																
	3b - Renewal	2 ≤ T< 5 Years	24 months					$$$$	0	$	$	($$$)	20	2																
	...	...	...			...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Fig. 3 Sample of process register (example)

matured asset management system that allows the systematic quantification and demonstration of value realization. It may also be used as a guideline, in a less formal way, towards progressively promoting asset management practices within the organization such as life cycle cost analysis, performance measurement frameworks and risk management techniques. The proposed process is to be further developed and validated with complementary empirical evidence through its application to a set of real case studies in selected asset intensive organizations. This is presently being undertaken as part of the PhD thesis of the first author of this paper. The resulting outcomes will be presented in future publications.

References

1. Abraham D, Gibbons P (2016) Preparing for the journey. Assets s.l.: s.n
2. Almeida N, Sousa V, Alves Dias L, Branco F (2010) A framework for combining risk-management and performance-based building approaches. Build Res Inf 38(2):157–174
3. Almeida N, Sousa V, Alves Dias L, Branco F (2015) Managing the tehnical risk of performance-based building structures. J Civ Eng Manag 21(3):384–394
4. AMC (2014) Framework for asset management, 2nd edn. s.l.: Asset Management Council
5. Broadleaf (2010) Value management. Broadleaf Capital International Pty Ltd., Pymble
6. Bekefi T, Epstein J, Yuthas K (2008) Managing opportunities and risks. In: Management accounting guideline, CMA, AICPA, CIMA
7. Copperleaf (2016) Value-based decision making (white paper). s.l.: Copperleaf Technologies Inc.
8. GFMAM (2016) The asset management landscape, 1st edn. s.l.: Global Forum on Maintenance and Asset Management
9. Haddadi A, Johansen A, Andersen B (2016) A conceptual framework to enhance value creation in construction projects. Proc Comput Sci 100:565–573
10. IAM (2015) Asset management—an anatomy. s.l.: The Institute of Asset Management
11. IP–SA (2016) Management report 2016. Available at http://www.infraestruturasdeportugal.pt/sites/default/files/files/files/2016_annual_report.pdf
12. IPWEA (2015) International infrastructure management manual, 5th edn. Institute of Public Works Engineering, Australia
13. IFIC (1998) FRISCO—a framework of information systems concept. The FRISCO Report. Web edition. International Federation for Information Processing
14. ISO (2015) ISO 9001:2015 Risk-based thinking. International Organization for Standardization
15. Jemana Electricity Networks (Vic) Lda (2015) 2016–20 electricity distribution price review regulatory proposal. Attachment 7–2. Asset management framework and governance. Jemena Limited
16. Kashiwagi D (2010) Case study: best value procurement/performance information procurement system development. Available at https://aiahouston.org/
17. NDA (2016) The NDA value framework. s.l.: Nuclear Decommissioning Authority
18. Olawuni T, Akinrata E, Arijeloye B (2016) Value management—creating functional value for construction project: an exploratory study. World Sci News 54(2016):40–59
19. Parlikad A, Srinivasan R (2016) Whole-life value-based decision-making in asset management. s.l.: Cambridge Centre for Smart Infrastructure & Construction
20. SAVE (2007) Value standard and body of knowledge. SAVE International
21. Stewart R, Brink G (2012) Function driven risk management. Available at http://www.valuefoundation.org/.e

22. Tamini I, Beullens DP, Sadnicki S (2016) Quantifying the benefits of investment portfolio optimisation versus prioritisation for asset intensive organisations
23. Too E (2009) Capabilities for strategic infrastructure asset management. s.l.: Submitted in fulfilment of the requirements for a degree of Doctoral of Philosophy. Faculty of Built Environment & Engineering. Queensland University of Technology
24. Trindade M, Almeida N (2016) Value-based management of constructed assets. 2nd encounter on quality and innovation in construction, QIC2016. Lisbon (in Portuguese), LNEC 21–23 Nov 2016
25. UIC (2016) Practical implementation of asset management through ISO 55001. s.l.: UIC Railway Application Guide
26. Wijnia YC (2016) Processing risk in asset management. Exploring the boundaries of risk based optimization under uncertainty for an energy infrastructure asset manager. s.l.: PhD thesis, TU Delft, NL
27. Woodhouse J (2014) Asset management decision-making: the salvo process, 1st edn. s.l.: The Woodhouse Partnership Ltd.

The Design and Performance of a Novel Vibration-Based Energy Harvester Adopted Various Machine Rotational Frequencies

Peter W. Tse and Shilong Sun

Abstract The research in vibration-based energy harvesting by using piezoelectric transducers (PZT) has been attracting high attention in the research field of green energy. It can absorb destructive vibrations generated by any operating machine and then convert the collected vibrations into useful electrical energy. Such electrical energy can be used to power up low-powered and wireless sensors that are used to monitor the health condition of the machine. To absorb maximum vibration energy generated by the machine, the criterion is to match the PZT's resonance frequency to the dominant vibration frequency of the machine, which is usually the fundamental rotational frequency of machine. This paper presents a novel design of PZT energy harvester whose resonance frequency can be tuned by using several cantilever-type beam structures attached to the PZT energy harvester. The beam structures include the T-folded, the E-folded without a tip mass, the E-folded with one tip mass and the E-folded with two tip masses. A finite element method was used to investigate the resonance frequency and the operational bandwidth of the harvester and its outputted voltage and power. After finding the characteristics from the finite element method, experiments were conducted to confirm the findings. The results show that the best structure is the structure of E-folded with two tip masses. It provided a tuneable resonant frequency of the harvester from 30.4 to 18.18 Hz. With help of such tuneable harvester, it can easily adapt to different machines that are running at different speeds and scavenge the highest electrical energy from collecting the destructive vibration. Most importantly, since it absorbs the destructive vibrations, less damage will be caused to the machine, making the life of machine longer.

P. W. Tse (✉) · S. Sun
City University of Hong Kong, HKSAR, Hong Kong, China
e-mail: meptse@cityu.edu.hk

S. Sun
e-mail: shilosun-c@my.cityu.edu.hk

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_61

1 Introduction

Energy harvesters have been widely studied by many researchers in recent years. An energy harvester is defined as harvesting ambient energy or wasted energy, might be acoustic noise, solar radiation, wind or mechanical vibrations, into usable energy, such as electric energy. It is used in many different disciplines including medical, biology, mechanical and also computer science. Therefore, the energy harvester field has a significant effect to the improvement of our society and people's daily life. Among these ambient energy, the mechanical vibration has relative high power density makes them the most accessible ambient energy source in many applications [8]. In order to scavenge maximized energy from the mechanical vibration, an electromechanical system, composed of piezoelectric material and mechanical structures must be designed, which the resonant frequency can be tuned to adapt the ambient vibration frequency sources or its bandwidth can be widened to adapt the system natural frequency. Over the past few years, various methods for resonance frequency tuning or widening of the bandwidth have already been described, including magnetic, electrical and mechanical methods [2, 4, 6, 10].

Firstly, some researchers have already demonstrated resonance frequency shift by utilizing an external magnetic field, for example the magnetic force [9, 13], attractive or repulsive force [11] to adjust the electromechanical system stiffness and then shift the resonance frequency. Secondly, the coefficient and parameters of the electromechanical structure used to tune the resonance frequency also have been widely studied [3, 7]. The mechanical stress [1], the electrical damping [5], and the number of degree freedom [12] have been studied to tune the resonant frequency and broaden the operational bandwidth. However, there still exists a great weakness for the above studies: it is hard to obtain a high power output at a lower resonance frequency simultaneously. The purpose of this research is to find the optimal structure design which can scavenge the maximum power output at a relative low resonance frequency.

In this research work, four different composite structures were developed to scavenge the mechanical vibrations and convert them into useful electric energy. These four beam structures are the T-folded, the E-folded without any tip mass, E-folded with one tip mass and E-folded with two tip masses. All the beam structures were designed based on the cantilevered concept. The first stage of research was to find the optimal design that could tune the resonance frequency in a relative low frequency range. The second stage was to find out the operational bandwidth, the possible voltage and power outputs for each of the four beam structures. Finally, the optimized design of the energy harvester was obtained and the output of maximum power output could be guaranteed. This reported paper is organized as follows: the design and model for each type of piezoelectric vibration energy harvester are introduced in Sect. 2. Section 3 describes the experimental setup and the process of investigation. The observed results and discussions are given in Sect. 4. Finally, conclusions are offered in Sect. 5.

2 Model of Vibration-Based Energy Harvester

Traditionally, a single degree of freedom (SDOF) piezoelectric energy harvester design is composed of a cantilever beam and a piezoelectric layer. In order to achieve the goal of optimizing the structure design, four different basic host structure designs were proposed based on the traditional piezoelectric energy harvester. The first design was called, T-folded model, consisted of one main beam and one tip mass beam. Then, we placed two wing beams at each end of the tip mass beam to develop an E-folded model. In order to study the symmetry and asymmetry structure's effect on the proposed piezoelectric energy harvester models' performance, one tip mass and two tip masses are added to the free end of wing beams, respectively.

Figure 1 shows the four proposed piezoelectric energy harvester models. The main beam, the tip mass beam and the wing beam are composed by aluminium material and the tip mass is composed by iron. The length, width, height, and the other parameters are shown in Table 2. The piezoelectric material was laminated on the main beam for all the four designs shown in Fig. 2 (Table 1).

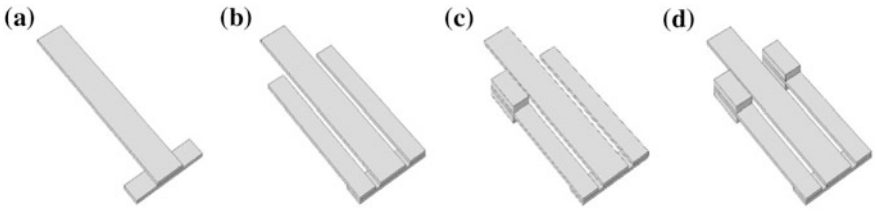


Fig. 1 The models for the four proposed piezoelectric energy harvester: **a** the T-folded, **b** the E-folded without any tip mass, **c** the E-folded with one tip mass and **d** the E-folded with two tip masses

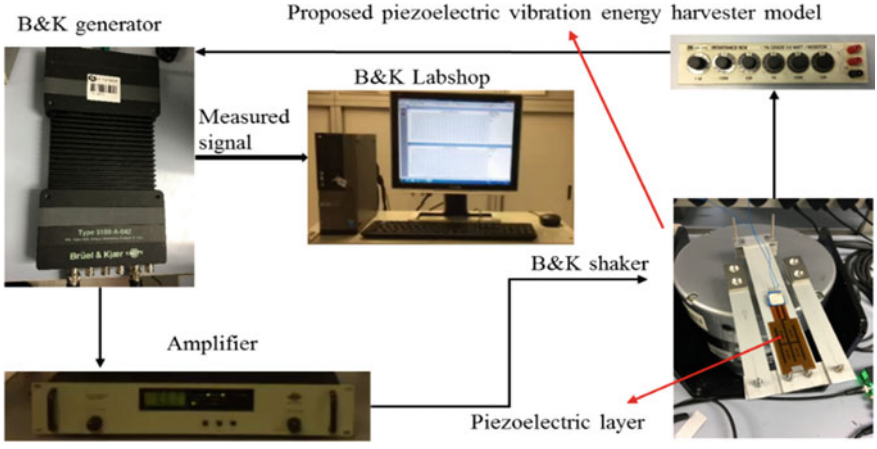


Fig. 2 Schematic of the experimental setup for testing the designed energy harvester

Table 1 Parameters of the designed piezoelectric energy harvester

Parameter and description		Value	Unit
Main beam	Length × Width × Height	200 × 30.5 × 2	mm
Tip mass beam	Length × Width × Height	80 × 19 × 2	mm
Wing beam	Length × Width × Height	150 × 19 × 2	mm
Tip mass	Length × Width × Height	39 × 19 × 4.5	mm
Aluminum	Density	2700	kg/m^3
	Young's modulus	70×10^9	Pa
	Poisson's ratio	0.35	–
Iron	Density	7860	kg/m^3
	Young's modulus	152×10^9	Pa
	Poisson's ratio	0.27	–
Piezoelectric material	Length × Width × Height	49.3 × 25.4 × 0.6	mm
	Young's modulus	55	GPa
	Density	7700	kg/m^3
	Area moment of inertia	4.57e-13	m^4

3 Experimental Study

The experimental testing system for the proposed models is shown in Fig. 2. It was composed by the B&K generator, the Pulse Labshop, the Amplifier and the Shaker. The excitation signal was produced by the generator and then transferred to the amplifier. After through the amplification effect, the excitation signal then was transmitted to the shaker and drove the proposed PVEH models vibrate. All the beam components were made up by the aluminium material except the tip masses at the end of the wing beams and the piezoelectric layer laminated on the top surface of the main beam as the collector of the electric energy converted from the system vibration.

The amplifier was used to refine and improve the signal generated from the B&K generator to drive the shaker to vibrate the proposed piezoelectric vibration energy harvester (PVEH) models. The signal level was set up to 0.2 mV/s. The excitation signal was a swept sine wave which frequency varies from 1 Hz to 200 Hz at a rate of 2 Hz/s. The voltage response and resonance frequency were measured by the B&K Pulse Labshop software which was installed on a windows computer.

4 Results and Discussions

4.1 The Study of Resonance Frequency

As well known, for an electrometrical system only when the resonance frequency is matched with the ambient vibration frequency, then the power output can be optimized and maximized. However, the operational bandwidth for a resonance

Table 2 Comparison of resonance frequency analysis both on simulation and experiment

Model type		Resonance frequency (Hz)		
		Simulation	Experiment	Error (%)
The T-folded	1 mode	28.095	30.4	8.20
The E-folded without any tip mass	1 mode	22.552	24.2	7.31
	2 mode	54.052	55.0	1.75
The E-folded with one tip mass	1 mode	20.448	20.0	2.19
	2 mode	23.87	23.6	1.13
	3 mode	68.81	63.8	7.28
The E-folded with two tip masses	1 mode	18.783	18.18	3.21
	2 mode	23.597	23.6	0.01
	3 mode	24.36	26.8	10.02

frequency usually is very narrow which is very bad for the power output. Therefore, we want to propose a novel model which can provide a resonance frequency under 50 Hz while it can exert a relative high energy density simultaneously because most of the industrial rotating vibration machine vibration frequency are under 50 Hz or around. Table 2 shows the comparison of resonance frequency analysis both on simulation and experiment analysis of the four different piezoelectric energy harvesters. The simulation results were calculated in COMSOL Multiphysics using frequency domain analysis.

For the T-folded model, the simulation results show that the first resonance frequency 28.095 Hz. For the E-folded model with two wings added on each end of the tip mass beam, the first and second resonance frequency are 22.552 and 54.052 Hz. Due to the second resonance frequency still is larger than our designed goal, so a tip mass at the free end of one of the wing beam is placed to study its performance further. The first three modal frequencies for the E-folded one tip mass are 20.448, 23.87 and 68.81 Hz and for the E-folded two tip masses, are 18.783, 23.597 and 24.36 Hz. Only E-folded with two tip masses can offer a multi-mode (3 mode) resonance frequencies are all under 50 Hz. Here we define the peak-to-peak frequency value under 50 Hz as the operational bandwidth. Therefore, the operational bandwidth corresponding to the asymmetric and symmetric structure are around 3.4 and 5.42 Hz which means adding two tip masses on the free end of both wing beams can broaden the bandwidth significantly compared to only one tip mass to the wing beam.

4.2 The RMS Voltage Response at Different Frequencies

Figure 3 shows the comparison of RMS voltage response as a function of swept sine frequency excitation range from 0 to 200 Hz. The left and right figures are experimental and simulated RMS voltage respectively. The RMS voltage

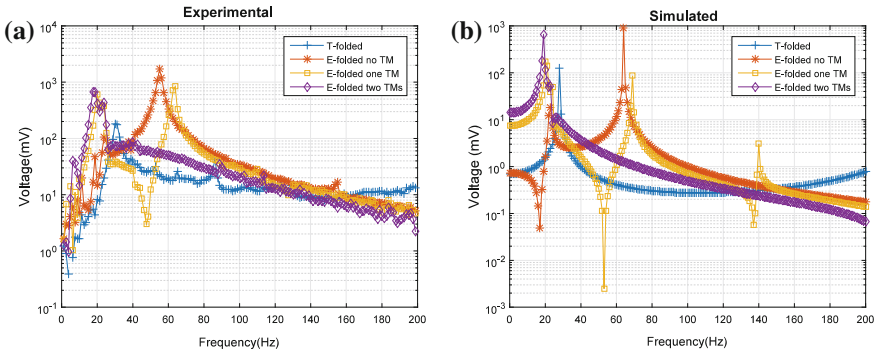


Fig. 3 Comparison of the RMS in voltage versus various frequencies generated by the swept sine frequency excited at a range from 0 to 200 Hz: **a** the experimental results and **b** the simulation results

performance of four proposed design structures are studied by both experimental and simulated methods. The results showed that the first peak value of RMS voltage is 800.1 mV at a frequency of 18.18 Hz and the second peak is 435.4 mV at 23.6 Hz for E-folded with two tip masses. Moreover, the experimental operational bandwidth is 5.42 Hz having a good agreement with the simulation.

In sum, the E-folded with two tip masses structure is the optimal electromechanical system among the four piezoelectric vibration energy harvesting designs. It has expressed a good performance that this design can collect high RMS voltage at a relative low frequency simultaneously which have been demonstrated by the simulation results. The next step is to study how to make the maximum electric power output and improve the performance of the energy density of the whole structure.

4.3 The Voltage and Power Output for Resistor Load

To investigate the performance of voltage and power output of the proposed E-folded two tip masses design, an external electric circuit was designed and added to the electromechanical system shown in Fig. 2. Here a simple resistor R is considered as the external electric load which can be used to optimize the maximum power output. Figure 4 shows the voltage and power output response versus different resistor load at frequencies of 18.18, 23.6, and 26.8 Hz under acceleration of 1 m/s^2. The experiment was conducted under the resistor load varying from 1 kΩ to 1 MΩ. The results show that the E-folded with two tip masses achieved its maximum voltage output at the resistor load of 1 MΩ. The voltage response changed a little when the external load is larger than 1 MΩ. The voltage output is

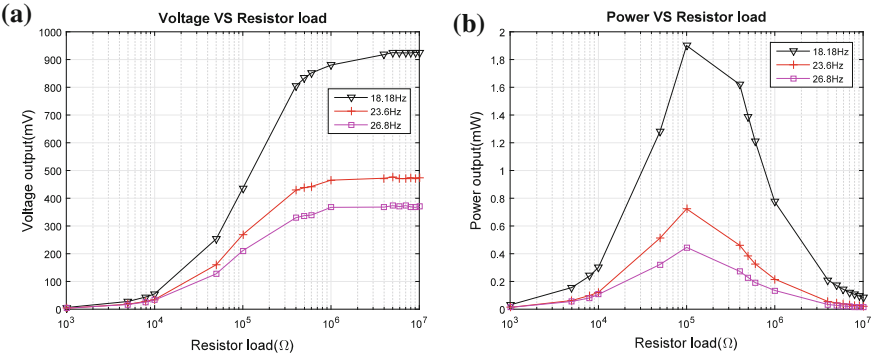


Fig. 4 The voltage and power output response versus different resistor load at frequencies of 18.18, 23.6, and 26.8 Hz under acceleration of 1 m/s^2

0.88 V@18.18 Hz, 0.465 V@23.6 Hz and 0.367 V@26.8 Hz respectively. Clearly, the first voltage is larger than the second and the third one based on the experimental results.

For the power response versus resistor load at frequencies of 18.18, 23.6, and 26.8 Hz, the maximized power was obtained at 100 kΩ for all the three resonance frequencies. The maximized power at frequency of 18.18 Hz is 1.9 mW while the power for the second frequency 23.6 Hz is 0.72 mW and the power for the third frequency 26.8 Hz is 0.45 mW. The power output at the first resonance frequency dominated the main power output of the whole structure compared to the other two resonance frequencies.

5 Conclusions

The performance of four different designs of vibration-based energy harvesters by using four different types of cantilevered beam structures were reported in this paper. The four designs are beam structures that include the T-folded, the E-folded without any tip mass, the E-folded with one tip mass and the E-folded two tip masses. The resonance frequency, the operational bandwidth, and the outputs of voltage and power were analysed and compared. The finite element analysis was used to simulate the four piezoelectric energy harvesters so as to investigate the possible tuneable range of resonance frequency of each harvester. The result showed that the E-folded with two tip masses exhibited a trend that the first three modal resonance frequencies were under 50 Hz. The voltage and power output were measured by applying a resistor load. In summary, the structure of the E-folded with two tip masses could yield the highest outputs in voltage and power at a relative low resonance frequency range and has a wider bandwidth that can easy match with most of the dominant vibration frequency or often the rotational frequency of rotary machines.

Acknowledgements 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_11201315).

References

1. Eichhorn C, Goldschmidtboeing F, Woias P (2009) Bidirectional frequency tuning of a piezoelectric energy converter based on a cantilever beam. J Micromech Microeng 19:094006
2. Erturk A, Inman DJ (2008) A distributed parameter electromechanical model for cantilevered piezoelectric energy harvesters. J Vibr Acoust 130:041002–041002
3. Eun Y, Kwon D-S, Kim M-O, Yoo I, Sim J, Ko H-J, Cho K-H, Kim J (2014) A flexible hybrid strain energy harvester using piezoelectric and electrostatic conversion. Smart Mater Struct 23:45040–45040
4. Ferrari M, Ferrari V, Marioli D, Taroni A (2006) Modeling, fabrication and performance measurements of a piezoelectric energy converter for power harvesting in autonomous microsystems. IEEE Trans Instrum Meas 55:2096–2101
5. Ooi BL, Gilbert JM, Aziz ARA (2015) Switching damping for a frequency-tunable electromagnetic energy harvester. Sens Actuators A 234:311–320
6. Ramezanpour R, Nahvi H, Ziaei-Rad S (2016) Electromechanical behavior of a pendulum-based piezoelectric frequency up-converting energy harvester. J Sound Vib 370:280–305
7. Rivadeneyra A, Soto-Rueda JM, O'Keeffe R, Banqueri J, Jackson N, Mathewson A, López-Villanueva JA (2016) Tunable MEMS piezoelectric energy harvesting device. Microsyst Technol 22:823–830
8. Shu Y, Lien I (2006) Efficiency of energy conversion for a piezoelectric power harvesting system. J Micromech Microeng 16:2429
9. Singh KA, Kumar R, Weber RJ (2015) A broadband bistable piezoelectric energy harvester with nonlinear high-power extraction. IEEE Trans Power Electron 30:6763–6774
10. Tan T, Yan Z, Hajj M (2016) Electromechanical decoupled model for cantilever-beam piezoelectric energy harvesters. Appl Phys Lett 109:101908–101908
11. Wickenheiser AM, Garcia E (2010) Broadband vibration-based energy harvesting improvement through frequency up-conversion by magnetic excitation. Smart Mater Struct 19:65020
12. Xiao H, Wang X, John S (2016) A multi-degree of freedom piezoelectric vibration energy harvester with piezoelectric elements inserted between two nearby oscillators. Mech Syst Signal Process 68–69:138–154
13. Zhou S, Cao J, Erturk A, Lin J (2013) Enhanced broadband piezoelectric energy harvesting using rotatable magnets. Appl Phys Lett 102:173901

The Design of a Novel Line-Array Type of Laser Source for Non-contact Guided Waves to Inspect the Integrity of Plates

Peter W. Tse and Jingming Chen

Abstract Popular guided waves (GW) transducers employed for emitting and receiving GW when inspecting the integrity of different objects include the contact and non-contact ones. Contact transducers that must be physically mounted on the objects are piezoelectric transducers (PZT) and magnetostrictive sensors (MsS). The non-contact method can consist of laser-based transducers. However, the contact transducers could be difficult to apply to some operating situations, such as measuring objects that have high surface temperature, moving target, complex geometries and rough surfaces, and hazardous environment etc. These operating situations limit the possibility of mounting the contact transducers on the inspected objects. Therefore, it is necessary to design a non-contact type of GW transducer. This paper reports the feasibility of using laser to generate the desired GW wave mode, the novel design of optical transducers based on an integrated optical Mach-Zehnder interferometer (IOMZ), and the results of using such novel laser-based GW to inspect the defects occurred in plates. With the help of such IOMZ system to generate the required laser-based GW wave mode, the location and geometrical extents of any defect occurred in an aluminium plate can be successfully determined.

P. W. Tse (✉)
The Smart Engineering Asset Management Laboratory (SEAM), Department of Systems Engineering & Engineering Management (SEEM), City University of Hong Kong, Hong Kong, China
e-mail: Peter.W.Tse@cityu.edu.hk

J. Chen
SEAM, Department of SEEM, City University of Hong Kong, Hong Kong, China
e-mail: jingmchen2-c@my.cityu.edu.hk

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_62

1 Introduction

Inspection using guided waves (GW) has multiple distinguishing features. It includes the long distance inspection with less energy attenuation and the use of a single transducer to emit GW and receive the reflected GW signals. One of the promising applications including the use of GW to detect defects occurred in aluminium alloy beams and plates. To use of GW methods for the detection or identification of macroscopic anomalies such as defect, corrosion and incipient buried micro-cracks, even in the early fracture stage is very important for ensuring structural safety and integrity. The popular transducers employed by GW include the contact and non-contact ones. The contact transducers that must be physically mounted on the objects are piezoelectric transducers (PZT) and magnetostrictive sensors (MsS). The non-contact ones can be laser-based transducers. In the past two decades, laser technology has been widely and rapidly developing. It has attracted great interest to engineers and scholars to employed laser technology in non-destructive testing (NDT) applications.

Normally, there are two kind of typical laser sources - the point and the line array. In 1980, the laser heated point source was considered as a thermoelectricity means to generate elastic waves that propagating inside metallic structures [9]. However, a major weakness associated with laser technology is its poor signal-to-noise ratio (SNR). To conquer this weakness, in 2000, line source was designed for laser systems. According to the experimental result, line source laser can provide amplitude directivity pattern, which is identical to that produced by point source in an aluminium half-cylinder [2]. Laser-based GW that employed point source to inspect thin tube was discussed in [8]. However, due to the broad bandwidth characteristics of the laser source, the emitted GW signal contained infinite GW wave modes that propagated in all directions, making complications in the analysis of the received GW signal. Fomitchov et al. made laser beam into a line rather than a circular spot so that more GW energy could be emitted into the inspecting object while keeping the energy density at a low level [4]. Sohn et al. used scanning laser-generated line source instead of the conventional GW's pulse-catch or pulse echo method for emitting and receiving GW signals [10]. Such source increased the performance in SNR of GW signals and proved that the thermo-elastic line source could be generated as a monopolar surface wave, which dramatically helped to indicate the presence of the defects. During the experiment, an arrayed line slit method was used in the laser beam as the GW emitter and a set of dual air-coupler as the GW receiver. Although the laser line source improves the directivity in GW wave propagation, the received GW signal still has a broad bandwidth characteristic.

In order to generate the appropriate GW mode by using the non-contact laser method, Kim et al. proposal a laser line array for GW inspection [5]. The arrayed laser beam was emitted to the surface of an aluminate plate in order to enhance the focusing ability and spatial resolution. For this purpose, a slit mask was designed to convert the laser point source to a line array source [5–7]. The mask could help to shorten the board bandwidth characteristics of laser so that desired GW mode in

narrower band could be emitted. However, different masks have to be made in order to emit different type of wave mode. Moreover, the mask would absorb part of the emitted GW energy. To avoid the shortcomings of mask, in this paper, an innovative and novel laser-based GW inspection has been developed. The emission of the desired GW mode is performed by an integrated optical Mach-Zehnder interferometer (IOMZ). It is made from totally optical lens to form the desired laser beam pattern for converting the laser point to desired line array, which ultimately can be used to generate the desired GW mode for inspection purpose.

2 The Design of IOMZ for Emitting the Desired GW Mode in a Non-contact Means

To achieve the proposed method of using laser interferometry, a green laser system based on IOMZ interferometer is proposed for the actuator to emit the GW. The experimental setup is containing of a pulsed Nd:YAG laser with wavelength of 532 nm with lenses as shown in Fig. 1. First, the laser beam is passing through a beam splitter, which divides the incident light and sends them in two line mirrors and then recombines them through the beam combiner.

The conventional Mach-Zehnder interferometer (MZ) uses a pair of beam splitters to separate the laser beam into two arms and then merges them back to form interference fringes [3]. The principle of MZ is used here again as the polarizing cube beam splitters divide and recombine different laser beams traveling from different directions to achieve the effect of laser interferometry. The resultant laser beams then pass through another pair of beam splitters/combiner and mirrors to ultimately form the expected laser lines. The convex lens are then used to adjust the shape and size of the final lines which are belonged to light interference fringes. The above setup by using totally optical lenses to convert the laser point source to line array with specific number of lines and width of each line is named by us as IOMZ, which is a completely new design to generate the expected laser-based GW mode and signal for inspecting the integrity of plates and beams.

To determination of the width of each laser line which is generated by the interference of laser light, the wavelength matching method [1] is used to select the desired GW mode according to the dispersion curves:

$$V = f\lambda = f\Delta s = fd\left(\frac{\Delta s}{d}\right) \tag{1}$$

where V is the phase velocity of the A_0 mode propagate in a thickness of d = 2 mm aluminum plate. fd is the frequency of desired mode propagating in 2 mm thickness and Δs is the space of line-arrayed source. The relation between the wavelength λ, frequency f, the phase velocity V refer to the desired mode of guided wave, thickness of material d, and then the Δs can be calculated. Once the desired propagation mode have confirmed, the relevant slit space of the interference of laser can be calculated.

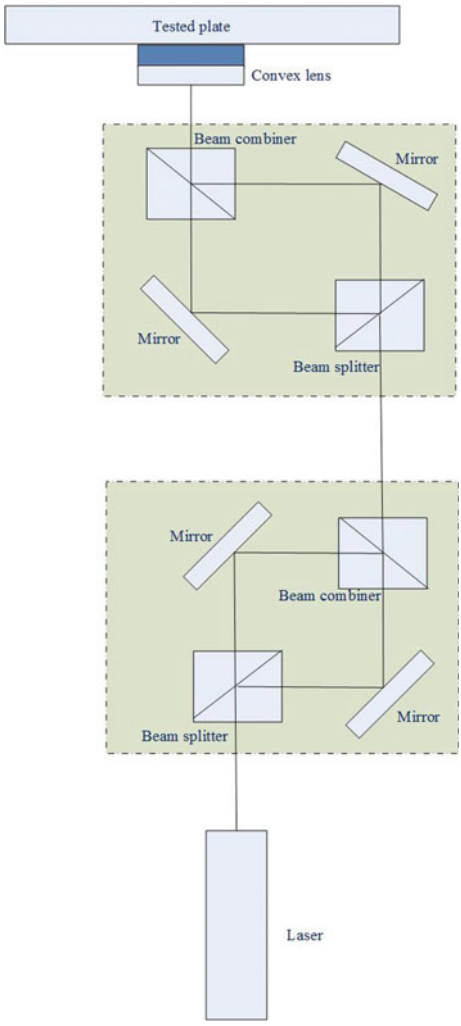


Fig. 1 The scheme of making the laser based IOMZ

The selection of desired GW mode is depended on the dispersion curves as shown in Fig. 2. The A_0 mode is characterized predominantly by the out-plane displacement, while the S_0 mode is characterized predominantly by the in-plane displacement. The reason we chosen A_0 mode is because the laser receiver is sensitive to the out-plane displacement. Therefore, the A_0 mode was applied in this study. A desired GW A_0 mode can be emitted by satisfying an appropriate spacing of line-arrayed source in terms of wavelength of A_0 mode. Equation 1 could be applied to other materials such as aluminum, steel, and other isotropic materials.

According to Eq. 1, the slop = $\Delta s/d$, the velocity, V = 2 mm/μs and the fd = 0.3 MHz ×2 mm = 0.6 MHz-mm, which is the space of line-array source, is 6.6 mm as indicated in Fig. 3 with the picture of the generated line-array laser source.

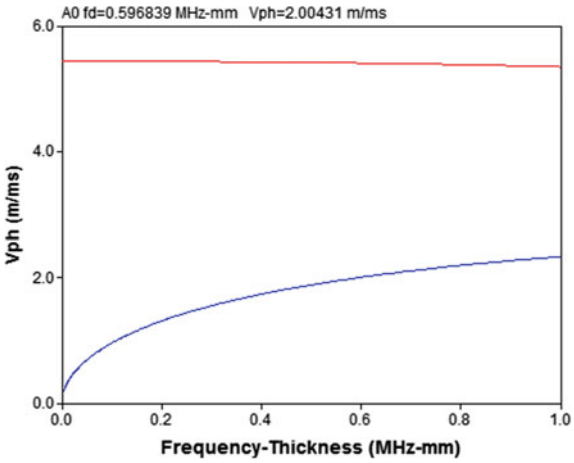


Fig. 2 The dispersion curve of the tested aluminium plate and the selected GW mode, A_0, and the expected emission frequency

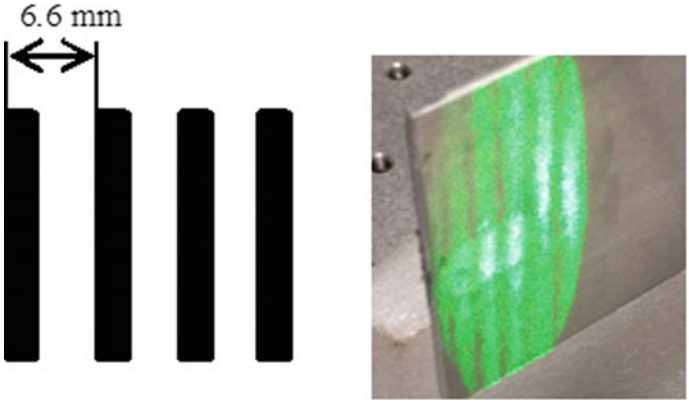


Fig. 3 The space width of the generated line-array laser source (left diagram) and the image of the line array by IOMZ (right diagram)

3 Experimental Verification and Result Analysis in Aluminum Plate

To validate the effectiveness of IOMZ, an aluminum plate with a thickness of 2 mm in health condition and in defective condition were investigated. A piece of iron with epoxy adhesives was stuck to part of the surface of the tested plate to simulate as a defect. The experiments were conducted using a Nd:YAG pulsed laser equipped with the IOMZ to emit the desired GW A_0 mode. The GW signal reflected the defect and the other end of the plate were received by the laser Doppler

vibrometer. The drawing of the tested plate, its length, the location of defect is shown in the top diagram of Fig. 4. The two GW propagating paths, which include the distance of the GW propagating from the laser emission point (excitation) to the laser measurement point (path A) and from the laser emission point to the defect location and then propagated back to the laser measurement point (path B) are shown in the bottom diagram of Fig. 4.

The experimental setup and instruments of the laser-based GW system for inspecting the plate in pulse-echo mode are depicted schematically in Fig. 5. The Nd:YAG pulsed laser was used to emit the GW at a normal angle. A laser vibrometer was used to receive the reflected GW signal. When the GW signal was emitting as a conventional laser point source into the tested plate, the GW signal, which was reflected by the defective plate, was received by the laser vibrometer. Figure 6a shows the temporal waveform of the propagating GW signal (top left diagram) and its corresponding frequency spectrum (bottom left diagram). To ease revealing the relationship of time and frequency of the received GW signal, the collected signal was further analysed by a signal processing method, called continuous wavelet transform (CWT) as shown in Fig. 6b. The received GW signal reflected by the defect was inevitably contaminated by noise and disturbance as can be seen in Fig. 6b. The existence of high noise and disturbances were due to the board bandwidth characteristic of the laser-based GW transducer that collected all signals from low to high frequencies. Hence, even if the reflected GW signal was filtered by CWT, the true defect-related GW signal was not easy to be revealed in Fig. 6b.

With the help of IOMZ, the emitted laser-based GW signal becomes narrower in bandwidth and the emitted GW frequency range can be selected by laser lines that are separated by a specific width. Again, Fig. 7a shows that the application of an IOMZ interferometer-based laser excitation system to the normal plate can greatly improve the SNR and facilitate the interpretation of the signals as compared to those generated by a laser spot source. The IOMZ interferometer-based laser

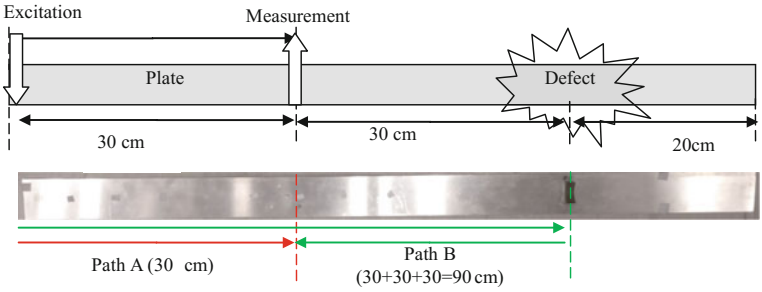


Fig. 4 The tested plate, its length and the location of defect (top diagram) and the GW propagating paths from laser emission point to the laser measurement point (path A) and to the defect location (path B) (bottom diagram)

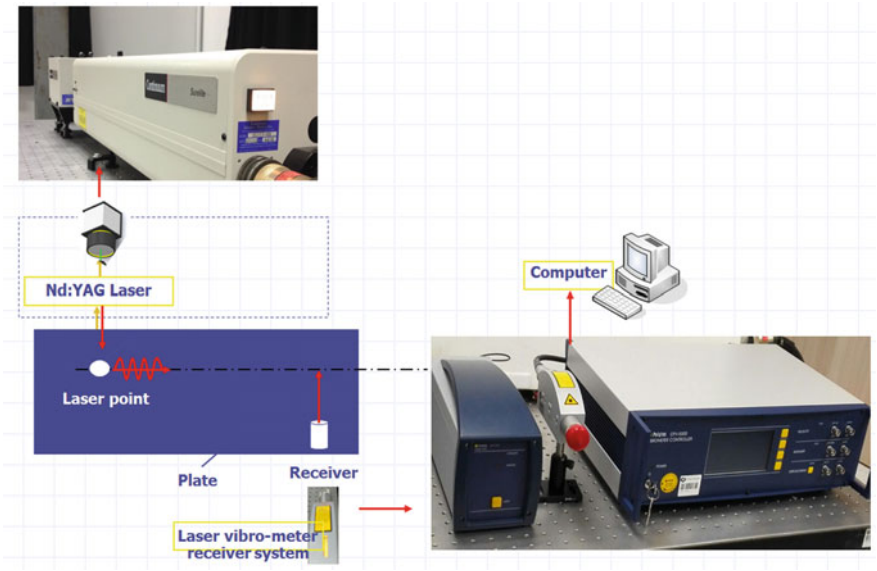


Fig. 5 Experimental set-up of the laser-based GW system for emitting (Nd:YAG pulsed laser, top-left diagram) and receiving by 1D vibrometer (bottom right diagram)

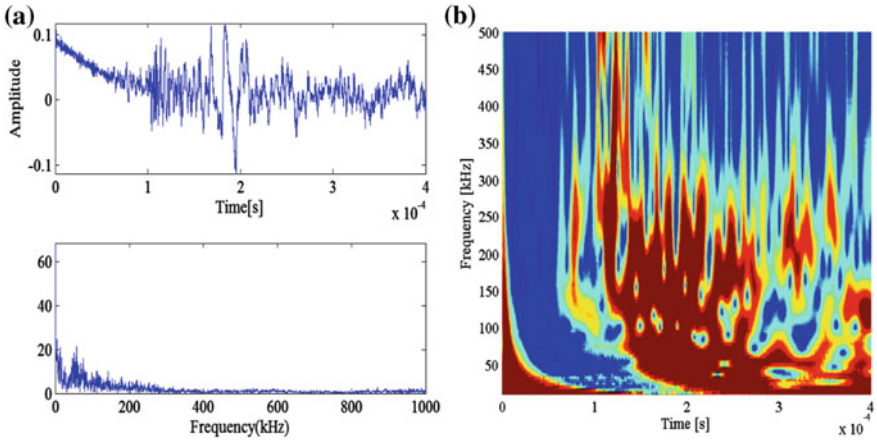


Fig. 6 GW signal emitted as a laser point source and then received from **a** a defective plate in its temporal waveform (top diagram) and its frequency spectrum (bottom diagram) and **b** its time-frequency in 2.5D signal after reconstructed by CWT

excitation system was also further applied to the defective plate as shown in Fig. 7b. This time, the received GW signal reflected by the defect can be observed due to the effect of minimizing noise and disturbance by the narrowed bandwidth of GW signal emitted by the IOMZ. From the time interval of the occurrence of defect

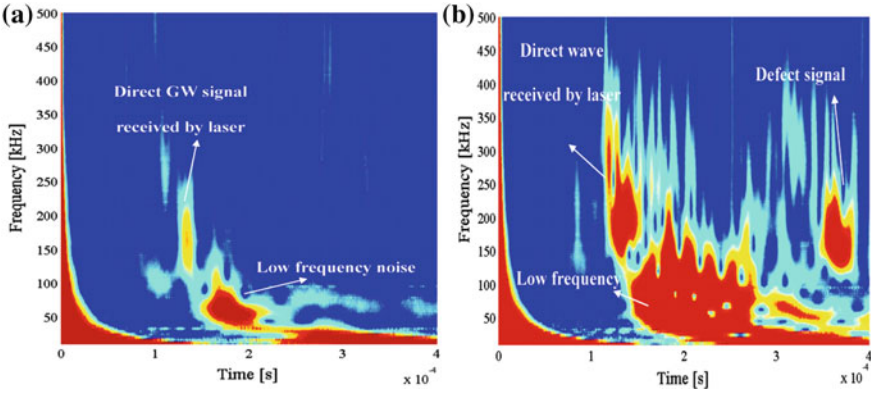


Fig. 7 GW signals received from **a** a normal plate after reconstructed by CWT and **b** a defective plate after reconstructed by CWT

signal in Fig. 7b, and with the known GW velocity of the A_0 mode, one can determine the starting and end time of the defective signal. Hence, the length and the location of the defect can be determined accordingly.

4 Conclusion

The research work on the design of a novel line-array laser source for producing a non-contact GW transducer to inspect the integrity of plates is reported here. The concept and method to design and implement such line-array laser source by using the novel IOMZ have been discussed in details. By using the conventional laser point source to emit the GW signal into the inspected plate, the true defect related signal can hardly be seen from the temporal and frequency plots of the collected GW signal. Even when the collected signal was processed by a popular signal processing method, the CWT, the defect's information could not be revealed from the time-frequency plot due to the existence of high noise and disturbances. The main reason is the broad bandwidth characteristic embedded in the emitted laser point source.

To overcome this deficiency, an innovative approach that totally used optical lens to create the line array like that generated by conventional GW's PZT transducers was invented. With the help of IOMZ that can convert the laser point source into line array source with the capability of selecting a proper GW frequency range by adjusting the width of each laser line, both the location and the axial length of the defect can be clearly revealed in the CWT time-frequency plot. Hence, the feasibility and capability of the innovative design of laser-based IOMZ for emitting the desired GW mode have been tested and verified respectively.

Acknowledgements 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_11201315).

References

1. Albiruni F, Cho Y, Lee J-H, Ahn B-Y (2012) Non-contact guided waves tomographic imaging of plate-like structures using a probabilistic algorithm. Mater Trans JIM 53(2):330
2. Bernstein JR, Spicer JB (2000) Line source representation for laser-generated ultrasound in aluminum. J Acoust Soc Am 107(3):1352–1357
3. Fokine M, Nilsson L-E, Claesson A, Berlemont D, Kjellberg L, Krummenacher L, Margulis W (2002) Integrated fiber Mach-Zehnder interferometer for electro-optic switching. Opt Lett 27(18):1643–1645
4. Fomitchov PA, Kromine AK, Krishnaswamy S, Achenbach JD (2000) Sagnac-type fiber-optic array sensor for detection of bulk ultrasonic waves. IEEE Trans Ultrason Ferroelectr Freq Control 47(3):584–590
5. Kim H, Jhang K, Shin M, Kim J (2006) A noncontact NDE method using a laser generated focused-Lamb wave with enhanced defect-detection ability and spatial resolution. NDT E Int 39(4):312–319
6. Kim HM, Lee TH, Jhang KY (2006) Non-contact single-mode guided wave technique by the combination of wavelength-matched laser generation and angle-matched leak wave detection. Paper presented at the Key Engineering Materials
7. Lee J-H, Lee S-J (2009) Application of laser-generated guided wave for evaluation of corrosion in carbon steel pipe. NDT E Int 42(3):222–227
8. Lee TH, Choi IH, Jhang KY (2008) Single-mode guided wave technique using ring-arrayed laser beam for thin-tube inspection. NDT E Int 41(8):632–637
9. Scruby C, Dewhurst R, Hutchins D, Palmer S (1980) Quantitative studies of thermally generated elastic waves in laser-irradiated metals. J Appl Phys 51(12):6210–6216
10. Sohn Y, Krishnaswamy S (2004) Interaction of a scanning laser-generated ultrasonic line source with a surface-breaking flaw. J Acoust Soc Am 115(1):172–181

Joint Optimization of Preventive Maintenance and Spare Parts Logistics for Multi-echelon Geographically Dispersed Systems

Keren Wang and Dragan Djurdjanovic

Abstract Preventive maintenance scheduling and optimization of logistic operations in geographically dispersed systems is an intricate decision-making problem, as maintenance activities have close interactions with spare parts logistics. Planning of spare parts logistics aims to get the right spare parts to the right places at right times for the necessary maintenance activities to take place, while keeping minimal requirements on maintenance resources and maintaining high utilization of assets. This paper considers such a problem by concurrently pursuing preventive maintenance scheduling and spare parts inventory planning problem for a set of geographically dispersed assets, each consisting of multiple degrading components. This paper also shows that the assumption of a multi-echelon maintenance network brings opportunities to improve the system performance through sharing inventories between maintenance facilities. Technically, the decision-making problem is modelled as a stochastic optimization problem and a novel simulation-based metaheuristic is proposed to solve it to obtain the integrated decision-making policy. Through numerical examples, the proposed integrated policy is shown to effectively reduce the operating cost by capturing the trade-offs between asset utilization, maintenance costs and consumption of maintenance resources.

1 Introduction

Preventive maintenance (PM) scheduling in a large and distributed system of degrading assets is a challenging decision-making problem because of inherent interactions between maintenance decisions and logistics operations. As PM

K. Wang (✉)
Program of Operations Research and Industrial Engineering,
University of Texas at Austin, Austin, USA
e-mail: wangkeren@utexas.edu

D. Djurdjanovic
Department of Mechanical Engineering, University of Texas at Austin, Austin, USA
e-mail: dragand@me.utexas.edu

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_63

operations are aimed at effectively restoring equipment reliability and reducing downtime costs by replacing degraded parts before they actually fail, getting the right amounts of spare parts available in the right places at the right time for maintenance operations is of paramount importance for success of those operations. Therefore, the interconnections and spare parts inventory levels in maintenance facilities should be considered simultaneously with maintenance schedules. This is especially important for large networks of geographically dispersed assets, such as oil/gas extraction companies or airlines.

Two general types of maintenance operations can be seen in practice: reactive maintenance (RM), which occurs after an asset starts behaving in an unacceptable manner, and preventive maintenance (PM), which is performed on an asset before unacceptable behavior occurs. Furthermore, PM policies can be classified as reliability-based maintenance (RBM), where maintenance is performed at certain age or usage interval of an asset, and condition-based maintenance (CBM), where maintenance is performed based on monitoring the actual condition of a system. Though, CBM decisions are more dynamic and efficient [8], the requirements on appropriate sensors and building of appropriate condition models make CBM both costly and challenging. Hence, RBM still dominates the PM practice.

According to a recent review on joint maintenance and inventory optimization systems [13], the early works on jointly PM scheduling and service parts logistics planning can be differentiated based on the underlying PM strategies into those that use age/usage-based PM policies [5, 7] and those that use block/period-based PM policies [1]. From the maintenance point of view, the complexity in asset structure is generally ignored in the aforementioned works, however, several recent studies go beyond the simple single-part asset structure and consider a serial-connected multi-part asset structure [6], *k*-out-of-*n* asset structure [2] and flexible-connect multi-part asset structure [9]. With focus on the degradation process and maintenance operations, these works inevitably assume simple logistic structure for service part management, with spare parts either being stocked locally, or being provided from a single source.

Other researchers seek to consider the joint PM and inventory optimization problems in a more complicated logistic environment, including human resources planning [10], equipment delivery decisions [14] and etc. Chen is the first, and to our best knowledge, the only work that conducted a study on a multi-echelon network [3], where there exists a distributor, multiple users and suppliers. These works, though enhancing the knowledge from the logistic side, oversimplified the maintenance decision-making process through assuming a simple asset structure.

To reflect the fact that real-world problems are complicated from both the maintenance and logistics sides, this paper proposed an integrated decision-making policy that jointly optimizes PM intervals and spare parts inventory levels in a multi-echelon network of geographically dispersed multi-part assets and multiple maintenance centers serving those assets with spare parts needed for their maintenance. Through modelling the decision-making process as a stochastic optimization problem and solving it via a discrete-event simulation-based metaheuristic approach, the proposed framework to obtain the integrated policies is flexible

enough to accommodate other more complicated logistic network and asset structures, as well as more elaborate system operations and cost functions.

The rest of the paper is organized as follows. In Sect. 2, the problem is described and the integrated decision-making approach is introduced. In Sect. 3, the proposed integrated decision-making policies are evaluated in a simulated environment. Section 4 provides conclusions of the research and outlines several possible avenues for future work.

2 Modelling Methodology

2.1 Problem Statement

The following terminology will be used in the remainder of this paper:

- The term *spare part* is referred to a type of maintenance resource stored in the maintenance facilities.
- The term *working part* is used to refer to a basic unit of an asset.
- The term *asset* is used to refer to a machine that can be operated independently to generate revenue. An asset is assumed to have multi-part structure and can be operated properly only if all its parts behave properly.
- A *maintenance center* (*MC*) fulfills maintenance orders from nearby assets by shipping new spare parts to their operating sites. Maintenance centers have finite inventory levels of spare parts and any maintenance order that cannot be immediately fulfilled is back-ordered to the central warehouse.
- A *central warehouse* replenishes spare parts for maintenance centers following a one-for-one (S-1, S) replenishment policy [12]. Also, the central warehouse can provide new spare parts directly to the assets as emergency orders in maintenance events. The central warehouse is the primary source of all new spare parts and is assumed to have infinite spare parts inventory levels.

As illustrated in Fig. 1, the topology of the system considered in this work is a multi-echelon spare parts logistic network with three levels of facilities. The maintenance centers (Level 2) provide service to nearby degrading assets (Level 1) through shipping the spare parts needed for maintenance, and get replenishment from the central warehouse (Level 3). In this paper, the central warehouse is connected to n maintenance centers, labeled $MC_1, MC_2, \ldots, MC_n$, and every maintenance center MC_i is further connected to a fleet of a_i nearby assets, $A_{i,1}, A_{i,2}, \ldots, A_{i,a_i}$. Due to the geographic dispersion of the degrading assets and maintenance facilities, the delivery delays of the necessary spare parts in maintenance events are considered and assumed to follow the symmetric triangular distributions.

The assets are assumed to have multi-part structure. For example, the asset $A_{i,j}$ is made up of $c_{i,j}$ serially connected parts, labeled $P_{i,j,1}, P_{i,j,2}, \ldots, P_{i,j,c_{i,j}}$, each of which

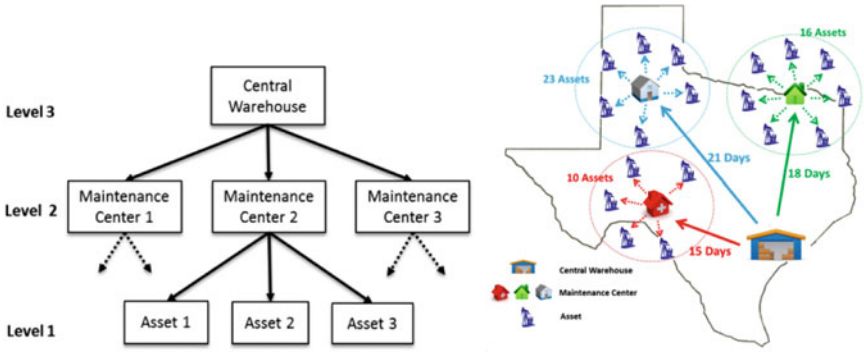


Fig. 1 Spare parts logistic network: multi-echelon structure (left) and geographic dispersion of the system (right)

degrades independently with its usage. The degradation dynamics of a part is described by a reliability function approximating the distribution of the part's usage to failure. The reliability function of a part is *part-and-asset* specific, and is estimated through analysis of the long-run statistical behavior of the part's degradation process on the asset, or parts of the same type on other similar assets.

A so-called *replacement maintenance policy* is assumed, that is to say, both PM and RM are assumed to consist of a new spare part replacing the broken or severely degraded working part on the asset. A complete maintenance order involves the following three steps: (1) Initialization: order a new spare part to replace the target working part during a planned PM or an unplanned RM; (2) Delivery: ship the ordered spare part from the connected maintenance center to the asset as a *normal order* or from the central warehouse as an *emergency order* (when no appropriate spare parts available in the connected MC); (3) Execution: replace the target working part on the asset with the delivered spare part, resulting in a maintenance intervention (repair time). Finally, we always assume a perfect maintenance operation.

From the side of maintenance, since a *usage-based PM replacement policy* is used, then PM triggering usage level $x_{i,j,k}$ is set for each part $P_{i,j,k}$, indicating the part's critical usage level at which a PM operation is initiated, and is assumed to take value in a discrete value set, $X_{i,j,k}$. From the side of logistic network, let $SP_1, SP_2, \ldots, SP_H$ denote all types of spare parts in this system, and a *one-for-one replenishment policy* for spare parts in maintenance centers is used. Therefore, the decision-variables, $\{y_{i,h}\}_{1 \le i \le n, 1 \le h \le H}$, indicate the critical levels of spare part $SP_h, h = 1, 2, \ldots, H$, to restore its current inventory level to the target level.

2.2 Inventory-Sharing Option

So far, if the corresponding maintenance center runs out of the necessary spare part during a maintenance event, a spare part is shipped from the central warehouse as an emergency order, leading to a much longer asset downtime than a normal order. A more common situation in reality is that, when an unplanned failure happens and the corresponding maintenance center is out of the necessary spare part, that spare part could be obtained faster from some other nearby MCs. This way of spare part delivery will be referred to in this paper as a "*complementary order*". From the logistic point of view, the implementation of complementary orders leads to a strategy of sharing inventory between maintenance centers in serving the assets, thus enabling avoidance of backorders to the central warehouse.

Following this strategy, there are multiple sources for an asset to obtain spare parts for a maintenance action—central warehouse and several nearby maintenance centers. The priority of these sources is determined by evaluating the costs associated with providing a RM service from that source, with those costs involving the cost to order RM and the expected penalty for the asset downtime.

2.3 Integrated MC-Dependent Decision-Making Policy

For a system considering the inventory-sharing option, we will seek an integrated decision-making policy for spare parts inventory levels $\{y_{i,h}\}_{1\le i\le n,1\le h\le H}$ and PM triggering usage levels $\{x_{i,j,k}\}_{1\le i\le n,1\le j\le a_i,1\le k\le c_{i,j}}$ that minimize the cost function representing the expected unit-time operating cost. This policy is referred as the *integrated MC-dependent decision-making policy* in this paper, and is obtained through solving the following stochastic optimization problem,

$$\begin{aligned}\underset{\substack{\{x_{i,j,k}\in X_{i,j,k}\}_{1\le i\le n,\,1\le j\le a_i,1\le k\le c_{i,j}}\\ \{y_{i,h}\in\{0,1,2,3,\ldots\}\}_{1\le i\le n,1\le h\le H}}}{\text{Minimize}} \quad & \frac{1}{T}\mathbb{E}\left\{\sum_{i,h}\left(h_{i,h}V^H_{i,h}+s_{i,h}V^S_{i,h}\right)\right.\\ &+\sum_{i,j,k}\left(r_{i,j,k}R_{i,j,k}+m_{i,j,k}M_{i,j,k}+e_{i,j,k}E_{i,j,k}\right)\\ &\left.+\sum_{i,j,k}\sum_{i'\neq i}\left(r^{i'}_{i,j,k}R^{i'}_{i,j,k}+m^{i'}_{i,j,k}M^{i'}_{i,j,k}\right)+\sum_{i,j}l_{i,j}L_{i,j}\right\}\end{aligned} \tag{1}$$

with the notations being described in Table 1.

As we can see that, three types of costs considered in this model: the inventory costs (holding/replenishment), the penalties for asset downtimes, and the maintenance costs (normal PM/RM, complementary PM/RM, emergency RM). To be

Table 1 Notations used in the optimization formulation (1)

Category	Symbol	Description
General	i, i'	Indices for maintenance center
	j, k, h	Indices for asset (j), working part (k), spare part type (h)
	T	Time horizon
Inventory	$V_{i,h}^{H}$	Inventory holding cost per unit time for the spare part S_h at MC_i
	$V_{i,h}^{S}$	Replenishment delivery cost per unit time for the spare part S_h to MC_i
	$h_{i,h}$	Cumulative inventory holding time of the spare part S_h at MC_i
	$s_{i,h}$	Cumulative replenishment delivery time of the spare part S_h to MC_i
Downtime penalty	$L_{i,j}$	Penalty cost per unit downtime of the asset $A_{i,j}$
	$l_{i,j}$	Cumulative downtime of the asset $A_{i,j}$
Normal PM/RM	$R_{i,j,k}$	Cost per RM, including all one-time costs to have RM on the part $P_{i,j,k}$
	$M_{i,j,k}$	Cost per PM, including all one-time costs to have PM on the part $P_{i,j,k}$
	$r_{i,j,k}$	Cumulative number of RM orders for the part $P_{i,j,k}$
	$m_{i,j,k}$	Cumulative number of PM orders for the part $P_{i,j,k}$
Emergency RM	$E_{i,j,k}$	Additional charge of an emergency RM on the part $P_{i,j,k}$
	$e_{i,j,k}$	Cumulative number of emergency RM orders for the part $P_{i,j,k}$
Complementary PM/RM	$R_{i,j,k}^{i'}$	Cost per complementary RM for the part $P_{i,j,k}$ serviced by $MC_{i'}$
	$M_{i,j,k}^{i'}$	Cost per complementary PM for the part $P_{i,j,k}$ serviced by $MC_{i'}$
	$r_{i,j,k}^{i'}$	Cumulative number of complementary RM for $P_{i,j,k}$ serviced by $MC_{i'}$
	$m_{i,j,k}^{i'}$	Cumulative number of complementary PM for $P_{i,j,k}$ serviced by $MC_{i'}$

noted that, different companies and even different parts of the same company will operate with different cost function and parameters, this optimization model can be easily adapted to reflect specificities of the cost model.

The expectation operator is applied due to the stochastic effects induced by randomness in the reliability of working parts and the time-related parameters (all repair times and shipping times). For each candidate solution, these random effects were captured by discrete-event simulations in the solution process.

2.4 Integrated MC-Independent Decision-Making Policy

Similarly, a so-called *integrated MC-independent decision-making policy* will be developed for a system without considering the inventory-sharing option. As sharing spare parts between MCs is no longer allowed, the cost for complementary PM/RM orders will not be involved in the cost function, and activities in different MCs are independent, which implies that minimizing the operating costs for each MC separately yields the same decisions as minimizing the cost for the entire system. Therefore, the stochastic optimization problem (Eq. 1) can be decomposed to n independent optimization sub-problems given in formulation (Eq. 2), each yielding the integrated decision of PM triggers and spare parts inventories for a MC.

$$\begin{aligned} &\underset{\{x_{i,j,k} \in X_{i,j,k}\}_{1 \le j \le a_i, 1 \le k \le c_{i,j}}}{\text{Minimize}} \quad \frac{1}{T}\mathbb{E}\left\{\sum_h \left(h_{i,h}V_{i,h}^H + s_{i,h}V_{i,h}^S\right)\right. \\ &\{y_{i,h} \in \{0,1,2,3,\ldots\}\}_{1 \le h \le H} \\ &\quad + \sum_{j,k}\left(r_{i,j,k}R_{i,j,k} + m_{i,j,k}M_{i,j,k} + e_{i,j,k}E_{i,j,k}\right) \\ &\quad \left. + \sum_j l_{i,j}L_{i,j}\right\} \quad \textit{for} \;\; i = 1,2,\ldots,n \end{aligned} \tag{2}$$

2.5 Simulation-Based Optimization Algorithm

In this section, we will describe a simulation-based metaheuristic optimization procedure that pursues the proposed integrated MC-dependent decision-making policy through solving the optimization problem (1) formulated for a system considering inventory-sharing option[1].

Simulation-based optimization is shown to be a powerful paradigm for decision-making in jointly PM and spare parts inventory systems [3, 6]. Consequently, discrete-event simulations are utilized in this paper to estimate the expected unit-time operating cost of the system under a candidate solution, denoted by (X, Y), where X is a decision-variable vector containing all PM triggering usage levels triggering PM orders, and Y contains all spare parts inventory levels triggering replenishment orders from the central warehouse to MCs. The cost function, $F(X, Y)$, is estimated via multiple replications of discrete-event simulations, and is

[1]By treating each sub-problem in formulation (2) as a special case of the optimization problem (1), this proposed optimization approach is also useful in obtaining the integrated MC-independent policy.

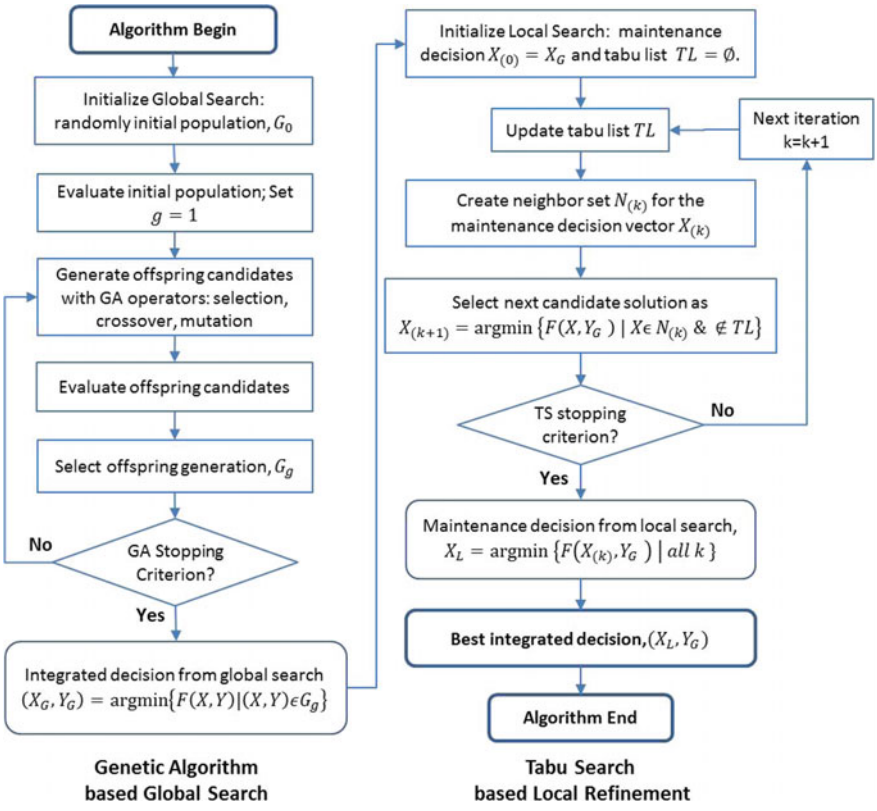


Fig. 2 Flow chart of the hybrid GA-TS based metaheuristic

then fed back into the search heuristic to guide the movements towards improved candidate decisions.

To search for the solution to the optimization problem (1), a hybrid approach is proposed, that commences using a genetic algorithm (GA) [11] to conduct a fast global search and the solution generated by it is further improved through a Tabu search (TS) [4] based optimization refinement. The flow chart in Fig. 2 highlights the proposed hybrid algorithm.

3 Simulated Example

We evaluated the newly proposed integrated decision-making policies (MC-dependent and MC-independent) in a simulated system. The two integrated policies will be benchmarked to a fragmented decision-making policy that is developed following the industrial common practice that the maintenance schedule

is usually determined before the decisions on the spare parts inventory levels. More details of this benchmark policy can be found in [15].

The simulated system has a central warehouse and three maintenance centers that respectively provide maintenance service to 10, 16 and 23 geographically dispersed assets. These 49 assets consist of all together 126 parts, belonging to 5 different types. The set of decision variables consists of 15 logistic decision variables and 126 maintenance decision variables.

With the planning horizon as 36,500 time units, the operating costs and other system statistics will be evaluated as averages over 50 replications. The integrated decision-making policies are obtained by the hybrid GA-TS based metaheuristic described in Sect. 2.5. In terms of computational costs, the implementation on an ordinary PC (Intel Core i5 CPU, 16 GB RAM, 64-bit Win 7) took less than 15 h to obtain an integrated policy, it would be feasible to achieve daily updates on the integrated decisions for systems of similar scale.

In the simulated system, the operating costs under the three decision-making policies are compared in Fig. 3. We can observe that both integrated policies outperform the fragmented policy with much lower total operating costs. Additionally, inventory sharing attached to the integrated MC-dependent policy enables a further 2.9% operating cost reduction from the MC-independent policy. Among the three types of operating costs, the reductions in inventory-related costs explain the most of the operating cost changes, implying a more efficient usage of maintenance resources.

Table 2 offers a more detailed comparison between the three policies. Compared to the fragmented policy, we can see that both integrated policies require less spare parts inventories but expect a more aggressive PM schedule, implying the fact that the joint decision-making process led to savings in maintenance resources and had the tendency to replace costly RMs with relatively cheaper PMs.

Also, it is visible that the fewest spare parts are required within the integrated MC-dependent policy, which is a clear indication of a more efficient usage of spare parts through sharing inventories among all maintenance centers. Furthermore, the

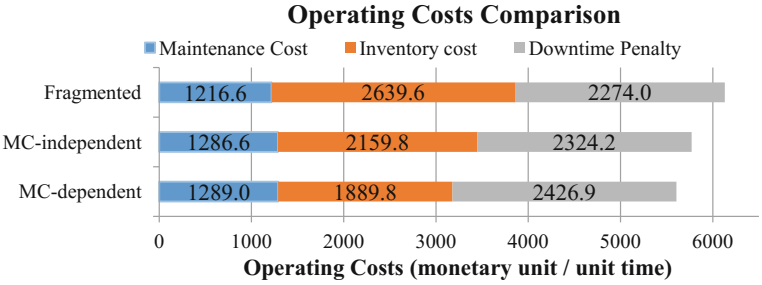


Fig. 3 Operating costs comparison between the fragmented and integrated (MC-dependent, MC-independent) policies

Table 2 Comparison of performance statistics between three decision-making policies

Policy name	Fragmented	Integrated MC-independent	Integrated MC-dependent
Expected overall operating cost (monetary unit/unit time)	6130.2	5770.6	5605.7
Sum of spare part inventory levels	88	72	63
Average PM triggering usage level	55.24	48.85	49.76
Total number of emergency RM	589.9	377.3	96.8

most expensive and time-consuming emergency RM orders serviced by the central warehouse are almost completely replaced by more cost effective RM orders serviced by nearby maintenance centers under the inventory-sharing option.

4 Conclusion

In this paper, a decision-making method is proposed for concurrent spare parts inventory planning and preventive maintenance scheduling in a system of geographically dispersed assets and maintenance facilities that serve those assets. The newly proposed integrated decision-making method considers a usage-based PM replacement policy on the geographically dispersed degrading assets and a one-for-one replenishment policy for the maintenance centers that provide those assets with the necessary spare parts. This optimization problem was solved via a simulation-based metaheuristic approach.

The integrated decisions-making policies introduced in this paper were implemented in a simulated environment and benchmarked against the traditionally used fragmented decision-making policy in which maintenance and logistic decisions are made sequentially. The results illustrate that the newly proposed integrated decision-making policies outperforms the fragmented decision-making policy by enabling lower overall system operating costs. Furthermore, both the joint maintenance/logistics decision-making process and the inventory-sharing option led to avoidance of excessive stocking of spare parts in maintenance centers and showed the preference on relatively cost-efficient maintenance operations.

Several research extensions of this work can be noted. Firstly, the methodology presented in the paper can be extended to consider more elaborate system operations, such as imperfect maintenance and multiple delivery options. Secondly, the robustness of integrated decision-making policy can be improved to uncertainties in the model. Finally, the labor resource planning also deserves further studies.

References

1. Acharya D, Nagabhushanam G, Alam S (1986) Jointly optimal block-replacement and spare provisioning policy. IEEE Trans Reliab 35:447–451
2. Bjarnason ETS, Taghipour S (2015) Periodic inspection frequency and inventory policies for a k-out-of-n system. IIE Trans 48(7):638–650
3. Chen M-C, Hsu C-M, Chen S-W (2006) Optimizing joint maintenance and stock provisioning policy for a multi-echelon spare part logistics network. J Chin Inst Ind Eng 23:289–302
4. Glover F, Laguna M (2013) Tabu search. Springer
5. Hu R, Yue C, Xie J (2008) Joint optimization of age replacement and spare ordering policy based on genetic algorithm. In: International conference on computational intelligence and security, CIS'08. IEEE, pp 156–161
6. Ilgin MA, Tunali S (2007) Joint optimization of spare parts inventory and maintenance policies using genetic algorithms. Int J Adv Manuf Technol 34:594–604
7. Kabir AZ, Al-Olayan AS (1996) A stocking policy for spare part provisioning under age based preventive replacement. Eur J Oper Res 90:171–181
8. Lee S, Djurdjanovic D, Ni J (2007) Optimal condition-based maintenance decision making for a cluster tool. In: TECHCON, Austin TX, 10–12 Sept
9. Lynch P, Adendorff K, Yadavalli V, Adetunji O (2013) Optimal spares and preventive maintenance frequencies for constrained industrial systems. Comput Ind Eng 65:378–387
10. Nguyen D, Bagajewicz M (2010) Optimization of preventive maintenance in chemical process plants. Ind Eng Chem Res 49:4329–4339
11. Reeves C (2003) Genetic algorithms. Springer
12. Svoronos A, Zipkin P (1991) Evaluation of one-for-one replenishment policies for multiechelon inventory systems. Manag Sci 37:68–83
13. Van Horenbeek A, Buré J, Cattrysse D, Pintelon L, Vansteenwegen P (2013) Joint maintenance and inventory optimization systems: a review. Int J Prod Econ 143:499–508
14. Zanjani MK, Nourelfath M (2014) Integrated spare parts logistics and operations planning for maintenance service providers. Int J Prod Econ 158:44–53
15. Wang K, Djurdjanovic, D (2017) Joint optimization of maintenance and spare parts logistics for a system of geographically distributed, multi-part assets. Working paper

Vane Pump Damage Detection via Analysis of Synchronously Averaged Vibration Signal

Wenyi Wang

Abstract Variable displacement vane pumps are gaining popularity in aerospace applications. Damage detection of this type of pump is traditionally carried out using measurements of pump performance parameters, such as the flow rate ripple which is difficult to measure directly. Vibration analysis is also employed to detect pump faults related to bearings and shafts, via spectral analysis of raw signals in most cases. Damage such as cavitation to the vanes and side plates of vane pumps is very common, which can be very challenging to detect using performance parameter and raw vibration spectral analysis. This paper presents an advanced technique of detecting such damage using features derived from synchronously averaged vibration signatures, which aims to isolate pump vibration signal from other dominating vibration sources. The technique is applied to the vibration data recorded by an on-board monitoring system on a modern military aircraft engine. The vibration data contained the failure event of a variable displacement vane pump attached to the accessory gearbox of the engine. The analysis results show that the technique can effectively detect the changes caused by the cavitation damage to the vanes. With appropriate threshold setting and further validation, this method can be readily implemented into real-time pump monitoring systems.

1 Introduction

Traditionally, pump condition monitoring focused on examining operation parameters such as the inlet and outlet pressures, flow rate, drive speed and power, and bearing temperature, etc. Often, advice on when a pump should be overhauled is based on increased internal clearance (e.g. doubled from the design value) and reduced pumping capacity (e.g. a drop of 4% or more), see [2]. However, the clearance-based guideline would require the dismantling of the pump and capacity-based advice would need head-flow tests [2] to identify such changes in

W. Wang (✉)
Aerospace Division, Defence Science and Technology Group, Melbourne, Australia
e-mail: wenyi.wang@dst.defence.gov.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_64

performance. Vibration and sound data analyses at selected frequency bands can be used as diagnostic means to detect faults due to unbalance, misalignment and resonant vibration excited by flow and bearing damages [1].

It is more and more popular in aerospace applications to use variable displacement vane pumps because less heat is generated by these pumps than the traditional gear pumps. Damage detection of this type of pumps is mostly carried out using measurements of pump performance parameters, such as the flow rate ripple which is difficult to measure directly [4, 6]. Vibration analysis is also employed to detect pump faults related to bearings and shafts, and spectral analysis of raw signals is normally used [3]. Cavitation damage to the vanes and side plates of vane pumps is very common, which could be very challenging to detect using performance parameter and raw vibration spectral analysis [2].

This paper presents an advanced technique of detecting vane damage caused by cavitation using features derived from synchronously averaged vibration signatures, which is focused on isolating pump vibration signal from other dominating vibration sources. From the averaged signal, two condition indicators (CI) are generated and trended over time. The two CI's are: (1) the energy ratio of the residual signal and the pump characteristic signal; (2) the scaled kurtosis in three different frequency bands of the change signal derived by the unified change detection approach [5]. The techniques are validated and compared using the vibration data recorded by an on-board monitoring system on a modern military aircraft engine. The vibration data contained the failure event of a variable displacement vane pump attached to the accessory gearbox of the engine. The analysis results show that the technique can effectively detect the changes caused by the cavitation damage to the vanes. This method could be readily implemented into real-time pump monitoring systems after appropriate threshold setting and further validation.

2 Vane Pump Failure and Vibration Data

During an accelerated test program of a military aircraft engine, a vane pump failed due to significant cavitation damage to all the vanes of the pump which was attached to the accessory gearbox of the engine. The engine on-line monitoring system did not detect the fault before failure occurred. The vibration data used for this paper were recorded at three stages of the test when the engine was running on full power. They were from (*a*) early stage—within the first 10% of the testing time; (*b*) late stage—between 80 and 90% of the testing time and (*c*) last stage—within the last 2% of the testing time of the accelerated test.

3 Data Analysis Methods

In this paper, two methods were employed to analyse the synchronously averaged vibration signal with respect to the pump drive shaft. The first method, where the novelty is for this paper, is to extract the pump characteristic signature, such as the vane pass frequency and harmonics, then to take the energy ratio between the residual signal and the pump characteristic signal. The second approach is to use the very first averaged signal as the reference and to work out the change signal of other signals under monitoring from the reference, which is the so-called unified change detection method originally proposed by the author in Wang et al. [5].

3.1 Energy of Pump Characteristic Signature

When a vane pump is in healthy condition, tight clearance exists between the vane tips and the pump casing (cutwater or throat). When each vane passes through the cutwater or throat area, it generates turbulent flow affecting vibration signature, hence the vane pass harmonics normally appear in the pump vibration spectrum. In faulty conditions, e.g. vane tip eroded by cavitation, the tip-throat or tip-cutwater clearance is increased and less turbulence will be generated; hence weaker vane pass harmonic content in the spectra is expected. On the other hand, synchronous vibration components may be affected by the development of faults in the vane pump. On this basis, we propose a condition index (CI) in this paper to track the deteriorating pump condition. Let $y(t)$ be the synchronously averaged vibration signal with the pump shaft, $y_{vp}(t)$ the harmonics of vane pass frequency and $y_r(t)$ the residual signal after removing the $y_{vp}(t)$ from $y(t)$, the mathematic expression of the condition index (CI_1) is given below:

$$\begin{aligned} y(t) &= y_{vp}(t) + y_r(t) \\ CI_1 &= \sigma[y_r(t)]/\sigma\left[y_{vp}(t)\right] \end{aligned} \tag{1}$$

where $\sigma[.]$ is the standard deviation (or root mean square—RMS) of the signal. For pumps in a complex mechanical system such as aircraft engines, some discrete spectral lines generated by other mechanical components in the system may need to be removed before calculating the CI using Eq. 1. An example of synchronous signal average (SSA) with the pump driving shaft is shown in Fig. 1. Figure 2 shows the spectrum of the SSA where the dominant spectral content is the vane pass harmonics, i.e. the integer multiples of 10 shaft orders. It is worthwhile to mention that the raw vibration spectrum was dominated by the blade pass harmonics of the engine and the gear mesh harmonics of the accessory gearbox, which are not shown here due to the limitations on paper length. Without synchronous averaging the vibration signal with respect to the pump's drive shaft, the pump

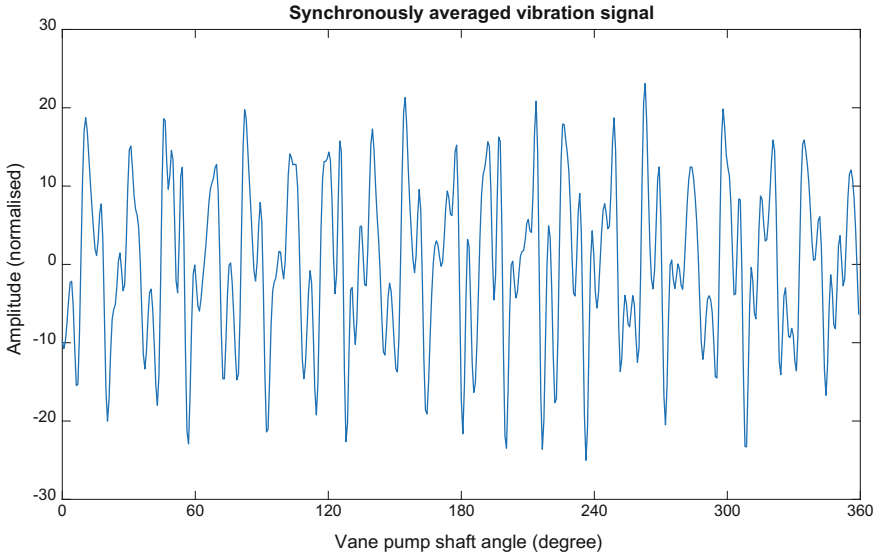


Fig. 1 Synchronous signal average (SSA) with respect to pump drive shaft

characteristic vibration components would be completely masked by these other dominate vibration sources.

The spectrum in Fig. 2 can be decomposed into the spectrum of the vane pass harmonics shown in Fig. 3 and the spectrum of the residual signal shown in Fig. 4.

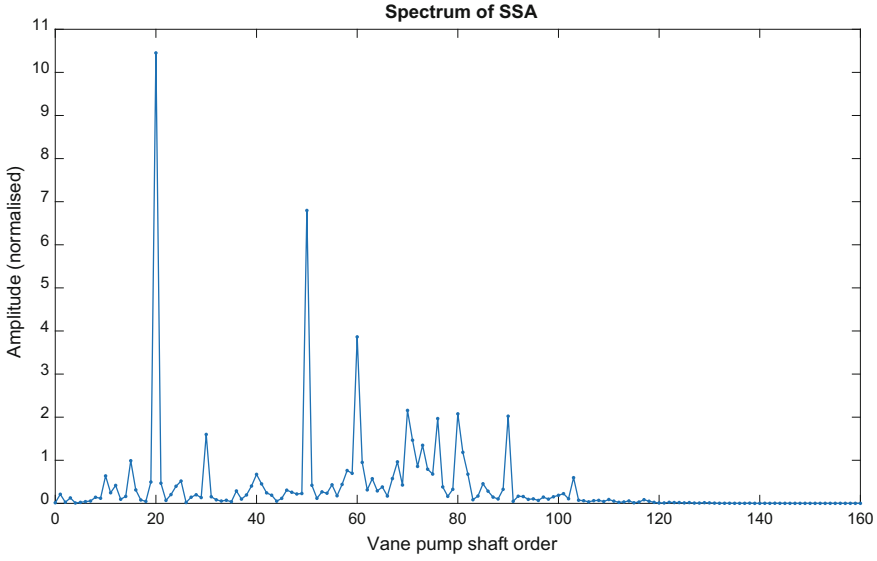


Fig. 2 Spectrum of synchronous signal average (SSA) shown in Fig. 1

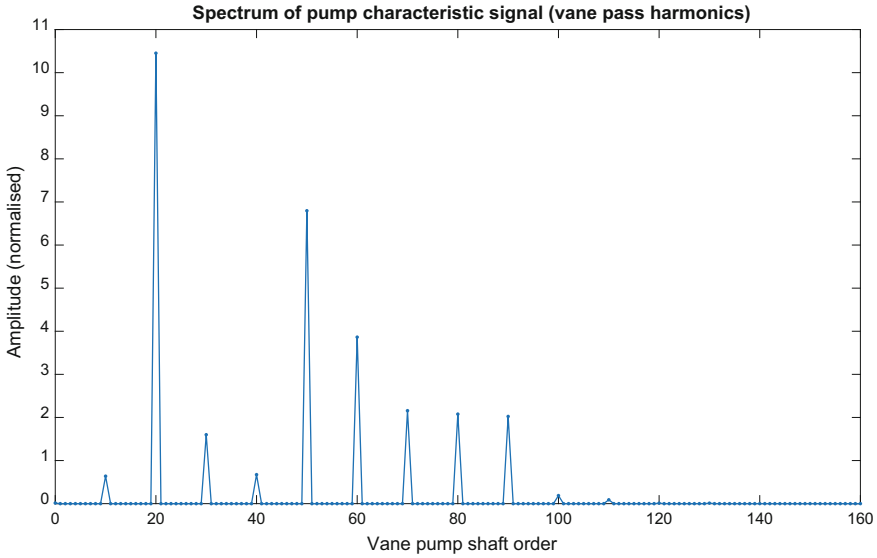


Fig. 3 Spectrum of pump characteristic signal (vane pass harmonics—VPH)

It is anticipated that, as the vane tips are damaged by cavitation, the energy of vane pass harmonics (or pump characteristic vibration) is reduced and the energy of residual signal is increased. Therefore, the energy ratio of the residual signal to the pump characteristic vibration signal will be on an increasing trend as the damage progresses.

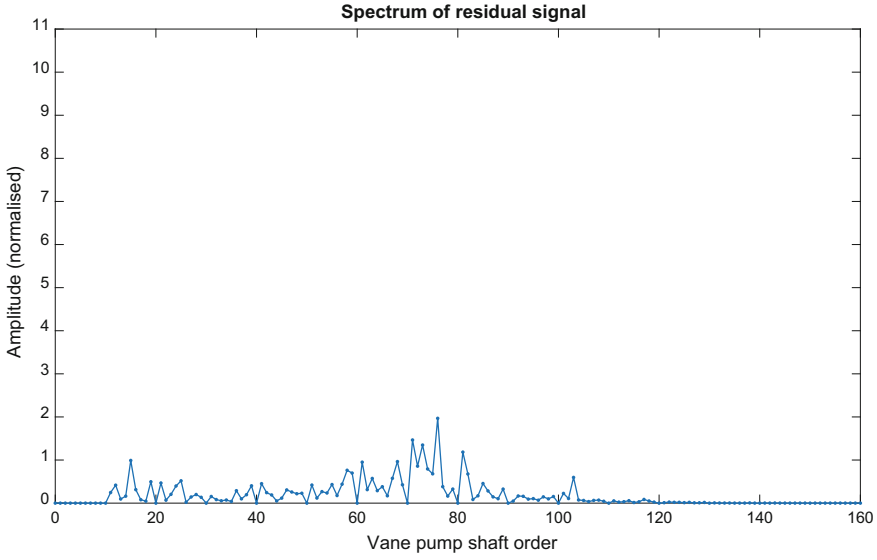


Fig. 4 Spectrum of residual signal after removing VPH signal

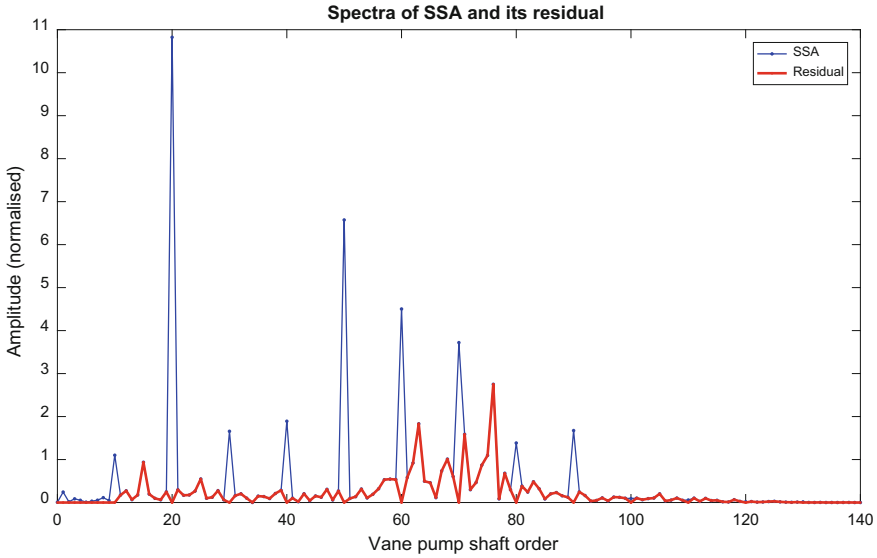


Fig. 5 Spectrum of an early stage averaged vibration signal (CI = 0.3423, residual RMS = 3.4836, VPH RMS = 10.1778)

Figures 5 and 6 show the vane pass signal and the residual signal at the early and later stages of the fault development, respectively, where the indication of condition change is evident. From the signals shown in Figs. 5 and 6, we can see that the CI has gone up by 88.2% (from normalised values of 0.3423 to 0.6442), i.e. residual RMS up by 33.8% whereas vane pass RMS down by 28.9%. As seen in the spectra, the change in the residual is not as visually obvious as that in vane pass harmonics.

A trending plot for CI_1 over time is shown in Fig. 7. This method appears to be very sensitive in detecting the late stage changes, as shown in the second step change in Fig. 7. For this pump failure, when the SSA is not used, the condition indices based on raw vibration signals, e.g. RMS and kurtosis and crest factor etc., are not able to detect the changes caused by the damage to the vane pump.

The threshold for this CI at high engine power should be lower than those at low engine power because the vane pass harmonics are inherently more pronounced (i.e. larger denominator) at high power settings. Assuming the CI value in a healthy condition is normally distributed and has a mean of 0.25 with standard deviation of 0.03, we can employ the 3-sigma principle to set the alert threshold, i.e. $0.25 + 3 \times 0.03 = 0.34$, and a 6-sigma level to set the alarm threshold at 0.43. By observation in Fig. 7, the late stage faulty condition has a mean CI of 0.35, and the last stage has a mean CI of 0.6, which far exceeds the presumed alarm threshold of 0.46. The vane pump actually failed very quickly after the last stage shown in Fig. 7. Sometimes, when the CI variability is changing with time, the alarm threshold could be set by the 3-sigma principle (assuming normally distributed CI)

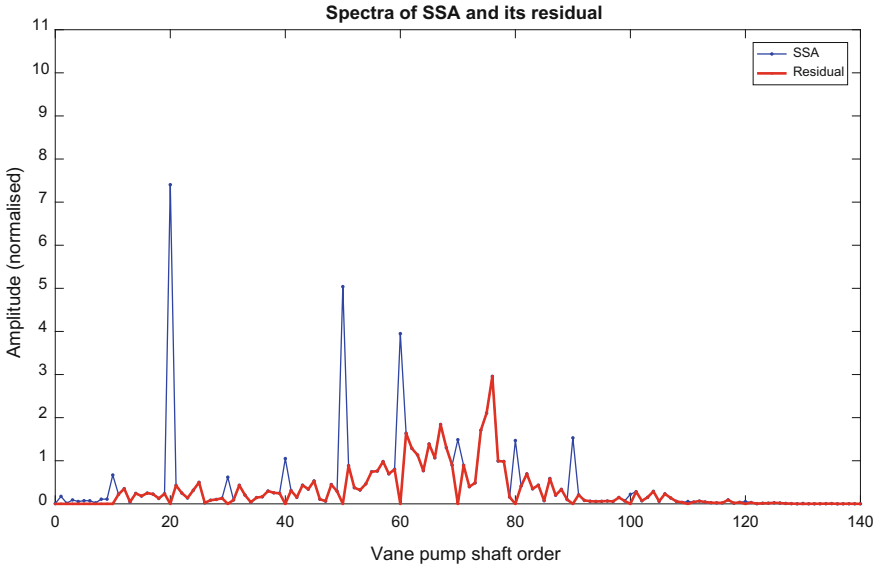


Fig. 6 Spectrum of a late stage averaged vibration signal (CI = 0.6442, residual RMS = 4.6613, VPH RMS = 7.2362)

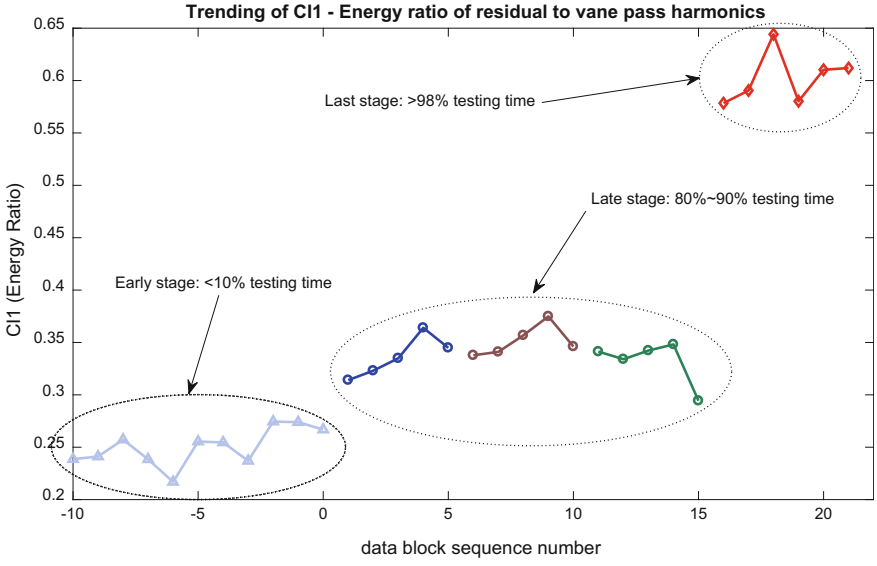


Fig. 7 Trending of condition index 1 (CI-1) over time

on the alert level, which produces a similar alarm threshold value to 0.46 mentioned above in the case of vane pump damage.

3.2 *Unified Change Detection*

A unified change detection method was recently proposed by the author [5] to detect and trend changes caused by various types of faults in rotating machinery, which is used here to compare with the approach proposed in this paper. A brief description of this method is given in this section. This method may have an advantage in detecting unexpected or new types of mechanical fault. The detection of changes due to machine faults often involves comparison of signals from the healthy-state to the faulty-state of the machine. However, a direct comparison in the time domain is often prohibited simply because these signals are in most cases not phase-aligned. The unified approach deals with the synchronously averaged vibration signals from a rotating component in the machine as it progresses from a healthy state to a faulty state. The healthy-state signal x is employed as a reference, and it is phase shifted by the phase difference to the future-state (healthy- or faulty-state) signals y. The phase shifted healthy-state signal x_s is then subtracted from future-state signals y to form the change signal, which is often referred to as baseline subtraction in structural health monitoring. It is expected that fault-induced changes will be captured by the change signal. Statistical measures can then be derived from the change signal as condition indicators, and trended over time for fault detection purposes. When the trending curve goes beyond a prescribed threshold, the change signals can be carefully examined to diagnose the type of fault.

The theoretical foundation of the proposed technique is based on extracting the differences between the synchronously averaged signals acquired under the changing health conditions after aligning their phases. The alignment enables a direct comparison between the amplitudes of the synchronously averaged signals and provides a trustworthy evaluation using a common reference datum. It has been shown [5] that the alignment of the phases can be performed in the frequency domain, as summarised in Eq. 2.

$$\delta_{xy}(t) = \int_{-\infty}^{\infty} [A_Y(f) - A_X(f)] \cdot e^{j \cdot \Phi_Y(f)} \cdot e^{j2\pi f t} \mathrm{d}f \tag{2}$$

Therefore, Wang et al. [5] defined the spectrum of the change signals as having an amplitude spectrum of $[A_Y(f)–A_X(f)]$, which is the amplitude spectrum of the difference signal, and a phase spectrum of $\Phi_Y(f)$, which is the phase spectrum of $y(t)$. The time domain change signal is then obtained by an inverse Fourier transform as shown in Eq. 2.

For the vane pump vibration data described in the previous section, we employed 36 tri-axial vibration data files used for the analysis using the unified change detection method where the first one in the early stage was used as the reference. As a comparison to Fig. 7, the Fig. 16 in Wang et al. [5] is re-produced and shown here as Fig. 8. It shows the trending curves of scaled kurtosis CI_2 (in three bands using the horizontal axial (the most sensitive direction) vibration data. The scaled kurtosis was defined by Wang et al. [5] as the product of kurtosis and the energy ratio of the change signal to the reference signal.

In Fig. 8, along the horizontal axis there are 35 columns of CI points; the first 11 files were from the early stage of testing; the following 18 files were from the late stage and the last 6 files were from the last stage of testing. The cross-over frequency for the low and high bands was selected at just above the 6th harmonic of the vane pass frequency. We can see in Fig. 8 that CI_2 for the full band (red) and low band (green) show a prominent step change across the three stages of testing. The high band (blue) CI shows some indication of change but it is not as prominent as the other two bands. This is because the signal changes caused by the vane damage are mostly likely located at the vane pass frequency and its lower harmonics, which was observable in Figs. 5 and 6 as the negative change or decreased vane pass harmonics as damage progressed. Obviously, the changes detected by the unified approach can give sufficient lead time to the failure of the vane pump. The pump actually failed on the very next run after the last data file was recorded.

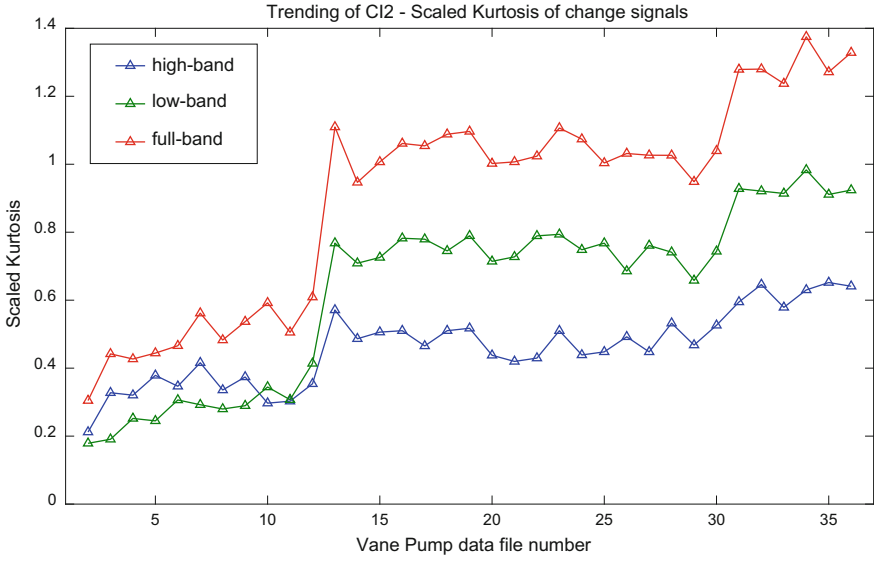


Fig. 8 Detection of changes caused by cavitation damage using unified change detection approach by trending the scaled-kurtosis (CI-2) over time with cross-over frequency of 65th shaft order. This is a reproduction of Fig. 16 in [5]

4 Discussion and Conclusion

We employed two methods, one proposed in this paper and the other proposed by the author in Wang et al. [5], for analysing the SSA vibration to detect changes caused by cavitation damages in a vane pump. Results show that both methods are capable of detecting the changes. The first method seems to be more sensitive in detecting the late stage changes shown as the second step change in Fig. 7, whereas the second method in full band appears very good at reflecting the early stage changes shown as the first step change in Fig. 8. Results presented here have indicated that the synchronous signal average (SSA) of vibration signals is an effective tool to extract the pump vibration signal from the more dominant engine background vibration, such as the gear mesh vibration and blade pass tonal vibration. It is very important to perform SSA prior to further signal processing.

On the basis of the analysis results, we can conclude that, by combining the two methods of analysing the SSA, changes caused by cavitation damage to vane pumps can be effectively detected. Late stage changes are best detected by the energy ratio between the residual signal and the pump characteristic signal. Early stage changes may be more efficiently detected in the full band scaled kurtosis of the change signal obtained via the unified change detection approach.

References

1. Badu GS, Das VC (2013) Condition monitoring and vibration analysis of boiler feed pump. Int J Sci Res Publ 3(6):1–7
2. Beebe RS (2004) Predictive maintenance of pumps using condition monitoring. Elsevier Science & Technology Books
3. Hancock KM, Zhang Q (2006) A hybrid approach to hydraulic vane pump condition monitoring and fault detection. Trans Am Soc Agric Biol Eng (ASABE) 49(4):1203–1211
4. Vijaykumar, S, Prabakkaran, Natesan R (2014) A condition monitoring system for vane pump using LArVa. Int J Sci Eng Res 5(5):904–907
5. Wang W, Forrester BD, Frith PC (2016) A unified approach to detecting and trending changes caused by mechanical faults in rotating machinery. J Struct Health Monit 15(2):204–222
6. Yang M, Edge KA, Johnston DN (2008) Condition monitoring and fault diagnosis for vane pumps using flow ripple measurement. In: Proceedings of BATH/ASME symposium on fluid power and motion control (FPMC), Bath, United Kingdom, Sept 2008

Novel Non-destructive Technique of Internal Deterioration in Concrete Deck with Elastic Wave Approaches

Kazuo Watabe, Hidefumi Takamine, Takahiro Nishida and Tomoki Shiotani

Abstract An innovative technique to visualize internal defects in concrete bridge decks has been developed. To achieve visualization of internal damage, acoustic emissions (AE) using the two following analysis methods is employed. Firstly, based on the data detected by AE sensors, the location of the AE sources is estimated assuming the unique elastic wave velocity over the measured concrete. Then, velocity distributions in the concrete are evaluated by an advanced ray-tracing method. The technology can be applied from the bottom surface of the deck without closing bridges and stopping traffic. We have successfully verified the effectiveness of this advanced technique through the application to an in-service bridge. The comparison between the analysis result and core samples taken from the slab revealed the high validity of this technology. Accordingly, the internal damage of concrete deck can be evaluated quantitatively by applying a simple procedure using two evaluation axes, that are AE source density and propagation velocity of AE wave in the concrete. Thus, the local deterioration of the concrete deck can be classified into several stages that represent the damage levels of internal defects.

K. Watabe (✉) · H. Takamine
Research & Development Center, Toshiba Corporation, Kawasaki, Japan
e-mail: kazuo.watabe@toshiba.co.jp

H. Takamine
e-mail: hidefumi1.takamine@toshiba.co.jp

K. Watabe · H. Takamine
NMEMS Technology Research Organization, Tokyo, Japan

T. Nishida · T. Shiotani
Department of Civil and Earth Resources Engineering, Kyoto University, Kyoto, Japan
e-mail: nishida.takahiro.6e@kyoto-u.ac.jp

T. Shiotani
e-mail: shiotani.tomoki.2v@kyoto-u.ac.jp

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_65

1 Introduction

The aging of infrastructure such as bridges, many of which were built during the high economic growth period that began in the late 1950s, is a pressing social problem in Japan. After ten years from now, nearly half of the bridges experience 50 years since construction. In order to maintain these aging bridges under the limited budget, preventive maintenance is considered to be essential. To address this need, the structure has to be repaired and reinforced before apparent deterioration is exposed. However, only visual observation is recognized to be valid and has been carried out to evaluate the damage of the structure. Therefore corrective maintenance is currently used for the structural management. In order to convert to the preventive maintenance from corrective maintenance, the technique to detect the internal damage of the structure shall be crucial.

Acoustic Emission (AE) is promising candidate for the non-destructive technique to be used for the preventive maintenance. The AE method can be applied for the damage evaluation of wide area of the structure [1], such as a concrete deck of road bridges. As for the AE measurement of a bridge deck, no special device is needed to induce elastic waves because traffic load can induce AEs, which enable measurements to be conducted on in-service bridges. AE methods is widely used to understand fracture mechanism of materials in lab experiments and the potential to adopt this method for real structure is well-known. However, the actual relationship between the damage condition and AE measurement data is rarely confirmed for a real concrete structure. For this reason, application of the AE method remains an issue of engineering interpretation.

In this context, the authors conducted the damage evaluation of an RC deck of a road bridge that had been in-service for approximately 40 years. AE activities induced by traffic loading of the vehicle were measured for about one week. As the bridge deck was scheduled to be replaced for renewal, the cut-out RC panels of the slab were tested with AE tomography by hammering to estimate the velocity structure of the specimens. As a result, an accurate damage evaluation of the actual RC bridge deck was obtained by combination of the AE method and the AE tomography.

2 Experiment

Measurements were conducted on the RC deck of the in-service high-way bridge. The bridge has been in service since 1975. The RC deck was once peeled on the top and thickened to be 235 mm in 1990. Two panels of the RC deck were chosen to be measured for about one week. Fifteen AE sensors (VS30-V, Vallen Systeme GmbH) were attached on the bottom surface of each panel. The sensors were connected to the multi-channel measurement system (AMSY-6, Vallen Systeme GmbH). The resonance frequency of the sensors is 30 kHz. The sensors had

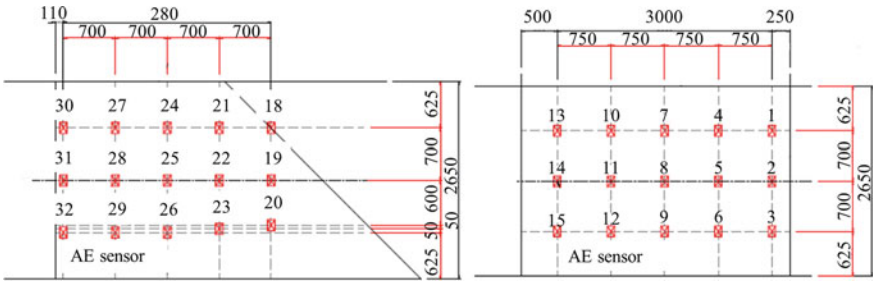


Fig. 1 Top view of the measured panels (left) panel A, (right) panel B. Installed AE sensors are numbered with red rectangular

Fig. 2 Appearance of measured RC panels. (Left) panel A, (right) panel B

captured AE signals induced by cracks in the RC deck which were activated by the vehicle load of the thru traffic. The sampling frequency was tuned to 10 MHz and the threshold amplitude was set to 53 dB. In the measurement, the bandwidth was limited between 25 and 850 kHz by the frequency filter embedded in the measurement system. The measured deck and the sensor allocation are shown in Fig. 1. The photographs of the appearance of the measured panels are shown in Fig. 2. A relatively large crack with water leakage trace is confirmed on Panel A, but there is no large crack on Panel B.

3 Result

3.1 AE Source Location Analysis

AE hits observed during the measurement period were analysed to determine the source location of each AE event. The source location of each event is calculated based on the arrival time differences between the sensors [2]. In the analysis, AE events that give high uncertainty of the location were regarded as noise and eliminated from the results. The evaluation index for the location accuracy is called "LUCY (location uncertainty)." LUCY is calculated based on the error determined by the estimated source location and measured time differences among the sensors [3]. In this analysis, the threshold for LUCY is set to 200 mm, which means the

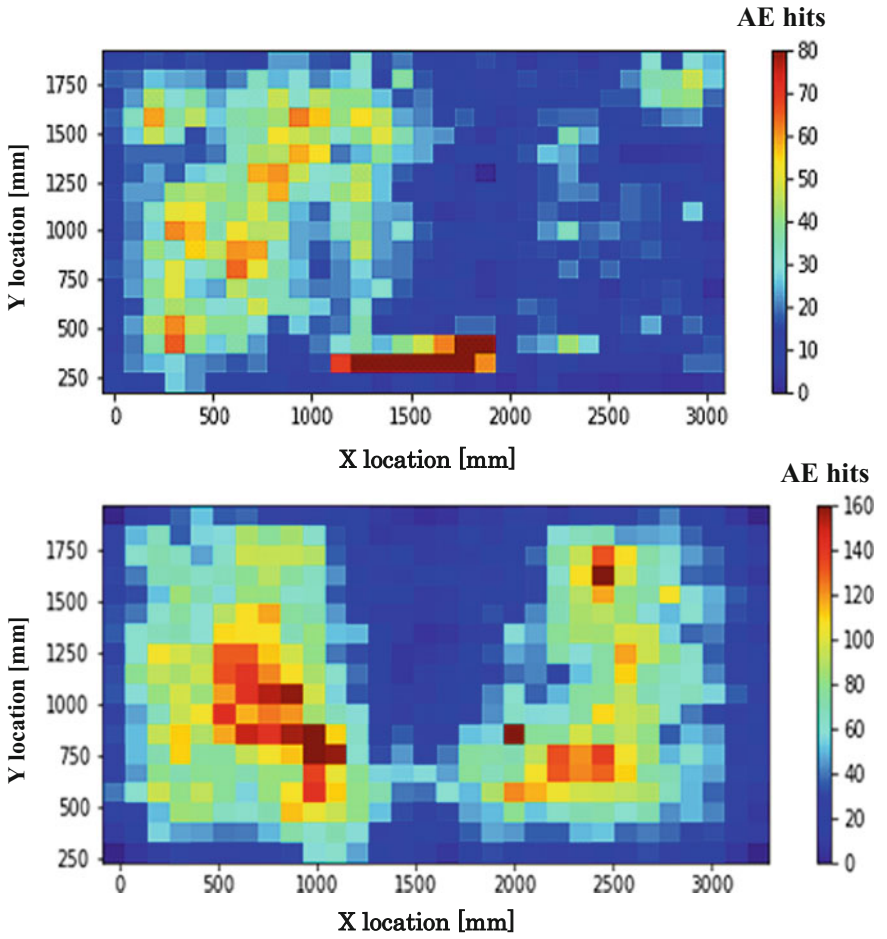


Fig. 3 Results of AE source location analysis. Direction X in the caption of the axes corresponds to the bridge axial direction. Direction Y corresponds to perpendicular to the bridge axial direction. (Top) panel A, (bottom) panel B

estimation accuracy of the location is about 200 mm. Figure 3 shows the results of the AE source location analysis. In the figure, each panel is divided into 100×100 mm^2 grids and the colour of the contour represents the AE hits density. The high AE-hit density area is coloured in warm colours and low density area is in cold colours.

Firstly, the number of AE events located in Panel B is larger than that of Panel A. For Panel B, concentrated clusters of AE events exist surrounding the central portion where shows sparse distribution of AE events. Panel A also contains sparse distribution of AE events in the right half side of the panel. These distributions of AE events do not necessarily have obvious relation with the appearance

of the under-surface of the deck. Considering the age of the RC deck, the sparsely distributed area does not correspond to intact condition as initial state of the degradation, but should be severely damaged inside the deck, which interrupt AE waves to propagate.

3.2 Characteristics Analysis of Located AE Events

Each AE events located by the previous (source location) analysis are characterized with regard to their first-hit waveforms. Peak frequency f_p, rise time t_r, and peak amplitude A_p are used to characterize the waveform. The FRA value $\frac{f_p}{t_r/A_p}$ is used to judge the nature of the AE source. It has been studied that tensile events are linked to higher FRA value, whereas shear events are linked to lower FRA value [4]. Figure 4 shows the AE source distribution of the tested panels, where each event is classified with FRA value. The event distribution of each panel is divided into two figures, which are the one with tensile events ($\mathrm{FRA} > 1.0\,\mathrm{kHz} \cdot \mathrm{V/ms}$) and the other with shear events ($\mathrm{FRA} \leq 1.0\,\mathrm{kHz} \cdot \mathrm{V/ms}$). The upper figure of each panel in Fig. 4 represents tensile events and the lower one represents shear events. The area with sparse AE events that pointed out in the previous subsection tends to have shear events for both panels. The shear events should correspond to existence of internal cracks, which generates secondary AE activities induced by traffic loadings. According to the literature [5], the average frequency of AE events drops and the RA value increases as experiencing the fracture to the final rupture. This finding is drawn from experiments on steel fibre reinforced concrete. This results support the estimation above, which suggests the existence of severe damage inside the deck.

3.3 AE Tomography Analysis

The measured RC deck was removed from the bridge for renewal of the road. The vicinities of the tested panels are cut in the shape of parallelogram as show in Fig. 5.

In general, implementing AE source location requires a number of AE sensors, which depends on the type of source location. As for three dimensional source location, when the material is assumed to be homogeneous and isotropic, namely the elastic wave velocity is constant and known, four unknown parameter of AE source as its coordinates x, y, z and t (AE generation time) shall be solved and therefore at least four AE sensors will be installed onto the materials of interest. Based on this theory, sensors could be placed in any positions within the material of interest [6].

The source locations in AE tomography are computed on the basis of the ray-trace technique [7]. The ray-trace is respectively conducted from individual receivers on a mesh that covers the material of interest, in order to compute travel times from the receiver to the nodal points. The potential emission time is the time

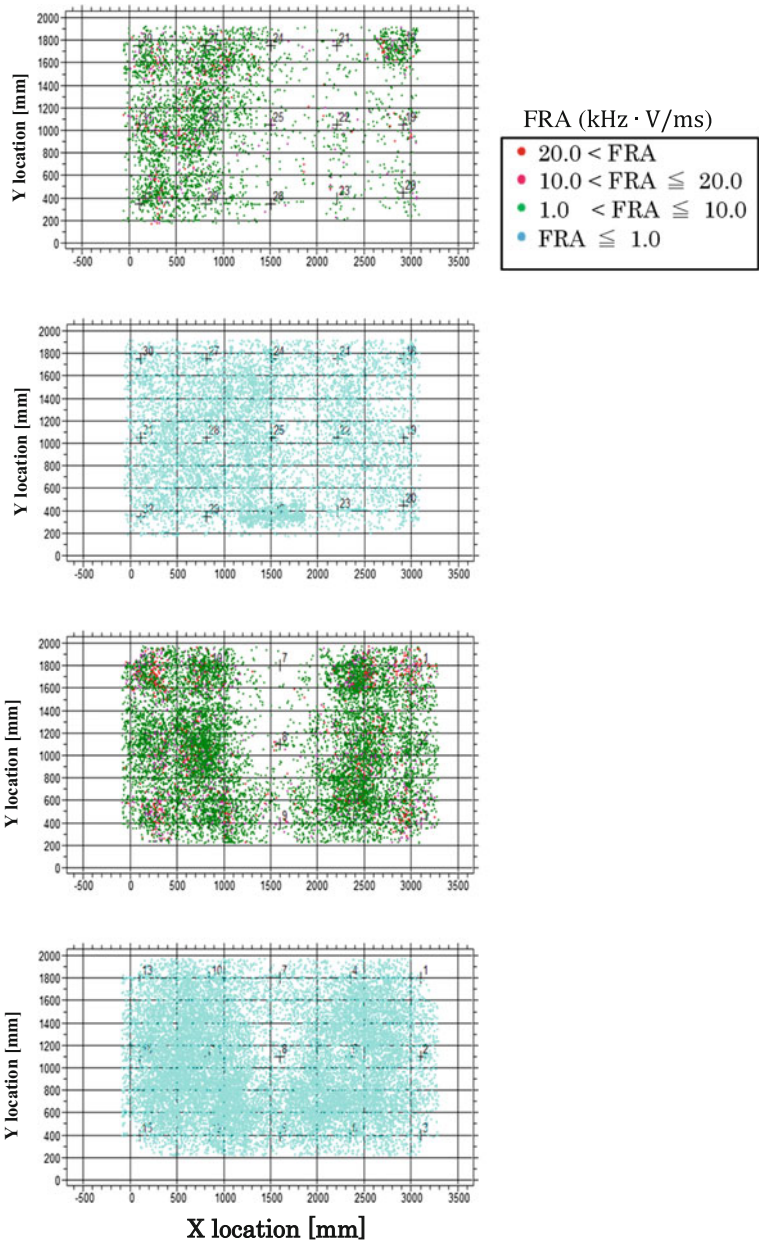


Fig. 4 Results of AE source characteristics analysis. (Top, 2nd) panel A, (3rd, bottom) panel B

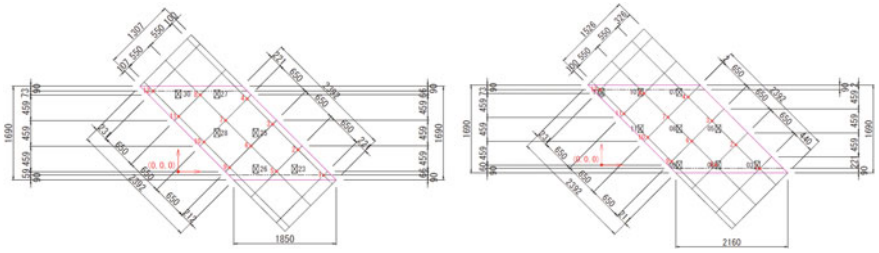


Fig. 5 Sensor location on the cut-out decks for velocity analysis (left) panel A, (right) panel B

when an AE event is to be emitted at the nodal point, so that it matches the arrival time observed at the receiver, and accordingly to the given elastic-wave velocity structure. This procedure is executed for every receiver on an individual event.

It should be noted that the potential emission times must be identical to AE generations at the source location, if the ray-paths and the elastic-wave velocity distribution are correct. However, because the accuracy of the velocity distribution is generally insufficient, the nodal point that gives the minimum variance of the potential emission times is chosen in this source location algorithm.

Since this technique is supposed to be applied for the evaluation of structural integrity and the heterogeneity of velocity distribution is frequently found in practice, the proposed procedure is quite effective. Especially, the heterogeneity would be stronger if the medium is seriously deteriorated.

It is noteworthy that the relay point introduced shall be another candidate in this algorithm. The resolution of the source location is controllable by changing the density of relay points. This helps to increase the accuracy of the identified source location even in the case that the mesh is coarse [8]. Consequently, since the 3D AE tomography technique in this study is performed with the least error, it enables to provide AE sources at the most appropriate positions even when sensors are placed only one side of targeted structure.

Three-dimensional AE tomography [9] was conducted for each cut panel to obtain velocity structures of AE wave propagation. Twelve AE sensors were attached for each cut-out panel as shown in Fig. 5. The sensors have a resonance frequency of 60 kHz, which is higher than the previous (bridge) measurement because the distance between adjacent sensors is shorter for the cut-out panels. The surface where the sensors are put is the same side as before removal from the bridge. The other side of the surface was randomly hit by an 11-mm-diameter steel ball hammer to input AE waves. The threshold for LUCY was set at 300 mm in this case. The number of nodes used in the analysis model is 12 (long axis direction) $\times$ 12 (short axis direction) $\times$ 4 (thickness direction) for both panels.

Figure 6 shows the analysis results. Three sectional areas in horizontal direction are calculated for both panels. For the Panel A, the right-center area has low velocity. The area corresponds to the area with sparse shear events in Fig. 4. The center column of the Panel B also shows low velocity, which corresponds to the

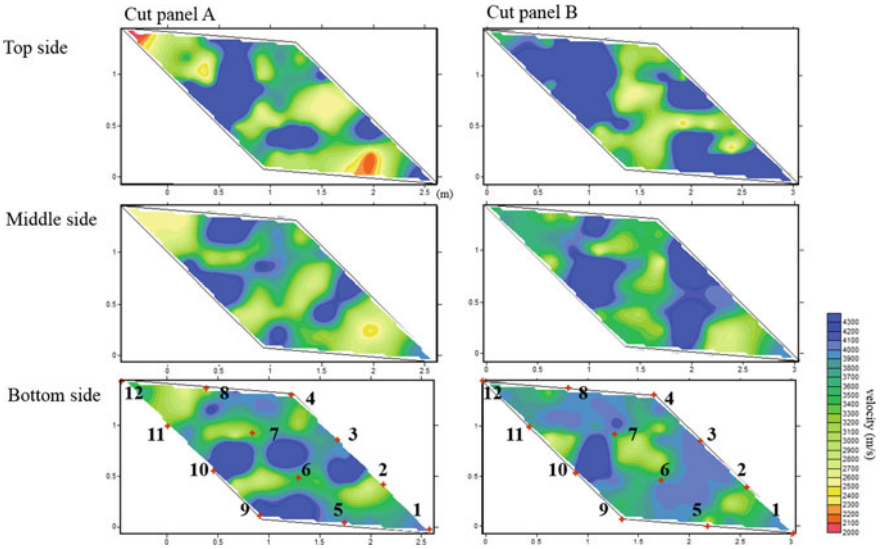


Fig. 6 Results of velocity distribution analysis. (Left) panel A, (right) panel B

area with sparse shear events in Fig. 4. These areas are presumed to contain large cracks that intercept the AE waves to propagate to the under-surface of the deck, resulting decrease of the propagation velocity in the analysis.

3.4 Core Sample Analysis

The above-estimated internal deterioration of the tested deck is verified by core sampling of the cut-out panels. Figure 7 depicts the locations of sampling points for both cut-out panels. Red filled circles in the figure represent the points where severe internal damage was suspected. The points are chosen from the portion where sparse AE events occurred and the propagation velocity is relatively low. On the other hand, blue filled circles correspond to the points where internal damage is estimated not to be so significant. These points are picked up from the potion that shows dense AE events and relatively high velocity.

Figure 8 shows the pictures of core samples cut from the panels. Surroundings of each picture has same colour with the coloured circle in Fig. 7. Almost all the samples from red points have large horizontal cracks in the upper half of them. This fact supports our above-mentioned estimation, that is large crack prevent the AE waves from propagation. Meanwhile, the rest of the samples (extracted from blue points in Fig. 7) show no apparent cracks. Elastic waves in these areas have no significant impedance to propagate, resulting in a high propagation velocity in the AE tomography analysis.

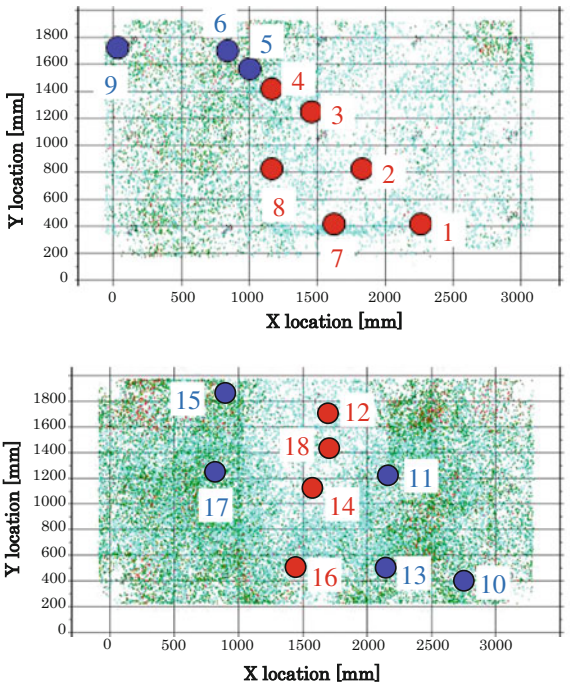


Fig. 7 Coring locations of each panel. (Top) panel A, (bottom) panel B

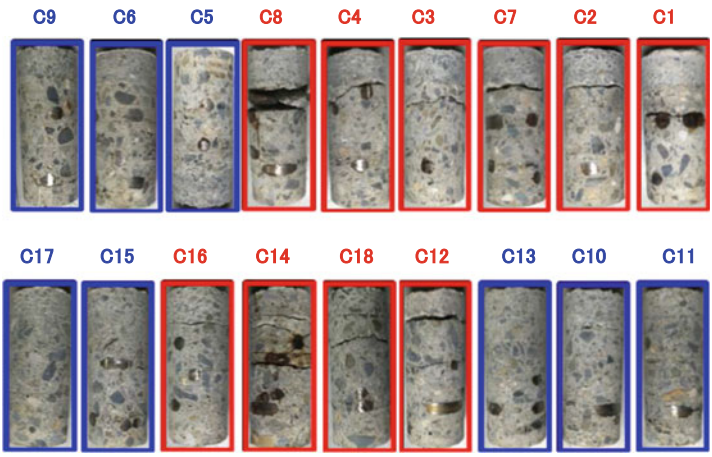


Fig. 8 Core samples extracted from the cut-out decks (top) panel A, (bottom) panel B

4 Discussion

A new approach to evaluate damage of concrete structure by AE testing is proposed based on the results obtained above. Figure 9 describes the concept of the new approach. The new method uses two evaluation axes, which are AE source density and elastic wave velocity in order to evaluate the damage of the tested structure. The above-mentioned results of experiment and analysis revealed that the severely damaged portion of the structure shows low propagation velocity and low AE source density, which corresponds to the third quadrant in Fig. 9. Therefore, the concrete deck should trace the dashed curve in the figure along experiencing fatigue by thru traffic.

Figure 10 shows the damage map of Panel A (top) and Panel B (bottom) reconstructed from the analysis results of Figs. 4 and 6. The areal density of AE source is calculated from the source location map (Fig. 4), and the analysed area is divided into two levels (dense-sparse) based on the threshold value of the source density. In this case the threshold was set to 30 hits per 100 mm^2. Furthermore the velocity distribution (Fig. 6) is also divided into two levels (high-low) based on the threshold of 3,800 m/s for example. The 4-level map (Fig. 10) is obtained by simply applying these two two-levels (dense-sparse, high-low velocity) for the whole area of the tested panel, applying the scheme depicted in Fig. 9.

As for the core samples of Panel A, all the samples from the red filled circles in Fig. 7 correspond to the level 4 except one (C5) from level 3 in Fig. 10. The samples from the blue filled circles in Fig. 7 correspond to the level 1 and 2 except one (C9) from level 4. As for the core samples of Panel B, the samples from red filled circles in Fig. 7 correspond to level 4 in Fig. 10. Most of the samples from blue filled circles in Fig. 7 correspond to level 2, which validated our new approach for the application to the damage evaluation of the concrete deck.

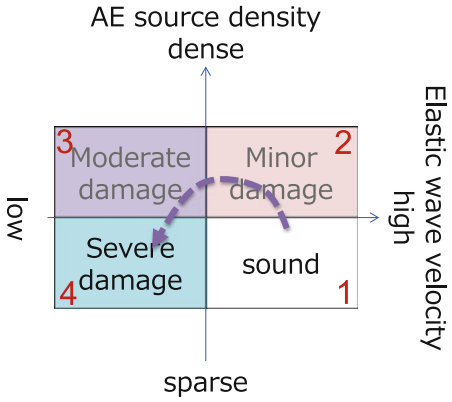


Fig. 9 Evaluation of damage based on AE source density and propagation velocity of AE wave in the concrete

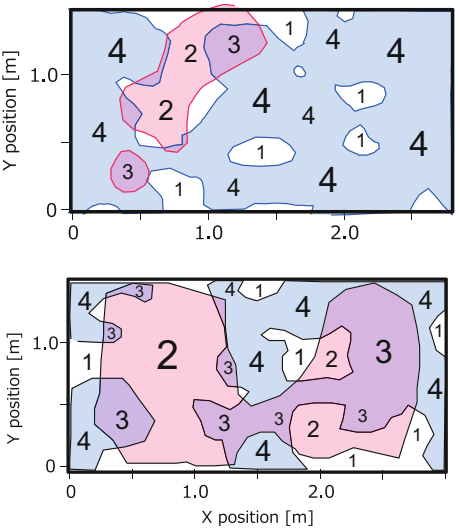


Fig. 10 Distribution of the damage level throughout Panel A (top) and Panel B (bottom)

5 Conclusion

In this study, the damage evaluation of an RC bridge deck was carried out. AE analysis is applied to evaluate the internal damage of the deck. Firstly, AE source analysis is conducted with measured AE data obtained from in-service bridge. Secondly, velocity distribution of the tested deck is estimated by three-dimensional AE tomography for the cut-out panel of the RC deck. The validity of the techniques is verified by comparing the analysis result with core samples taken from the slab. The following outcomes were obtained from this study.

- The internal damage evaluation using AE source location analysis and velocity analysis of elastic wave propagation is confirmed to be valid through the comparison between the analysis result and core sampling.
- The evaluation can be applied to the in-service bridge without closing bridges and stopping traffic.
- A new approach for assessing the damage of the concrete slabs is proposed to classify the damage into quadrant levels.

Acknowledgements This work was supported by the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

References

1. Shiotani T, Aggelis DG, Makishima O (2009) Global monitoring of large concrete structures using acoustic emission and ultrasonic techniques: case study. J Bridge Eng 14(3):188–192
2. Vallen H (2006) Acoustic emission testing, pp 27–28. Castell Publication
3. Thenikl T, Altmann D, Vallen H (2016) Qualifying location errors. In: Proceedings of 32nd European conference on acoustic emission testing, pp 495–502
4. Polyzos D, Papacharalampopoulos A, Shiotani T, Aggelis D (2011) Dependence of AE parameters on the propagation distance. J. Acoust Emiss 29:57–67
5. Aggelis DG, Soulioti DV, Sapouridis N, Barkoula NM, Paipetis AS, Matikas TE (2011) Acoustic emission characterization of the fracture process in fibre reinforced concrete. Constr Build Mater 25:4126–4131
6. Grosse CU, Ohtsu M (2008) Source location (chapter 6). In: Acoustic emission testing, p 396. Springer
7. Kobayashi Y, Shiotani T, (2016) Innovative AE and NDT techniques for on-site measurement of concrete and masonry structures. State-of-the-art report of the RILEM technical committee 239-MCM, pp 47–68
8. Kobayashi Y (2012) Mesh-independent ray-trace algorithm for concrete structures. Constr Build Mater 48:1309–1317
9. Kobayashi Y, Shiotani T, Aggelis, DG, Shiojiri H (2007) Three dimensional seismic tomography for existing concrete structures. In: Proceedings of second international operational analysis conference, vol. 2, pp 595–600

Evaluation of Condition and Damage in Reinforced Concrete by Elastic Wave Method

Takeshi Watanabe, Hayato Fukutomi, Kohei Nishiyama, Akari Suzuki and Chikanori Hashimoto

Abstract In this study acoustic emission and ultrasonic testing were applied to evaluate cracking induced rebar corrosion. It is clarified that AE activity becomes high and wave velocity detected by ultrasonic testing decreases due to cracking-induced rebar corrosion. In addition, impact echo and ultrasonic testing are applied in order to develop an evaluation formula and estimate the concrete strength for several W/C concrete specimens. As a result, it was observed that there is a good correlation between compressive strength and estimated strength in both testing methods.

1 Introduction

The evaluation of concrete condition and damage is very important to estimate the performance of reinforced concrete (RC) structures for maintenance and asset management. When these cracks reach the concrete surface, damage of rebar cor-

T. Watanabe (✉) · C. Hashimoto
Department of Civil and Environmental Engineering, Tokushima University, Tokushima, Japan
e-mail: watanabe@ce.tokushima-u.ac.jp

C. Hashimoto
e-mail: chika@ce.tokushima-u.ac.jp

H. Fukutomi
Honshu-Shikoku Bridge Expressway Company Limited, Kobe, Japan
e-mail: hayato-fukutomi@jb-honshi.co.jp

K. Nishiyama
West Nippon Expressway Company Limited, Osaka, Japan
e-mail: k.nishiyama.ab@w-nexco.co.jp

A. Suzuki
Graduate School of Advance Technorogy and Science, Tokushima University, Tokushima, Japan
e-mail: c501731033@tokushima-u.ac.jp

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_66

rosion is readily recognized by visual inspection. In this case, however the damage is referred to as the accelerated stage. In order to evaluate the performance of concrete structures, it is important to ascertain the corrosion of rebar as early as possible. In addition, compressive strength of concrete is required as fundamental data. In order to evaluate compressive strength, core samples are collected from the concrete structure. In this case, the coring holes become defects of the structure, and the holes must be repaired carefully to ensure durability of the structure. Non-destructive tests can obtain many points of data and understand distributions on a greater scale without compromising the structure.

Therefore, Acoustic Emission (AE) and Ultrasonic Testing (UT) are applied to evaluate rebar corrosion in concrete specimens. The impact echo (IE) method and UT are conducted to estimate compressive strength.

2 Evaluation of Damage-Induced Rebar Corrosion and Concrete Strength by NDT Based on Elastic Waves

Recently, several NDE methods using elastic-waves have been studied to evaluate corrosion damage of reinforced concrete. The AE method was introduced to detect corrosion-induced cracking in concrete [1, 2]. UT is applied to evaluate corrosion induced cracks and rebar corrosion [3, 4]. In this study, both AE and UT are conducted for RC specimens to detect internal cracks induced rebar corrosion.

In Japan, JSNDI (The Japanese Society for Non-destructive Inspection) standardizes NDT for concrete. UT, Impact elastic wave method and Impact Acoustics Method were standardized as NDIS 2426 Part1, Part 2 and Part 3: Method of elastic wave test for concrete structures in 2009. NDIS 2426 Part 2: Impact elastic wave method was revised in 2014. In this standard, a method which evaluates compressive strength of concrete in new concrete structures is shown as Annex D. Relationship of Ultrasonic pulse velocity and concrete properties have been studied by many researchers. Ultrasonic pulse velocity was used to estimate concrete strength and quality [5, 6]. In this study, in order to clarify the applicability of the method which is shown by the Annex D in NDIS2424, was applied to estimate concrete strength of cylindrical specimens which cast several W/C by IE and UT.

3 Experiment

3.1 Test to Evaluation Rebar Corrosion

Three RC specimens are prepared. These are concrete prisms with dimensions of 100 × 100 × 400 mm. A deformed steel-bar of nominal 13 mm diameter is embedded at 30 mm depth. The water-to-cement ratio is 55%. High early strength

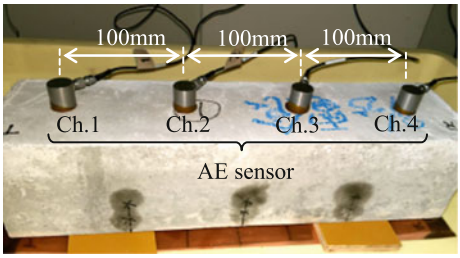


Fig. 1 AE sensor arrangements

Portland cement, crushed sand and coarse aggregate are used. Specimens were moisture-cured for 14 days. An electrolytic corrosion test was applied after curing. Specimens were soaked in water solution with 5% sodium chloride solution and electric current density charge was controlled at $2A/m^2$. A visible crack appeared on the surface 25 days after onset of the test.

The AE method was applied to one specimen to detect micro-cracks induced by rebar corrosion during the electrolytic corrosion test. Because the micro cracks penetrate from inside the surface of the specimen, AE methods were expect to be able to recognize when the cracking starts in a specimen. Four sensors were set on the side of the specimen as presented in Fig. 1. The sensors were 150 kHz resonance type. The threshold level was 40 dB.

UT was conducted once every week by using an ultrasonic measurement system. Two sensors were attached on the surface of the specimen with contact medium as shown in Fig. 2. In this case, the receiver mainly detects reflected waves from rebar as well as the boundary of the specimen. The driving frequency was set as 200 kHz and pulse voltage at 400 V.

3.2 *Test of Concrete Strength*

Table 1 shows the mixture proportions of cylindrical specimens. In order to observe the relationship between compressive strength and wave velocity, water cement ratios were set at 45, 50, 55 and 60%.

Uniaxial compressive tests were conducted for these specimens. In the tests, the age of the concrete specimens were 3 days, 7 days, 8 days and 56 days, respectively.

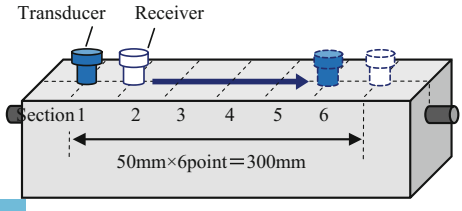


Fig. 2 Sketch of sensor arrangement of UT

Table 1 Mixture proportion of concrete specimens

W/C (%)	s/a (%)	Unit weight (kg/m^3)						SL (cm)	Air (%)
		W	C	S	G	AEA	SP		
45	45	157	349	790	965	0.007	4.2	10.0	5.2
50		164	327			0.008	3.3	12.5	5.0
55		170	308			0.009	3.1	7.5	4.5
60		175	292			0.010	2.6	10.0	3.0

Before the compressive tests, IE and UT methods were performed in order to detect wave velocity. In the impact echo test, the impactor and accelerometer were set on the surface as given in Fig. 3. The diameter of the steel ball was 10 mm and sampling time was 10 μs. The measured time-domain waveform was transformed to frequency domain by an FFT to determine the peak frequency. The peak frequency was used to detect the wave velocity using Eq. 1.

$$V_p = 2 \cdot f_0 \cdot L \tag{1}$$

V_p: wave velocity, f_o: peak frequency, L: Length of specimen.

Ultrasonic tests were conducted using ultrasonic measurement equipment as shown in Fig. 3. The equipment measured propagation time. Wave velocity was calculated by Eq. 2.

$$V_p = L/t \tag{2}$$

V_p: wave velocity, L: Length of specimen, t: propagation time.

In NDIS 2426 part 2, an evaluation formula, which shows relation between wave velocity and compressive strength, is proposed as in Eq. 3.

$$F_C = \beta \times V_p^{\alpha} \tag{3}$$

F_c: compressive strength, V_p: wave velocity, $\alpha\,\beta$: coefficient of regression.

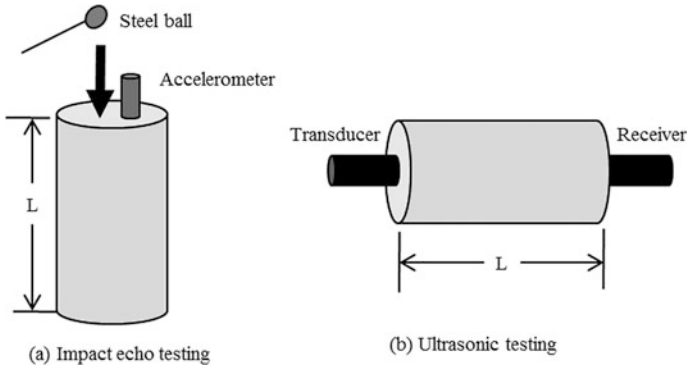


Fig. 3 Sketch of measurement by IE and UT

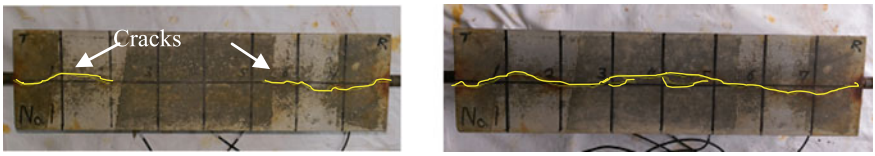


Fig. 4 Views of cracks on RC specimen (left: 25 days after, right: 30 days after)

4 Results and Discussion

4.1 Results of Evaluation Rebar Corrosion

Figure 4 shows visible cracks on the surface of an RC specimen. When the period of electrolytic corrosion test reached 25 days, visible cracks were observed on all specimens.

The relationship between AE hits and test period are shown in Fig. 5. It is observed that the cumulative number of AE hits was constant from 0 to about 11 days, and AE hits increased from 12 days rapidly. When the test period was 25 days, the increase of AE hits accelerated. Therefore it may be inferred that corrosion started after several days and cracking induced rebar corrosion in cover concrete occurred from 12 days to 25 days and, finally, cracks reached the surface.

Figure 6 shows the relationship between ratio of wave velocity detected by UT and test period. The ratio of wave velocity was calculated as velocity of test periods divided by initial velocity before electrolytic corrosion. Decrease of the ratio is recognized about 15 days after initiating the experiment. This decrease started as the number of AE hits increased as can be seen in Fig. 5. It is clear that wave velocity decreases due to corrosion induced cracks. These results suggest that AE and UT methods are able to detect cracks prior to the acceleration stage.

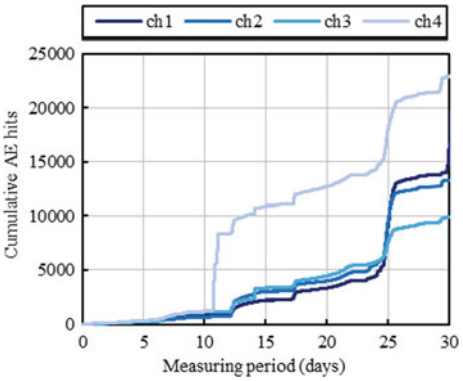


Fig. 5 Relation between cumulative AE hits and measuring period

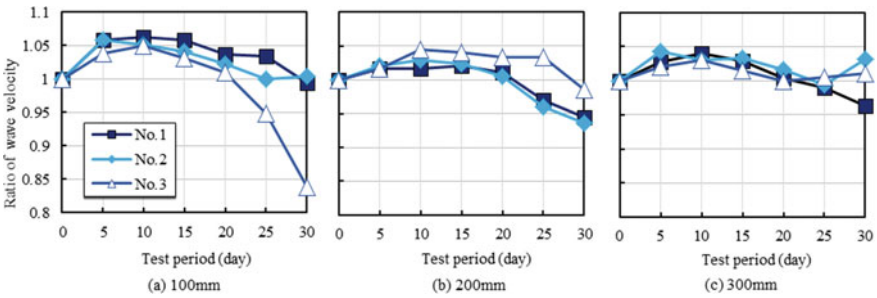


Fig. 6 Ratio of wave velocity

4.2 *Results to Estimate Concrete Strength*

The relationship between compressive strength and wave velocity measured by IE and UT are shown in Fig. 7. The curves represent the evaluation formulas of each W/C given by Eq. 3. It is observed that wave velocities have a high correlation with compressive strength in the case of the same W/C. In NDIS 2426, it is recommended that evaluation formula should take into account each mixture's proportion. Table 2 shows coefficients of evaluation formula of IE and UT. It is understood that the coefficients of IE and UT are not matched.

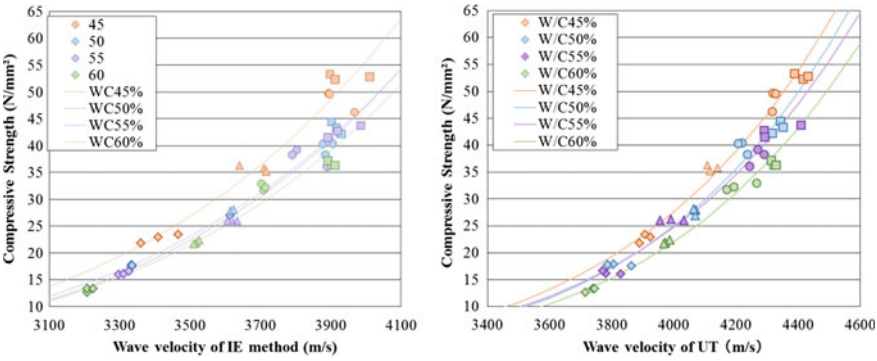


Fig. 7 Relation between compressive strength and wave velocity measured by IE and UT (left: IE, right: UT)

Table 2 Coefficients of evaluation formula of IE and UT

	Curing period (days)							
	3		7		28		56	
	α	β	α	β	α	β	α	β
IE	5.479	1.00×10^{-18}	5.416	1.50×10^{-18}	5.629	0.25×10^{-18}	5.519	0.59×10^{-18}
UT	6.995	1.80×10^{-24}	7.258	0.18×10^{-24}	6.885	3.90×10^{-24}	7.057	0.84×10^{-24}

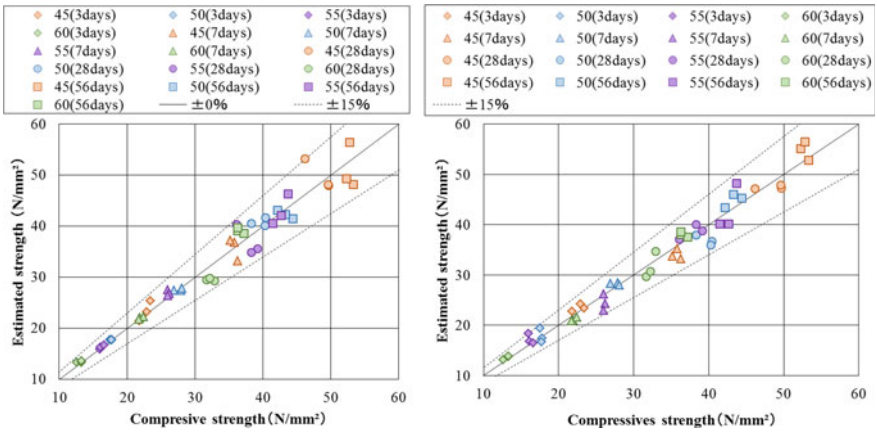


Fig. 8 Relationship between estimated strength and compressive strength (left: IE, right: UT)

Relationship between estimated strength and compressive strength is shown in Fig. 8. It is observed that the estimated strength corresponds to the actual compressive strength. Variations of IE and UT are almost of the same magnitude.

5 Conclusions

In this study, corrosion-induced cracks were experimentally evaluated. It was clarified that AE activity became high and wave parameter of ultrasonic testing decreased due to cracking induced rebar corrosion. In addition, IE and UT were applied to estimate the compressive strength of concrete. As a result, it was observed that wave velocities have a high correlation with compressive strength and the evaluation formula can estimate compressive strength with high fidelity.

Acknowledgements Part of this research was supported by the Grant-in-Aid for Scientific Research (C) (No. 15K06166).

References

1. Ohtsu M (2003) Detection and identification of concrete cracking in reinforced concrete by AE. In: Review of progress in quantitative NDE, vol 657. AIP conference Proceedings. vol 22B, pp 1455–1462
2. Kawasaki Y, Tomoda Y, Ohtsu M (2010) AE monitoring of corrosion process in cyclic wet-dry test. J Constr Build Mater 24(12):2353–2357
3. Watanabe T, Huyen Trang HT, Harada K, Hashimoto C (2014) Evaluation of corrosion-induced crack and rebar corrosion by ultrasonic testing. Constr Build Mater 67, Part B:197–201

4. Fukutomi H, Watanabe T, Hashimoto C, Miyazaki K, Ishimaru K (2016) Evaluation of ultrasonic propagation properties in reinforced concrete that reproduced rebar corrosion by artificial defect and chloride damage. Prog Acoust Emiss XVIII:539–544
5. Demirboğa R, Türkmen İ, Karakoc MB (2004) Relationship between ultrasonic velocity and compressive strength for high-volume mineral-admixtured concrete. Cem Concr Res 34 (12):2329–2336
6. Trtnik G, Kavčič F, Turk G (2009) Prediction of concrete strength using ultrasonic pulse velocity and artificial neural networks. Ultrasonics 49(1):53–60

Risk Application on Infrastructure in Conventional Contract and Performance Based Contract from Perspective of Owner

Mochammad Agung Wibowo, Evita Indrayanti, Bagus Hario Setiadji and Asri Nurdiana

Abstract Performance Based Contract (PBC) is a contract that integrates the activities of technical planning and construction. This contrasts with the conventional contract, a contract that separates the parts of the job based on the project life cycle. The aims of this study are to find the factors of potential risks that often occur in the project, then analyze the global weight of each risk factor. Primary data was obtained by distributing questionnaires and interviews. Purposive sampling method was used to distribute the questionnaires to experts. The data was then processed to obtain the criteria and sub-criteria of risk that were used to construct a hierarchy and then processed using the AHP method. Risks involved in the project were analyzed using a conventional contract and Performance Based Contract, identified by the project life cycle: The Development And Concept Phase, Design Phase, Procurement Phase, Construction Phase, and Management Phase. The identified risks are the risk that occurs in the two types of contracts. Based Contracts with Analytical Hierarchy Process (AHP), the most risk is the poor of quality control in both of conventional contract and Performance Based Contract.

M. A. Wibowo · E. Indrayanti · B. H. Setiadji
Department of Civil Engineering, Diponegoro University, 50275 Semarang, Indonesia
e-mail: agungwibowo360@gmail.com

E. Indrayanti
e-mail: evita_indrayanti@yahoo.com

B. H. Setiadji
e-mail: bhsetiadji@gmail.com

A. Nurdiana (✉)
Department of Civil Engineering, Vocational School, Diponegoro University, 50275 Semarang, Indonesia
e-mail: asrii.nurdiana@gmail.com

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_67

1 Introduction

Road infrastructure is built to support the distribution of goods and service in order to accelerate the movement of the economy in all fields. An increase in road maintenance demonstrates that more people prefer to use road transportation. One of the most fundamental efforts to realize the management of quality roads is applying the role of quality control as a guarantor of quality by supervisory consultants together with directors of engineering. According to Rahadian [12], implementation and supervision should apply innovative forms of procurement and contract practices, to bring the owner and the service provider in the scheme of risk sharing in order to minimize the interest gap and harmonize the interests of the owner and the service provider as close as possible.

Among the innovative methods of contract is Performance Based Contracts, as part of the reform of the conventional contract, which was used in the procurement of goods and services in Indonesia. A Performance Based Contract is a contract that integrates the activities of technical planning and construction. These contracts are based on a performance-based specification, where the emphasis is not on the methods and material specifications but on the performance of the work (output oriented) which is measured by the standard of road operators. In contrast, the conventional contract is a contract which separates the parts of the job based on project life cycle. The principle of a conventional contract is based on the specification method: the method of implementation, material specifications, and tools predetermined by road operators, so there is no innovation of technology to support the effectiveness of implementation and the efficiency by the service providers.

The aims of this study are to find the factors of potential risks that often occur in the project, using the method of literature study coupled with validation to the owner, then analyze the global weight of each risk factor on conventional contract and Performance Based Contract (PBC) with Analytical Hierarchy Process (AHP).

2 Literature Study

2.1 Performance Based Contract

The variance in the conventional contracts and Performance Based Contracts is on the transfer of responsibility of the work in proportion, where normally the responsibility is on the owner. The division of roles in the management of the road disclosed by Wirahadikusumah and dan Abduh [16]. It was divided based on life cycle project; the phase of planning, design, construction, maintenance and management. At Conventional Contract, the division of roles at the phase of planning is owner, in phase of design is owner, in the phase of construction is contractor, in the phase of maintenance is owner, and in the phase of management is owner. While at Performance Based Contract, the division of roles at the phase of planning is owner,

in phase of design is contractor, in the phase of construction is contractor, in the phase of maintenance is contractor, and in the phase of management is owner. At Performance Based Contract, the role of the owner in the design and maintenance phase is taken over by the contractor.

Wirahadikusumah and dan Abduh [16] explained that contract for the roadworks is generally distinguished by the following characteristics: the contract form/manner of payment (cost-based vs. price-based); consideration of risk allocation and innovation (method-based specification vs. performance-based specification); and the term of the contract (short term, long term). In the terms of form of contract/payment method; at conventional contracts the payment method is based on the actual cost plus overhead and profit (cost-based), at performance based contract accordance with the performance (performance-based). In the terms of allocation of risk; at conventional contracts is method-based specification, at performance based contract is performance-based specification. In the terms of time period; at conventional contracts is short-term (up to 1 year), at performance based contract is long-term (several years, typically up to 5 years).

2.2 Risk Management

The definition of risk management as outlined by the Project Management Institute Body of Knowledge [9] is:

1. A formal process where risk factors are systematically identified, analyzed, handled.
2. A systematic method of managing the formal that concentrates on identifying and controlling the areas or events where unwanted changes potentially occur.

The stages of risk management at construction according to Duffield and Trigunarsyah [5] are:

1. Identification of risk,
2. Evaluation of risk,
3. Allocation of risk,
4. Reduction of risk.

2.3 Analytical Hierarchy Process

Risk Evaluation utilized the Analytical Hierarchy Process (AHP), a Decision Support System developed by Thomas L. Saaty. According to Saaty [13] the steps of AHP are:

1. Determination of the components (goals/objectives, criteria, sub-criteria and alternatives) and the preparation of the component hierarchy of decision; complex issues are easily understood broken down into various substantial elements and then arranged hierarchically.
2. Assessment criteria, sub-criteria and alternatives.

Criteria and alternatives are assessed through paired comparisons. For many problems, a scale of 1 to 9 is optimal to express the variance of opinions. The values and definition of qualitative opinion from per-comparison Saaty scale can be seen in Tables 1 and 2.

3. Prioritization of criteria, sub-criteria and alternatives

Each criterion and alternatives should be evaluated via paired comparisons (pairwise comparisons). The value of relative comparison is then processed to determine the ranking of all alternatives. The weight or priority are calculated by matrix or through mathematical equations.

The considerations on pairwise comparisons to obtain the overall priorities was through the following stages:

a. Multiply this matrix of pairwise comparison,
b. Calculate the sum of the values of each row, then do the normalization matrix.

4. Checking the consistency of ratings

All elements are grouped consistently according to a logical criterion. The matrix of weight that is obtained from the pairwise comparison should relate to the cardinal and ordinal. The steps of calculating logical consistency are:

a. Multiply the matrix with the corresponding priority,
b. Summing up the results of multiplications in a line,
c. The sum of each row is divided by the concerned priorities and the scores are added,
d. Results c divided by the number of elements, then will be obtained the maximum value of χ,
e. Consistency Index (CI) = $(\chi_{max}-n)/(n-1)$.

Consistency Ratio (CR) = CI/RI, where RI is a random index consistency. If the consistency ratio $\leq$ 0.1, the calculation can be justified.

Table 1 Grading scale pairwise comparisons

The intensity of interest	Definition
1	Both elements are equally important
3	Element that one a little bit more important than any other element
5	Element which one is more important than any other element
7	One element is absolute more important than the other elements
9	One absolutely essential element than the other elements
2, 4, 6, 8	Values between two values adjacent consideration

Source Saaty [13]

Table 2 Pairwise comparison matrix element thickness

	A1	A2	…	An
A1	1	A12	…	A1n
A2	A21	1	…	A2n
…	…	…	…	…
An	An1	An2	…	1

Source Saaty [13]

3 Methodology

The method used in this research was the descriptive qualitative method. The method aims to create a description, a systematic picture and factual and accurate information on the event or the relationship between risk that will be investigated. The qualitative descriptive method that used was a survey method.

In this study, researchers identify the factors of potential risks based on the project life cycle in the conventional method of contracts and Performance Based Contracts. The identified risks were analyzed using AHP to determine the priority of risks in each type of contract (conventional contract and Performance Based Contracts) from the perception of owner.

The samples in this research consisted of similar project. The sample of conventional contracts was the roadworks in Bawen—Salatiga and the sample of Performance Based Contracts was the roadworks in Semarang—Bawen.

Primary data was obtained by distributing questionnaires and interviews to identify risk, the weight of pairwise comparison of risk, the validation of criteria and sub criteria of analysis and the validation of the result from risk analysis using AHP. Purposive sampling method was used to distribute the questionnaires to experts, while secondary data was obtained from text books, journals, theses, relevant research, contract documents and relevant regulations.

The data was processed to obtain the criteria and sub-criteria of risk that were used to construct a hierarchy. There were 5 risk criteria for this research: the development and concept phase, design phase, procurement phase, construction phase and maintenance phase.

4 Data and Analysis

The method used to analyze risk was Analytical Hierarchy Process (AHP). analysis of pairwise comparisons were conducted to determine the priority of several criteria.

Risk identification was obtained from a review of previous research then validated by respondents. The results of the identification of potential risks based on the study of literature and validation by owner is shown in Table 3.

Table 3 The results of the identification of potential risks based on the study of literature and validation by owner

No.	Project life cycle	Potential risk	Source
1.	The development and concept phase	Feasibility study	Sandyafitri and Saputra [14]
2.		Land acquisition	Sandyafitri and Saputra [14]
3.		Incompatibility scope of work	Dipohusodo [4]
4.		Limitations of data or lack of data inaccuracies and utilities	Nurdiana [8]
5.	Design phase	Revision of financial budget	Interview to owner (2016)
6.		Delay in DED process and EE or basic design	Nurdiana [8], Dziadosz and Rejmentz [6]
7.		The procedure is not fixed	Interview to owner (2016)
8.		Less meticulous in delivery of EE and DED or basic design	Interview to owner (2016), Andi [1]
9.	Procurement phase	Less meticulous in classification process	Interview to owner (2016)
10.		Limitations of human resources to understand the reference/guidelines/rules used	Queiroz et al. [10]
11.		Delay in the preparation of procurement documents	Interview to owner (2016)
12.		Long auction process	Queiroz et al. [10]
13.		Political intervention	Sigmund and Radujkovic [11], Chowdhury et al. [2]
14.		Failed auctions	Interview to owner (2016)
15.	Construction phase	Design changes	Andi [1]
16.		Poor of quality control	Interview to owner (2016)
17.		Force majeur	Sigmund and Radujkovic [11], Chowdhury et al. [3]
18.		Claim contractor for additional costs beyond the contract	Nurdiana [8]
19.	Maintenance phase	Premature deterioration of the road before the design life	Interview to owner (2016)
20.		Permitting the use of road space	Chowdhury et al. (2012)

Source Data processed

From the risk identification that has been obtained, the data was then arranged into a hierarchical structure of the risk based on each phase as shown in Fig. 1. The results of pairwise comparisons for each criteria and sub-criteria is shown in Table 4.

From the analysis in Table 4, it is known that in the conventional contract, the greatest risk is of poor of quality control, with a global weight of 18.91% and the smallest risk is feasibility study that has global weight 1.03%. In Performance Based Contracts, the greatest risk is of poor quality control, with a global weight of 11.38% and the smallest risk is the risk of permitting the use of road space with a global weight 1.71%.

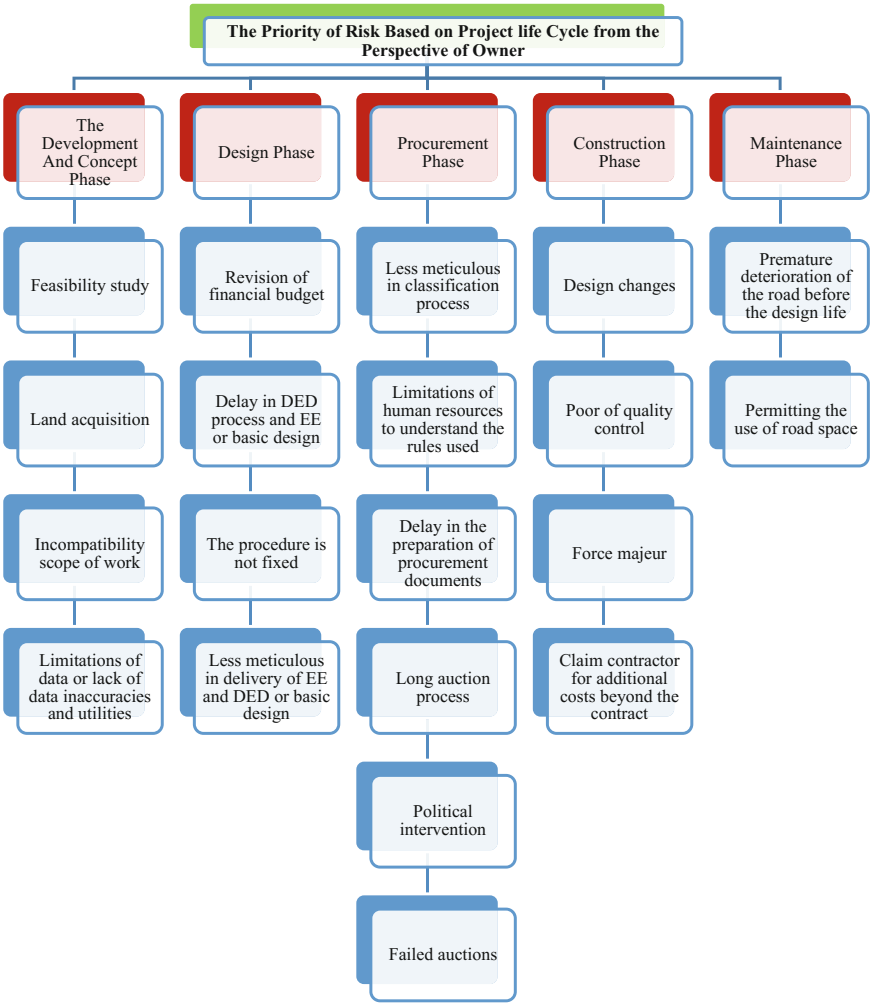


Fig. 1 Hierarchical structure for the comparison of each risk of sub-criteria on the project life cycle (*Source* data processed)

5 Discussion

Based on the results shown in Table 4, the Top Five priority risk can be seen in Table 5. These risks may cause construction failure from the validation to the owner.

From the different type of contract in construction, that is Private Private Partnership (PPP), Marques and Berg [7] found that risk in PPP project in Portugal were risks on regulation, financial, consumption, and other areas, a different model

Table 4 Value of global weight for conventional contracts and PBC

Rank	Conventional contract	% Weight	Performance base contract	% Weight
1.	Poor of quality control	18.91	Poor of quality control	11.38
2.	Premature deterioration of the road before the design life	11.55	Delay in DED process and EE or basic design	8.05
3.	Less meticulous in delivery of EE and DED or basic design	9.41	Less meticulous in delivery of EE and DED or basic design	7.38
4.	Design changes	8.21	Premature deterioration of the road before the design life	7.23
5.	Claim contractor for additional costs beyond the contract	7.30	Design changes	6.66
6.	Limitations of human resources to understand the rules used	5.77	Limitations of human resources to understand the rules used	6.40
7.	The procedure is not fixed	5.19	Limitations of data or lack of data inaccuracies and utilities	6.39
8.	Force majeur	4.30	The procedure is not fixed	6.24
9.	Less meticulous in classification process	4.16	Land acquisition	6.00
10.	Delay in DED process and EE or basic design	3.56	Incompatibility scope of work	5.06
11.	Revision of financial budget	3.36	Feasibility study	4.78
12.	Permitting the use of road space	2.61	Revision of financial budget	4.82
13.	Long auction process	2.40	Less meticulous in classification process	3.70
14.	Land acquisition	2.34	Claim contractor for additional costs beyond the contract	2.28
15.	Delay in the preparation of procurement documents	2.25	Force majeur	2.31
16.	Failed auctions	2.18	Long auction process	2.29
17.	Incompatibility scope of work	1.96	Political intervention	2.23
18.	Limitations of data or lack of data inaccuracies and utilities	1.84	Delay in the preparation of procurement documents	2.14
19.	Political intervention	1.70	Failed auctions	1.82
20.	Feasibility study	1.03	Permitting the use of road space	1.71

of risk compared to the result of this research. Different projects will have different project risk indices, different actions will be taken to minimise those risks, and different impacts will affect the project's performance. Because the project risk index had an indirect effect on the schedule performance index through progress performance, to obtain better performance, not only should risk factors be assessed at the beginning of the construction stage but also effective strategies should be carefully prepared to minimise those risks [15].

Table 5 Top 5 the significant risk

Rank	Conventional contract	% weight	Performance base contract	% weight
1.	Poor of quality control	18.91	Poor of quality control	11.38
2.	Premature deterioration of the road before the design life	11.55	Delay in DED process and EE or basic design	8.05
3.	Less meticulous in delivery of EE and DED or basic design	9.41	Less meticulous in delivery of EE and DED or basic design	7.38
4.	Design changes	8.21	Premature deterioration of the road before the design life	7.23
5.	Claim contractor for additional costs beyond the contract	7.30	Design changes	6.66

6 Conclusion

The risks involved in the project with a conventional contract and Performace Based Contract, identified by the project life cycle are as follows: The Development And Concept Phase, Design Phase, Procurement Phase, Construction Phase, and Management Phase. The identified risks are the risk that occur in the two types of contracts.

Based on the analysis of the global weight of each risk factor on a conventional contract with Analytical Hierarchy Process (AHP), the most risk is the poor of quality control while smallest risk is the feasibility study. Quality control according to the owner is a representation of the results of a work product to the specifications required and the attainment of the age of the plan that can reduce the cost of maintenance. The feasibility study is a planning of the construction work that is conducted during of new construction roads.

Based on the analysis of global weight in each risks factor on Performance Based Contracts with Analytical Hierarchy Process (AHP), the most risk is the poor of quality control while the smallest risk is the risk of permitting the use of road space. Quality is maintained service performance continuously in the duration of the contract. When the contract period is relatively long (in average seven years), the necessary quality control is the internal oversight of the service providers on performance-based contracts which is implemented by establishing a Quality Assurance Unit that is responsible for the implementation of quality programs and measuring performance using road appropriate reference standard that has been set by the contract. The road space utilization licensing is the responsibility of the owner.

Further research is the need for risk management assessment based on the perception of the performance-based contract by comparing the clause of FIDIC with the contract document, so it can be known the strengths and weaknesses of regulation that will be used as a reference document fixes for the PBC.

References

1. Andi (2006) The Importance and allocation of risks in Indonesian construction projects. J Constr Manag Econ 24:69–80
2. Chowdhury S, Rahman A, Sultana M (2012) Performance based maintenance of road infrastructure by contracting—a challenging for developing countries. J Serv Sci Manag 5:118–123
3. Chowdhury S, Rahman, A, Sultana M (2012b). An overview of issue to consider before introducing performance based road maintenance contracting. In: World academy of science, engineering and technology 62th
4. Dipohusodo I (1996) Manajemen Proyek dan Konstruksi Jilid 2, Penerbit Kanisius, Yogyakarta
5. Duffield C, dan Trigunarsyah B (1999) Project management conception to completion. Engineering Education Australia (EEA), Australia
6. Dziadosz A, dan Rejment M (2015) Risk analysis in construction project-chosen methods. J Procedia Eng 122:258–265
7. Marques RC, Berg S (2011) Risks, contracts and private-sector participation in infrastructure. J Constr Eng Manag 137(November):925–932. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000347
8. Nurdiana A (2011) Aplikasi Manajemen Risiko dari Persepsi Stakeholders. Magister Teknik Sipil, Universitas Diponegoro, Semarang, Tesis
9. Project Management Institute (2000) A guide to the project management body of knowledge, 3rd edn. PMBOK, Project Management Institute Inc, USA, Pennsylvania
10. Queiroz C, Qureshi N, Stankevich N (2005). Performance based contracting for preservation and improvement of road assets. Transport Note No. TN-27, pp. 1–11
11. Radujkovic M, Sismund Z (2014) Risk breakdown structure for construction project on existing buildings. J Soc Behav Sci 119: 894–901 (27th IPMA World Congress)
12. Rahadian H (2008) Langkah Awal Menuju Performance Based Contract Melalui Extended Warranty Period. Direktorat Jenderal Bina Marga, Jakarta
13. Saaty TL (1993) Pengambilan Keputusan Bagi Para Pemimpin. Pustaka Binaman Pressindo, Jakarta, PT
14. Sandyavitri A, dan Saputra N (2013) Analisis Risiko Jalan Tol Terhadap Prakonstruksi Studi Kasus Jalan Tol Pekanbaru—Dumai. Jurnal Teknik Sipil 9(1):1–83
15. Wiguna IPA, Scott S (2006) Relating risk to project performance in Indonesian building contracts. Constr Manag Econ 24(11):1125–1135. https://doi.org/10.1080/01446190600799760
16. Wirahadikusumah R, dan Abduh M (2003) Metode Kontrak Inovatif untuk Peningkatan Kualitas Jalan: Peluang dan Tantangan, Buku Konstruksi: Industri, Pengelolaan dan Rekayasa, ITB

Strategic Asset Planning: Balancing Cost, Performance and Risk in an Ageing Asset Base

Ype Wijnia and John de Croon

Abstract In asset intensive organizations ageing of the asset base is a key concern, especially if a major part of the asset base is built in a relatively short timeframe. Even if current performance is adequate, it may deteriorate quickly if the assets approach the end of life. To address this concern, knowledge is needed on the development of failure rate, failure consequences, and how much capital is needed to maintain the current system performance. To help asset managers address their concerns we have developed a simplified approach to model the long term development of the asset base. This approach divides the asset base in a limited number of asset types, each with their own risk and age profile, and optimization of the replacement timing. Summing over the asset types results in the optimal capital requirement and associated total system performance. The effect of budget restrictions on risk and performance can also be demonstrated. This simplified approach provides adequate results with limited effort. In this paper we will describe the approach and discuss the rationale of the applied model.

1 Introduction

In asset intensive organizations ageing of the asset base is a key concern, especially if a major part of the asset base is built in a relatively short timeframe. This is for example the case in western infrastructure asset bases, like electricity, gas, water, roads and sewage [1]. A significant part of these infrastructures was constructed in the economic boom of the 60s and 70s. Up until today most of these assets have been functioning well, but it is uncertain how much remaining lifetime there is, how the failure rate will develop when assets approach end of life, what the conse-

Y. Wijnia (✉) · J. de Croon
AssetResolutions, Zwolle, The Netherlands
e-mail: ype.wijnia@assetresolutions.nl

J. de Croon
e-mail: john.de.croon@assetresolutions.nl

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_68

quences of failure will be and what the required capital is to maintain the current performance of the system.

A first order estimate on capital requirement may be derived from replacing assets once they are fully depreciated. However, this tends to result in an exaggeration. The depreciation period is more like the minimum age an asset will reach, in order to prevent disinvestments in the books. That means most assets are not worn out at the moment they reach this age. This strategy would thus be virtually risk free but also costly. Another estimate for future costs can be derived from replacing assets at their average expected lifetime. Unfortunately this typically results in an underestimate, as a significant fraction of the assets will fail before they reach this age. This replacement strategy therefore is uncertain in its execution, though it may be less costly than replacing after being depreciated. A flaw of both estimates is that they assume replacement at a fixed age. In reality this may be true for (most) planned replacements, but certainly not for assets that are run to failure. In order to arrive at a more realistic prognosis of future expenses a probabilistic model of future failure is needed. Probabilistic approaches can easily get very complicated, if many uncertain parameters are involved. This may result in findings beyond the comprehension of the asset manager and thus without any value to the organization.

To help asset managers address this challenge, we have developed a simplified approach to model the long term development of the asset base. Key in this approach is embracing uncertainty by means of what-if analysis, instead of focusing on the precise numbers. This approach is currently being implemented at several infrastructure asset managers in the Netherlands. In this paper we will describe the approach and discuss the rationale of the applied model, required data, model calibration, accuracy and robustness, and achieved results.

2 Approach

The fundamental idea of our approach is that the future is inherently uncertain [2, 3], so that focussing on precise forecasts is meaningless. Instead, it is much more valuable to understand what the effect of uncertainties is on the performance of the asset base. To achieve this, our approach consists of 5 separate steps.

1. Developing the valuation model
2. Decomposing the asset base
3. Failure modelling
4. Data collection and validation
5. What-if analysis

Each of these steps will be further detailed in this chapter.

2.1 The Valuation Model

The valuation model is about the translation of impacts on several values (e.g. financials, safety, reliability, reputation) into a single scale, so that comparisons can be made on the value of strategies [4]. The simplest and most straightforward method to do so is to translate every effect in its monetary equivalent. After all, people are trained in comparing values to prices [5]. This allows a formal trade-off between several effects of risks and potential interventions. A practical approach for such a fully monetized value system is the risk matrix that complies with basic design rules [6]. Figure 1 holds such a well-designed risk matrix. The price per unit of the quantified consequences is constant over the severity levels. Qualitative consequences can be replaced by the financial equivalent. For example, National Commotion would be regarded as a serious consequence. The financial consequences equally bad are costs between 1 and 10 million euro.

2.2 Decomposition of the Asset Base

The asset base of any asset intensive organization easily consists of several 100s to several 1000s of asset types (= asset of specific make and model). Per asset type there may be many instances of assets, with the total numbering in millions. A utility company (electricity, gas, water) for example has at least 3 distinct assets per customer: joint, connection and meter. Each of these individual assets in theory has its own failure behaviour in terms of failure probability and failure

Potential consequences					Likelihood					
					Unlikely	Remote	Probable	Annually	Monthly	Weekly
Severity class	Finance	Safety	Reliability (customer days)	Reputation	<0,003	0,003-0,03	0,03-0,3	0,3-3	3-30	>=30
Extreme	> 10 M€	Several fatalities	> 100k CD	International commotion, >100k complaints	M	H	VH	U	U	U
Serious	1-10 M€	Single fatality or disability	10k-100k CD	National commotion, 10k-100k complaints	L	M	H	VH	U	U
Considerable	100k-1M €	Serious injuries and significant lost time	1k-10k CD	Regional commotion, 1k-10k complaints	N	L	M	H	VH	U
Moderate	10k-100k €	Lost time incidents	100-1000 CD	Local commotion, 100-1000 complaints	N	N	L	M	H	VH
Small	1k-10k€	Near misses, first aid	10-100 CD	Non-public discourse, 10-100 complaints	N	N	N	L	M	H
Negligible	<1k€	Unsafe situations	<10 CD	Internal discourse, <10 complaints	N	N	N	N	L	M

Fig. 1 Example risk matrix after [6]

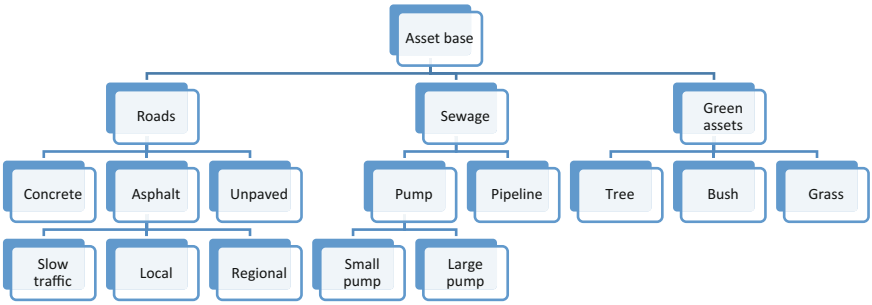


Fig. 2 Example of asset hierarchy

consequences. Yet, modelling millions of assets to forecast total asset base behaviour may be a bit extravagant. From a distance, assets of the same type will behave more or less similar. Even several types of assets may demonstrate very similar behaviour, creating some meta class of asset type. To reduce the modelling effort, we typically work down the asset hierarchy that is used by the organisation. In Fig. 2 a very high level representation is given for a typical municipality in the Netherlands.

The structure of such a hierarchy is not trivial. Roads for example could as well be clustered by their significance first (e.g. slow traffic, local, regional) and by their construction (asphalt, concrete, unpaved) afterwards. Two stop criteria are used in drilling down this asset hierarchy. The first is when the group can be described reasonably accurately by a single failure model (comparable aging, consequences and costs). The second is when the value represented by the group becomes insignificant, typically less than about 1% of the total asset base. In an asset base worth in total 1G euro, asset groups smaller than 10M euro do not need to be modelled to make a reasonably accurate forecast for the total.

2.3 Failure Modelling

For each of the asset types the development of asset costs is modelled. In its most basic form, only the time (age) dependent drivers need to be regarded: ageing/wear-and-tear resulting in repairs and corrective replacements, changing requirements (which often correlate with asset age) leading to functional replacements and planned preventive replacements. The costs resulting from the other drivers like growth, third party interventions and routine operations can be regarded as constant over the lifetime. Yet, in order to achieve more recognition (and options for validation of the outcomes!) it may be wise to include these non-time-dependent drivers. The BowTie like Fig. 3 shows these drivers and their associated reactions to restore asset adequacy.

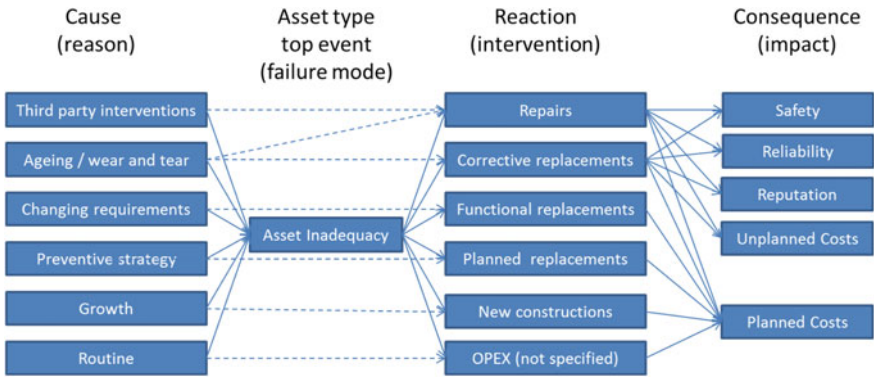


Fig. 3 BowTie diagram of asset inadequacy. The dashed lines indicate the dominant response

All causes in theory can be linked to all reactions. This is indicated by the solid lines running to and from asset inadequacy. However, in practice in most cases there is a dominant reaction, indicated by the dashed lines. There is one exception. Wear and tear for younger assets typically results in repairs, whereas the same inadequacy for an old asset may well result in a replacement. Because of the time dependence this is explicitly modelled in our strategic approach.

The typical function to describe end of life failures is the Weibull distribution [7]. However, for the assets at hand the data to derive such a distribution often are lacking. Furthermore, the field engineers find it very difficult to understand (let alone estimate) the parameters of the distribution.

Therefore we used a slightly different approach, based on a constantly growing failure rate. In this approach, two different parameters can be estimated: the undisturbed lifetime (T_{und}) which virtually all assets will reach (the depreciation period typically is a good first estimate) and the maximum age for the asset (T_{max}).

$$h(t) = c_0 e^{c_1 t}, \quad \text{conditions } h(T_{\max}) = 1,\ h(T_{\text{und}}) = 0,001 \tag{1}$$

Fitting a constantly growing failure rate to these points results in a Weibull like distribution as shown in Fig. 4.

Under the assumption that corrective intervention is more expensive (across all values) than a preventive intervention, this approach immediately provides a clue with regard to the optimal intervention strategy. That is when the cost of postponing the intervention (the risk) is larger than the benefit (i.e. depreciation and interest)[1]. This is shown in Fig. 5. This figure also shows a very important characteristic of optimization. As the optimum is a balance point between costs and benefits, which

[1]It does not matter whether absolute or conditional costs are used, as both depend on the same condition.

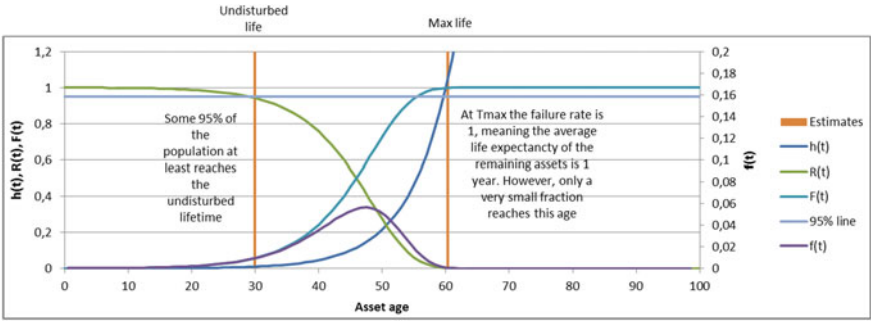


Fig. 4 Failure model based on undisturbed and maximum life

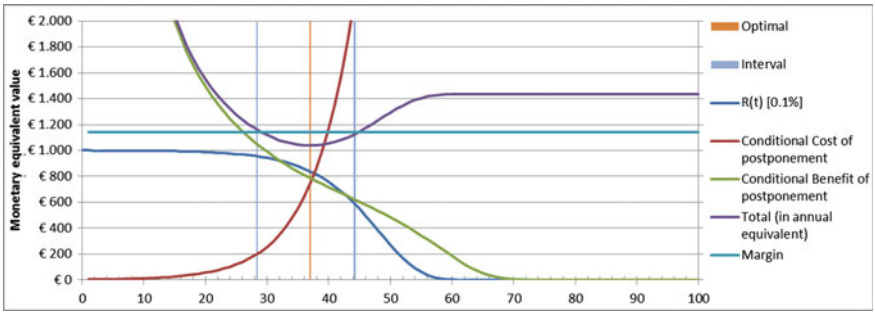


Fig. 5 Optimization of intervention interval

both develop slowly over time, being a few years off does not impact the total value much [8]. Typical intervals within 10% of the optimum span 10 years or more.

2.4 Data Collection and Validation

The approach needs relatively little data to produce a forecast. The required data is listed below per asset type.

- Asset data: Number of assets and its distribution over construction years
- Financial parameters: costs of new constructions, depreciation period, cost of planned interventions, costs of unplanned interventions, replacement value of group, book value of group
- Non-financial parameters: Undisturbed life, maximum (technical) life, non-financial consequences of failures, number of unplanned interventions over the past years

With these parameters, several validation options become available. For example, combining asset numbers with new construction cost should result in the replacement value. Correcting for depreciation (based on age profile) then should result in book value. The actual recorded number of failures can be validated by multiplying the age profile with conditional failure probability. In some cases, it may be difficult to validate the data for each individual asset. This happens for example if the asset hierarchy used in this approach does not match the hierarchy used in accounting. Failure costs for example may only be recorded as a lump sum, and not per asset type. Validation then happens on the asset portfolio instead of on the asset type.

2.5 What-if Analysis

To facilitate understanding the development of the asset base in an uncertain future, we use the approach of what-if analysis This allows the decision makers to compare several scenarios and alternative strategies within those scenarios. Typical alternatives are: a corrective strategy (reference), replacing at depreciation date, replacing at optimal age and a condition based replacement. These preventive strategies may be constrained by budget limits in several forms (corrective replacement cannot be constrained).

Typical scenarios address the following aspects:

- Growth rate of the asset base (reduction, constant, growth)
- Uncertainty in the value system, i.e. ratio between financials and non-financials
- Uncertainty in the input parameters (worst case, best guess and best case estimates)
- Innovations (coherent changes in parameters, e.g. life extension for certain maintenance action)

Because of the relatively simple approach, formulation and testing of strategies and scenarios can be done in a very short time. The typical evaluation time of a strategy or scenario is measured in seconds. This basically allows on the spot what-if analysis with the decision makers.

3 Results

3.1 Insight in Required Accuracy of the Input

Many asset managers presume that for a meaningful forecast lots of accurate data are needed. Using the what-if analysis of our approach, they learn that not all data is needed with the same accuracy, but that in many cases estimates can be used. For

example, with regard to the age profile, the gradual development of the failure rate acts as a low-pass filter. This dampens year to year differences. As long as the age profile is not skewed too much, replacing the age profile by a uniform distribution of ages between the oldest and youngest assets often is good enough for accurate predictions.

Understanding the required accuracy of the input allows for a drastic increase of the speed of implementation. The first asset types to be modeled typically are discussed in full detail, with all arguments why every single asset is special. But aggregating the data into averages, and demonstrating that using averages has almost the same quality of an answer as a very detailed assessment removes the perceived need for details. The typical time spent on the first asset types is about half a day, whereas the speed at the end is more an hour per type. Combined with the limited number of asset types (several tens) this means the whole data collection, modeling and forecasting can be performed in several weeks. If more details are needed, this can be conducted in a (second phase) scaling up of the effort, in which every single asset is taken into account.

3.2 *Developing the Long Term Strategy*

Implementing our approach provides the asset managers with an all value insight in the long term development of their asset base, which they previously found hard to achieve. This allows them to compare the value of several long term strategies in order to select the best option. Figure 6 shows 3 typical strategies that are compared:

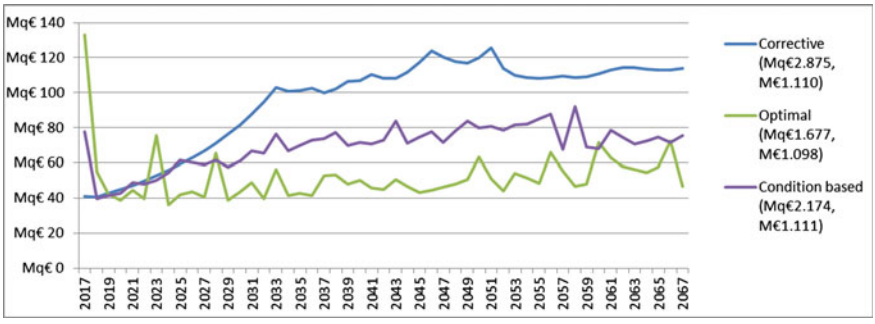


Fig. 6 Comparing several replacement strategies. The numbers in the legend represent the total present value of the strategies both in monetary equivalent (q€) as in true expenditure (€)

- Corrective, replacing assets upon failure
- Condition based, replacing assets at the minimally required condition
- Optimal, replacing assets when risks exceed benefits of postponing

As is clearly visible, the optimal strategy requires more capital upfront, but in the long run this reduces total costs of the system. The spike in 2017 is the backlog of overdue assets. This is also partly visible in the Condition based strategy. Interestingly, the strategies in the diagram do not differ much (if any at all) in terms of their Present Value costs. Financially, the strategies thus are comparable, though the timing of expenses differs. However, in terms of the total impacted value (safety, reputation, reliability) the difference is quite significant.

3.3 *Understanding Limits for the Decision Maker*

A relevant question in many organisations is how much can be squeezed out of the preventive budget. Yet, many asset managers feel that this is mortgaging the future: reduce the costs of maintenance will increase the future corrective costs. Our approach allows to demonstrate this, by implementing a budget constraint that is too tight. This is shown in Fig. 7. The implemented budget constraint is good enough for the first years. But after about 10 years, the corrective costs start to increase, and because of the budget constraint, this translates into a reduction of preventive actions. In about 25 years, there will be no preventive actions at all, so any increase in corrective costs will result in a budget overrun. The organization at this point has no control over its costs.

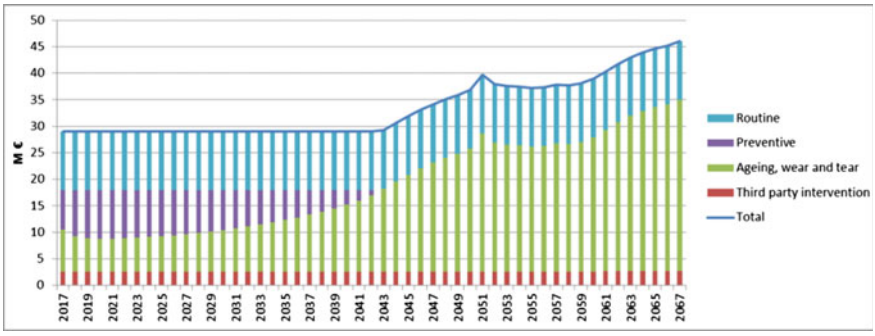


Fig. 7 Demonstrating the spiral of decline

4 Conclusion and Discussion

Many asset managers struggle with forecasting the development of their asset base in terms of costs, performance and risks. Available methods often are either too demanding in terms of effort and data, or too simple to provide a credible outcome. The resolution to this gap can be found in regarding the forecast as an optimization of the long term strategy. Optimization of slow developing risks by nature is robust in its outcome. Systematically testing for the effect of variability by means of what-if analysis allows for a controlled growth of complicatedness. The modelling effort can stop once adding more detail does not change the optimal decision anymore. For adequate strategic decisions, the required accuracy of data and model is often much less than anticipated. This means relevant outcomes can be delivered with relatively little effort in a relatively short time.

References

1. IAM (2002) International infrastructure managament manual, UK Institute of Asset Management
2. 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. Springer, Netherlands
3. Morgan MG, Henrion M (1992) Uncertainty—a guide todealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge (UK)
4. Jonkman SN, van Gelder PHAJM, Vrijling JK (2003) An overview of quantitative risk measures for loss of life and economic damage. J Hazard Mater 99:1–30
5. Sugden R, Williams A (1978) The principles of practical cost benefit analysis. Oxford University Press, Oxford
6. Wijnia Y (2012) Asset risk management: issues in the design and use of the risk matrix. In: Mathew J, Ma L, Tan A, Weijnen M, Lee J (eds) Engineering asset management and infrastructure sustainability. Springer, London
7. Gorjian N, ML, Mittinty M., Yarlagadda P., Sun Y. 2009. A review on degradation models in reliability analysis In: Kiritsis D, EC, Koronios A, Mathew J (eds) 4th world congress on engineering asset management. Springer, Athens
8. Wijnia YC (2016) Processing risk in asset management: exploring the boundaries of risk based optimization under uncertainty for an energy infrastructure asset manager. PhD, Delft University of Technology

Configuration Management—Why Asset Management Can't Do Without It

Greg Wilcock and Peter Knights

Abstract A clear understanding of the design status, the effect of any design change, and the performance of systems, equipment and facilities, is essential for effective asset management. Configuration Management (CM) provides a sound engineering framework for design identification, design change control, data management, and the audit of the safety status and performance of systems, equipment and facilities to assure their ongoing performance and safe operation. In this paper, the major elements of CM are explained. How a CM system can meet a number of the key requirements for asset management is also discussed. Two apparently disparate examples of the use of CM are provided: the first showing how CM can provide necessary data to assure safe and effective operation of military aircraft ejection seat systems, whilst the second outlines the industry requirements for the adoption of CM for the management of mine tailings dams. Properly implemented within an operation's asset management system, adherence to CM can mitigate many of the risks associated with managing complex and safety critical systems (SCS) and provides a sound basis for business decisions that can assure capability. CM should be considered as an enabler for, and an integral part of, any operation's asset management system.

1 Introduction

Asset management requires 'effective control and governance of assets … to realize value' from balancing 'cost, risk and performance' [6]. There are a number of recent cases where a lack of effective asset management of engineering systems, equipment and/or facilities has resulted in catastrophic outcomes and significant

G. Wilcock (✉) · P. Knights
Department of Mechanical and Mining Engineering, The University of Queensland, Brisbane, Australia
e-mail: gregw09@sent.com

P. Knights
e-mail: p.knights@uq.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_69

financial consequences. Examples include the recent Samarco tailings dam collapse, which resulted in 19 deaths and led to an immediate settlement of $US6 billion, and the recent $US9 billion recall of Takata car impact bag inflators [2, 10]. Conversely, an engineering management system that can provide effective design control, data management and performance reporting, particularly of complex and safety critical systems (SCS), can be a key enabler for effective asset management ensuring the maintenance of asset value.

One engineering approach to ensure design control, provide effective data management, and provide data to characterise performance, is to adopt a rigorous Configuration Management (CM) system that includes, among other aspects, audit or surveillance programs whose results provide necessary data to management teams of ongoing safety and performance.

The adoption of CM is particularly useful where systems, equipment and/or facilities are expected to have life spans of greater than 25 years. In order to maintain asset value, improve productivity and reduce life cycle costs, these systems, equipment and facilities are often subject to design modification, changes to maintenance plans, and/or operating conditions. Where this occurs CM provides a rigorous system to incorporate and maintain these design changes efficiently, effectively, safely and with minimal risk.

2 Scope

This paper provides background on CM and discusses the advantages of CM in supporting the effective control, operation and governance of assets. Specifically, this paper:

- provides a background to, and explains the elements of, CM;
- provides a discussion of how CM can achieve a number of the objectives of asset management;
- provides a case study on the approaches to the analysis of, and benefits gained from using, performance data generated from a CM-based audit program of military aircraft ejection seat system components;
- discusses the use of CM in managing the ongoing safety of mine tailings dams; and
- lists advantages for the adoption of CM within an asset management framework.

3 CM Background

Configuration Management assists with the establishment and maintenance of the technical integrity of systems, equipment and facilities and is typically part of an organisation's engineering and asset management system. CM was developed by,

and adopted within, the aerospace industry and the US Department of Defense in the late 1960s and 1970s to, in general terms, improve system availability and reliability, assure design safety, and assist in monitoring and maintaining system performance over a system's life. Since that time CM has been widely adopted, in addition to the aerospace industry, by the automotive, large infrastructure construction and Information Technology (IT) industries. CM is detailed in the current international industry standards ANSI/EIA-649B and ISO 10007:2017 [7].

CM has been implemented within complex multi-million dollar facilities and projects in disparate industries by organisations such as Airbus, the European Organization for Nuclear Research (CERN), and the £14.8 billion London Crossrail underground rail project [9]. CM is also seen as an essential tool to address project issues where un-checked changes in one sub-system can have significant consequences for other sub-systems [5]. Similarly, CM is increasingly finding more generic applications in project and operations management and is also now being recognised as an essential part of an engineering management system [3, 8]. It is also widely recognised that CM plays a vital part in system, equipment and facility life cycles to provide visibility and control on levels of performance and status [1].

4 Configuration Management

CM can be considered as a technical discipline applied to manage the evolving design of systems, equipment and facilities over their life and, in general terms, can include associated software, support and test equipment, and documentation.

Importantly, the scope of CM covers management activities applying to 'technical and administrative direction over the life cycle of a product and service' and 'its configuration identification and status'. Consequently, CM can be used to 'meet the product and serviceability identification and traceability requirements specified in other international standards', such as that for asset management [7].

In general terms, CM is made up of the following four elements:

1. identifying and recording the physical and functional characteristics of items, including large scale equipment, systems and facilities (known as configuration identification);
2. controlling design changes to those items, usually within an engineering design change management system (known as configuration control);
3. recording the status of the configuration of those items, including documentation and data that describes those items (known as configuration status accounting); and
4. regularly auditing and verifying physical characteristics against design documents, and the functional performance of the item against identified design baselines and the relevant specification (known as configuration verification and audit). This activity is also known as surveillance.

In addition to those CM elements listed above, a CM system would also typically identify:

- build standards for items that reflect authorised design configurations;
- authorised design change processes, including requirements for the review of the original item specification and methods for verifying and validating the design change;
- personnel authorised to develop, approve and accept design changes; and
- processes and responsibilities for the review of approved design changes before implementation into the operations, requiring the consideration of logistic support requirements and effects on other areas of the asset being managed. This wider view of an entire asset operation is central to mitigating risk associated with changes in one system having significant effects on other, apparently unrelated, systems.

5 CM within Asset Management

ISO 55000:2014 [6] recognises that supporting systems may be required to meet organisational objectives. For example, the standard states that asset management is data intensive and 'new tools are often necessary to collect, assemble, manage, analyse and use asset data'. Additionally, the standard states that '…the use of these tools can…improve organisational knowledge and decision making'. That is, ISO 55001:2014 [6] allows the option for organisations to use other tools to meet asset management objectives. Given that an organisation should be implanting 'planning, control…and monitoring activities to exploit opportunities and reduce risks to an acceptable level' the opportunity exists to identify systems that provide those activities. CM is one such system.

Table 1 shows the major areas in which CM can meet the requirements of an asset management system applied in accordance ISO 55001:2014 [6].

Table 1 shows that CM satisfies part of the solution for an asset management system. That is, if ISO 55001 specifies 'what' needs to occur for effective asset management, then CM can be considered as specifying 'how' a management activity occurs. Table 1 is not meant to be an exhaustive compliance checklist but rather provides an indication of the suitability of CM to meet key asset management requirements. Of course, ISO 10007:2017 [7] can only meet that part of any individual asset management requirement to the extent of the scope of the CM system. Obviously, other complementary systems and processes will need to be identified and implemented to satisfy all asset management requirements.

CM is a well-developed engineering management philosophy and can be seen as, at a minimum, complementary to an asset management system and, more likely, an integral component of, though subordinate to, an organisation's asset management system. For example, the CM system's Configuration Management Plan (CMP), which details how CM is to occur within an organisation, could be a

Table 1 ISO 10007 compliance against ISO 55001 requirements

ISO 55001 requirement[a]	ISO 10007 compliance[b]
4.3 Determining the scope of the asset management systems	5.2 Configuration management planning. The CMP[c] includes the scope of systems/equipment/facilities to be included for configuration management
4.4 Asset management system	5.2 Configuration management planning. A CMP is developed and implemented that supports the asset management objectives
5.3 Organizational roles, responsibilities and authorities	5.2 Configuration management planning. The CMP identifies and appoints personnel to leadership positions with engineering and logistics authorities 4. Configuration management responsibility. The organization identifies, describes and assigns responsibilities and authorities (including for the logistic implications of any design change), including accountability, related to the CM process
6.2 Asset management objectives and planning to achieve them (6.2.1 and 6.2.2)	5.2 Configuration management planning. The CMP details technical, regulatory and organizational requirements consistent with asset management objectives. Additionally, the CMP documents method and criteria for decision making for development, approval and incorporation of design changes[d]
7.2 Competence	5.2 Configuration management planning. The CMP details competency requirements for all staff involved in CM activities
7.6.3 Control of documented information	5.5.2 Documented information (configuration status accounting). A CM system provides necessary documentation control (including change control), storage and retrieval
8.2 Management of change	5.4 Change control. This includes a process for managing design changes including evaluating design change proposals the incorporation of changes and implementation and verification of changes
9.1 Monitoring, measurement, analysis and evaluation	5.6 Configuration verification and audit

Notes

[a]A number of the requirements of each of these sections may be applicable to CM. For simplicity, only the high level requirement has been listed

[b]This compliance may be met by either the 10007 requirement or is typically included in a configuration management plan under that general heading

[c]Configuration management plan

[d]These terms have special meaning in a design change management system and would be defined in the relevant CMP

component of the asset management system's strategic asset management plan (SAMP). Consequently, CM can be seen as an enabler for those aspects of an asset management system dealing with configuration identification, design change control, data management and system/equipment/facility performance.

6 Application of CM in the Analysis of Military Ejection Seat System Components

As an example of how CM can be implemented into a complex asset management system, data from a CM system covering a military jet aircraft's ejection seat system was considered. The performance of a number of energetic components (such as propellants) fitted to the system were monitored and analysed within a CM-based surveillance (or functional configuration audit) plan. That is, the functional performance of the component was measured against identified design baselines and the relevant specification. The analysis was to establish, amongst other requirements, the effects of age and operating temperature on component reliability. Based on this data, key asset management decisions, such as expected life and likely reliability, could be taken to ensure that the safety and capability of the system could be both assured, and managed cost effectively, over an extended period.

A number of Critical Performance Parameters (CPP), based on a component's Function and Performance Specification (FPS), were identified, whose trend in operation would provide data on which to assess the component's ongoing serviceability and, subsequently, on which larger asset management decisions could be made. Based on these surveillance results Fig. 1 shows the performance of one component as a function of total age. The horizontal lines are the upper and lower specification limits.

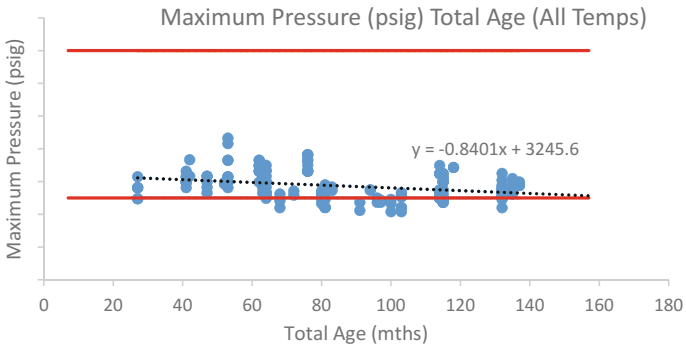


Fig. 1 Maximum pressure versus total age (all temps)

Whilst the trend in Fig. 1 suggests some dependence of the functional failures on the total age of the component, a Weibull function was subsequently fitted to the failure data to determine the extent of the dependence. The extent of this dependence then informs the asset managers on the need to, or otherwise, replace components, which has an asset cost associated with additional spare parts and maintenance to replace those items (in this case removing the ejection seat from the aircraft, over the entire fleet) and the commensurate loss of capability due to unavailability (which may or may not also have a financial value). The results at Fig. 2 suggest a clear dependence on total age ($R^2 = 0.958$) of the likelihood of the component operating outside specification limits.

To establish the effect of temperature on performance (as compared to age) a proportional hazards model (PHM) was developed using a two parameter Weibull function, given that the source data recorded two variables (time (t) (or age of component), and test temperature (T)). In this case the hazard rate is given by:

$$h(t,T) = \frac{\beta}{\eta}\left(\frac{t-t_0}{\eta}\right)^{\beta-1} e^{\alpha T} \tag{1}$$

and it can be shown that the Failure Function can now be given by:

$$ln\, ln\left(\frac{1}{1-F(t,T)}\right) = \beta\, ln(t-t_0) - \beta\, ln\,\eta + \alpha T \tag{2}$$

where α will indicate the effect of temperature on likely specification failure (and and are the Weibull shape and scale factors). Solving this equation (via a multiple regression giving an R^2 value of 0.96, indicating good correlation) confirms that age is a significant contributing cause of component failure, whilst operating temperature ($\alpha = -0.0004$) has negligible effect on failure (at any component age).

These conclusions then inform the requirements for asset management activities such as determining optimum maintenance intervals for component replacement and the (potentially expensive) requirement for restricting exposure of the

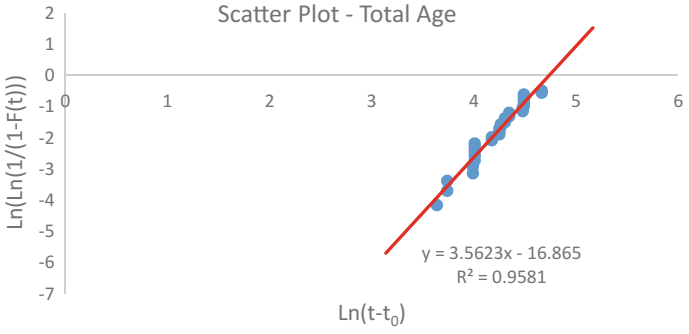


Fig. 2 Scatter plot of failure function

component to temperature variations over its life. Importantly, this data also provides the basis for quantifying the risk associated with current maintenance intervals or the risk that would be accepted should there be a requirement, for example, to temporarily extend a maintenance interval for logistics or operational reasons.

Consequently, in this example, the use of surveillance/audit program within a CM system allows an asset manager to:

- quantify the risk with the use of extant life;
- quantify the risk associated with any maintenance interval extension;
- optimise asset management activities and costs around temperature control of components; and
- reduce acquisition costs of replacement lifed components within such systems by accurately determining the component's life.

This example of the use of CM in an ejection seat system to achieve positive asset management outcomes achieves many of the benefits listed in ISO 55000:2014 [6] (Section 2.2) including improved financial performance, managed risk, and improved efficiency and effectiveness.

In summary, and for this example, the detailed analysis of data from a CM system can provide necessary assurance of ongoing asset capability, reduce asset management costs, and quantify risk associated with continued component operation.

7 Application of CM in the Management of Mine Tailings Dams

The Fundão dam was a tailings dam constructed as part of the Germano iron ore mine and operated by Samarco Mineracão SA, a joint venture between BHP Billiton and the Brazilian mining company Vale SA. The tailings dam collapsed in November 2015, resulting in the discharge of iron ore tailings and the deaths of 19 people. The incident is colloquially known as the Samarco dam collapse. As of May 2017 the circumstances of the Samarco collapse are still before the courts and, consequently, only limited comment can be made on likely causes of the dam collapse.

However, following this tailings dam collapse and a tailings dam collapse at the Mount Polley copper and gold mine in British Columbia, Canada in August 2014, the International Council on Mining and Metals (ICMM) released a report reviewing the management of tailings dams based on the Samarco, Mount Polley and Los Frailes (1998) collapses. The ICMM report provided recommendations for improvement of the management of tailings dams [4]. Those recommendations included, inter alia, the adoption of a 'formal change management process' and 'independent review (of designs) by suitably qualified and experienced professionals'. Additionally, the report recommended that an assurance program be

established 'that provides…for the frequency and scope of various levels of inspections, audits and reviews'. Whilst not using the term configuration management, these recommendations are for the establishment of processes that are key components of any CM system (ISO 10007:2017 [7] Section 5.4 *Change Control*) to manage or mitigate risk, and to improve efficiency and effectiveness.

Consequently, and based on of the recommendations of the ICMM report, the adherence to the principles of CM may have significantly reduced the effects, or potentially eliminated the cause, of this incident and the Mount Polley tailings dam collapse. That is, CM can also be considered as a risk mitigator or, alternatively, an incident prevention tool, not unlike business quality systems though with an even more direct effect on engineering operations. Of course, as with any risk mitigation tool, it is difficult to quantify the direct financial benefit of CM implementation since the incidents no longer occur and associated financial costs are not incurred.

Irrespective, and given the costs associated with incidents such as the Samarco dam collapse, CM can be seen, and has effectively been recognised as such by the ICMM, as an essential tool, in this case in the mining industry, in reducing operation risks and the larger parent business risks to levels which are considered to be acceptably low.

The recognition of the need for CM activities by the ICMM for mine tailings dam management shows how the application of CM as part of a dam asset management system can be used to manage risk, demonstrate compliance, and improve efficiency and effectiveness.

8 Advantages of Adopting CM for Asset Management

The adoption of CM principles to support any asset management system provides a number of significant advantages. Importantly, those advantages are consistent with the identified benefits of asset management as promulgated in ISO 55000:2014 [6]. That is, CM can be seen as an enabler for positive asset management outcomes. The advantages of using CM are:

- a clear understanding of the physical and functional characteristics of systems, equipment and facilities, including their current performance against the original function and performance specification;
- assurance that the engineering aspects of design changes have been approved by appropriately trained, qualified, experienced, and authorised personnel;
- assurance that logistics implications of design changes, before implementation, have been approved by appropriately experienced and authorised personnel who have adequately considered the effects on through-life support and other, apparently unrelated, systems;
- a single data repository containing all data that describes all systems, equipment and facilities subject to CM used at an operation. This includes records of design

changes, performance data, CM surveillance and results, and engineering drawings and specifications;

- a rigorous surveillance or audit program that constantly monitors the design safety and performance status of all CM-identified systems, equipment and facilities. The outcome of these surveillance programs provide assurance of ongoing safe operation and is the basis for:
  - effective and efficient system, equipment or facility management;
  - quantifying and mitigating risk using data analysis techniques;
  - reporting performance and configuration metrics, a number of which may not linked directly to business outputs but do provide confidence of ongoing safety of equipment and facilities; and
  - documented configuration and performance levels that provide a legal basis to protect an operation from suppliers whose equipment or services may not be meeting contracted of levels of performance, as well as demonstrating on-going compliance to statutory and regulatory requirements, including company policies (such as ISO 9000 compliance).

9 Conclusion

Configuration Management provides a readily developed engineering management philosophy and system to support the asset management of systems, equipment and facilities. Specifically, CM provides a system of design identification, design change control and data management, as well as a sound engineering framework through surveillance and audit for determining the safety status and performance of systems, equipment and facilities. An example using the analysis of surveillance data from ejection seat system components shows that adherence to the principles of CM can provide necessary data to assure safe and effective operation of components, and quantify risk associated with ongoing operation. In the case of mine tailings dams, the implementation of the principles of CM have been identified as essential in mitigating risks associated with the continued operation of these types of facilities.

A well-developed CM system can meet a number of the key requirements for asset management and CM can be seen as an integral component of an organisation's asset management system. That is, if ISO 55001 specifies 'what' needs to occur for effective asset management, then CM can be considered as specifying 'how' a management activity occurs. Successfully implemented, CM can mitigate many of the operational risks, and the larger parent business risks, of the operation of any system, equipment or facility, to levels which can be considered acceptably low.

CM should be considered as a key enabler for, and an integral part of, any operation's asset management system.

References

1. Ali U, Kidd C (2013) Critical success factors for configuration management implementation. Ind Manag Data Syst 113(2):250–264
2. BHP Billiton (2016) Samarco update. http://www.bhpbilliton.com/media-and-insights/news-releases/2016/12/samarco-update. Accessed 14 Dec 2016
3. Fowler A (1996) Case experience of implementing configuration management in a UK shipbuilding organization. Int J Proj Manag 14(4):221–230
4. Golder Associates (2016) Review of tailings management guidelines and recommendations for improvement. Int Counc Min Met (ICMM). London, UK
5. Hameri A (1997) Project management in a long-term and global one-of-a-kind project. Int J Proj Manag 15(3):155–157
6. ISO 55000:2014 (2014) Asset management—overview, principles and terminology. ISO (International Organization for Standardization), Geneva, 15 Mar 2014
7. ISO 10007:2017 (2017) Quality management—guidelines for configuration control. International Standards Organization, Geneva, Switzerland
8. Kluge O, Klement R (2005) Configuration management for electronics in automotive engineering. SAE International Systems Engineering, Electronics Simulation, Advanced Electronics Packaging, and Electromagnetic Compatibility, Detroit, Michigan
9. Lindkvist C, Stasis A, Whyte J (2013) Configuration management in complex engineering projects. In: 2nd international through-life engineering services conference, Elsevier, Reading, UK
10. National Highway Traffic Safety Administration (NHTSA) (2016) U.S. Department of Transportation expands and accelerates Takata air bag inflator recall to protect American drivers and passengers. NHTSA. https://www.nhtsa.gov/press-releases/. Washington DC, 4 May 2016
11. SAE International (2011) EIA 649B configuration management standard

Involving Property Practitioners for Improving Information Gathering and Distribution of Sustainability Features

Shi Yee Wong, Connie Susilawati, Wendy Miller and Diaswati Mardiasmo

Abstract Sustainable development is a critical challenge in the Australian residential property sector with different stakeholders' involvements throughout the lifecycle of a property. While substantial attention has been given in the design and construction stage, little is known about how property practitioners can contribute to the decision of home buyers in choosing a property with sustainability features. This paper examines the perspectives of property practitioners on sustainability features and their roles in distributing more complete information to buyers. A mixed-methods approach, which includes questionnaire survey and semi-structured interviews with real estate agents, property valuers and financiers, is adopted in this research. The results suggested property practitioners had different opinions on the importance of sustainability features. A sustainability-related information distribution framework which involved the roles of property practitioners, is developed to provide more effective gathering and distribution of sustainability-related information to property practitioners, and help home buyers in making an informed-decision in selecting a property.

S. Y. Wong (✉) · C. Susilawati · W. Miller
Queensland University of Techonology, Brisbane, Australia
e-mail: s31.wong@hdr.qut.edu.au

C. Susilawati
e-mail: c.susilawati@qut.edu.au

W. Miller
e-mail: w2.miller@qut.edu.au

D. Mardiasmo
PRDnationwide, Brisbane, Australia
e-mail: astimardiasmo@prd.com.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_70

1 Introduction

Housing can be seen as the most important long-term investment for a person. At the end of February 2017, 51.6% of Australian household wealth was held in residential property and the outstanding mortgage debt was AUD 1.64 trillion [6]. This high amount of the wealth investment in the residential property may potentially alert the household about decision making in spending [1, 2].

Sustainable development in residential property is facing the challenges in ensuring the building quality and continuous improvement in their environmental performance [4]. The Australian residential property sector alone accounts for 7.7% of national energy consumption in 2014–15 and it showed an increment of 1.8% compared to 2013–14 [3]. This situation may worsen as the Australian residential property supply is forecasted to face its upturn in years 2016–2019 [2] and Coleman [5] forecasted the continuous increments in energy demand for space heating and cooling. There is possibility that the home buyers may not make informed-decision in selecting a property that could help to save the operational cost of the property as they may not have sufficient knowledge about the full housing features. The complexity of housing information flow also increases the difficulties for home buyers to know thoroughly about the housing features.

Housing information is created at different stages of a building's life cycle and flow from one stage to another stage involves different stakeholders [9]. It depicts part of the common practices of government, construction industry, and property industry relating to the flow of information of a house from site development to dwelling sales. The stages of the house process include site development, landsale, purchase/settlement, planning approval, building approval, building finance, construction and dwelling sale, while the key stakeholders involved in different stages are developer, council, real estate agent, house owner, financier, valuer and architect [15]. For each stage of the process, the stakeholders hold the housing information and passed the information to the stakeholders who involved in the subsequent stage. One of the identified key relationships in the housing information flow was related to the real estate/valuation/finance and this is the key focus area in this paper.

Property practitioners, including real estate agents, valuers and financiers involve in the landsale, building finance and/or dwelling sale stages of the building's life cycle [15]. As sustainability is a complex system that involved social, economic and environmental dimensions [12], households require the involvement of property practitioners to help them recognise the importance of selecting property with sustainability features.

An Australian real estate company, LJ Hooker, worked with a group of sustainable housing design consultants to develop The 17 Things™, which form the basis for a new checklist for sustainable housing design and construction [7]. The copyright of this checklist was acquired by Commonwealth Scientific and Industrial Research Organisation (CSIRO) and incorporated these 17 features into the Liveability Real Estate Property Marketing Framework [8]. A building industry

approach identified approximately 150 distinct pieces of information that are typically created during the design and construction processes of a residential building [9]. Taking into consideration of these two approaches, Wong et al. [13, 14] suggested for a list of 50 sustainability features to be introduced to the real estate agents to provide more complete information related to sustainabilty features to the home buyers.

However, the lack of additional benefits associated with sustainable housing reflected in property valuations, which results in discouraging buyer demand and supply chain agent investments towards the purchase and/or construction of sustainable housing. There is a lack of research and analysis on how property practitioners could include information related to sustainability features during the property purchase selection phase [13]. This paper aims to present the research finding of a framework for more successful achievement of information gathering and distribution about sustainability features so property practitioners can highlight sustainability features during the property purchase selection phase in Australia.

2 Research Methodology

Exploratory study is criticised for the diversity of participants' experiences and familiarity with the topics [10]. Questionnaire faces the limitation of not having in-depth discussion, but it can reach out to a broad number of respondents and have fast responses [11]. Therefore, to collect more comprehensive data and provide more reliable results to reduce the chance of bias, this research adopted a mixed methods approach, which includes an online questionnaire survey and semi-structured interviews.

The questionnaire survey and interview involved real estate agents, property valuers and financiers (who include mortgage brokers and mortgage lenders), who have connection to the home buyers and may potentially affect their decision making in purchasing a property. The purpose of the questionnaire and interview was to ascertain the practices of the property practitioners, and identify the potential barriers in collecting and distributing information about sustainability features. This led to the collection of data through 78 questionnaire surveys (i.e. 60% real estate agents, 21% property valuers and 19% financiers) and 28 semi-structured interviews (i.e. 53% real estate agents, 18% property valuers and 29% financiers) across major Australian states with majority of responses from Queensland.

The questionnaire survey asked the property practitioners to rank the importance of 50 sustainability features based on five categories: spatial planning, occupants health and safely, occupants comfort, operations and services, and building durability, from three different perspectives: dwelling occupants; dwelling owners/investors; and professional. The semi-structured interviews identified the challenges and possible ways to improve property practitioners' involvement in the information gathering and distribution about sustainability features.

3 Results

This paper only presents part of the findings from questionnaire and interviews, which related to the ranking of sustainability features, source of information, challenges and suggestions to have better information gathering and distribution about sustainability features. The discussion on this part of the findings would lead to the development of a sustainability-related information distribution framework.

3.1 Questionnaire Survey

Property practitioners mentioned that they have quite substantial knowledge in the area of sustainability. Most of them regarded that the property price could increase due to the existence of sustainability features. Based on the analysis, location, property price, number of bedrooms, number of bathrooms and number of garages are still the priority housing features for property practitioners compared to information about sustainability features.

In terms of the ranking of 50 sustainability features based on five building categories, property practitioners have different opinions. The top three features ranked highest by the participants is shown in Table 1.

Table 1 shows that these three practitioners had different perspectives on the importance of such features and hence there is possibility that they are delivering different information to their clients. Therefore, their clients, i.e. home buyers, maybe confused on the type of information that are important to them. The housing feature of "smoke alarm" was considered as important features to both real estate agents and financiers. This may due to the fact that smoke alarm is mandatorily required by the regulations. This brought out the point of making some features mandatory to become the driver in motivating property practitioners to include such features in their practices.

Table 1 Three most important features to property practitioners

Property practitioners	Housing features	Category
Real estate agents	Smoke alarm	Occupant health and safety
	Site area Zoning/land use	Spatial planning
Property valuers	Building orientation Insulation	Occupant comfort
	Durability of building material	Occupant health and safety
Financiers	Smoke alarm	Occupant health and safety
	Type of energy services connections Water services connections	Operations and services

3.2 Semi-structured Interviews

To further examine the importance of the top three ranked features by each profession (see Table 1), semi-structured interview was employed to examine the source of information used by the property practitioners. All of the real estate agents recognised that the information about the smoke alarm, site area and zoning could be collected from site inspection and/or local government. Property valuers stated that building orientation, insulation and durability of building material could be determined through site inspection and/or building and pest inspection. Financiers, on the other hand, mentioned that they are collecting all the housing information based on the valuation report and information related to sustainability features is not their main focus. Financiers appear to not be interested in housing features in general as they may focus on the ability of the borrowers to pay back the money and their credibility. This part of the interview results seem to suggest that sustainability features is not the main focus to property practitioners. Their practices in promoting and collecting information related to sustainability features are driven by the demand of their clients and the amount of increments in property price.

The interviews proceeded with the challenges in collecting information related to sustainability features and the strategies to reduce the current challenges. The interviewees identified four key challenges in collecting information related to sustainability features and proposed five strategies that could be deployed to overcome some of these challenges, such as regulation, upskilling and modifying business practices. The strategies proposed by the interviewees seem to be able to solve the current challenges and identify different groups of stakeholders who can be involved in providing better information to the home buyers (see Table 2).

Table 2 Linkage between identified challenges, proposed strategies and stakeholders

Challenges	Suggested strategies	Stakeholders
Lack of information about sustainability features	Supply chain agents in providing more information	Group 1: Regulators and supply chain agents
Time and cost challenge	Government efforts in mandating the collection and distribution of information about sustainability features	
Lack of demand for sustainable housing	Property practitioners need to have better knowledge about sustainability features and less technical jargon	Group 2: The delivery channel (real estate agents, property valuers and financiers)
	Financial institutions could have more favourable lending criteria for sustainable housing	Group 3: Market adaptation (e.g. home buyers, home renters)
Lack of data about economic benefits of sustainability feature	Property valuers to reflect value of sustainability features in property value	

It appears that property practitioners are facing the challenges from the lack of information, low market demand and the low economic benefits in promoting the information related to sustainability features. These maybe due to the lack of transparency of the housing information that located in the local government repository and supply chain agents, and the lack of understanding from property practitioners in promoting sustainability features. Therefore, it is identified that three groups of stakeholders (i.e. regulators and supply chain agents, and property practitioners, homeowners) could have better involvement in distributing information related to sustainability features to the homebuyers.

4 Framework to Enhance Effective Information Gathering and Distribution on Sustainability Features

As the property practitioners identified their opinions on the importance of different sustainability features and challenges in collecting information about sustainability features, it is proposed to have a framework to allow more complete information gathering and distribution form the local government level to buyers through the involvement of property practitioners. A framework which combined the findings from questionnaire and interviews is proposed for improving information gathering and distribution by involving property practitioners in the process (see Fig. 1). This framework is created based on identification of the stakeholders from the proposed strategies collected from the interviewees (see Table 2).

4.1 The Regulators and Supply Chain Agents (Group 1)

The revised proposed framework was outcome from the case study databases, questionnaire and interview results. Group 1 consists of the regulators (government) and supply chain agents. It requires the government to make the information gathering and distribution about sustainability features as mandatory to property practitioners. Government wants to address the sustainability challenges to fulfil both national and international goals (e.g. amount of greenhouse gasses reduction). As a result, government seems to be the best party to enforce the rules. If government is making the information gathering and distribution about sustainability features as mandatory, property practitioners who are driven by the commissions and not interested in the sustainability features has no choice but to fulfil the requirements from government. This appears to be able to monitor the implementation activities of property practitioners.

The supply chain agents, including developers, contractors and builders are the parties that hold the building information during the building design and construction stage. The supply chain agents could provide better information to the property

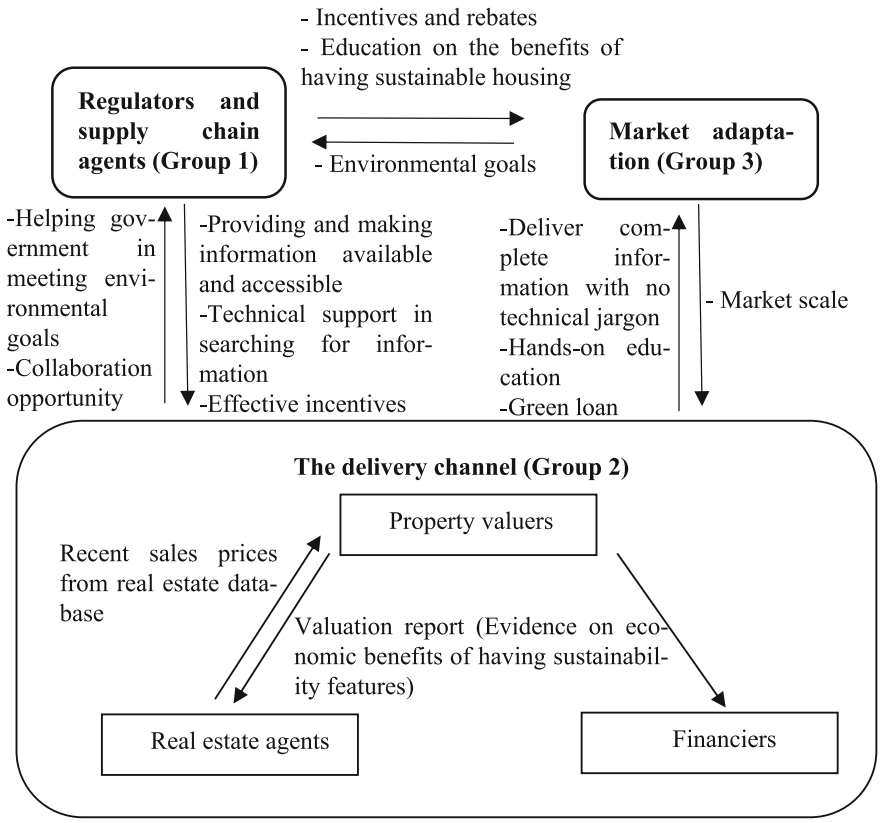


Fig. 1 A framework of improving information gathering and distribution about sustainability features

practitioners related to the housing features and collaborate with property practitioners to allow more effective distribution of information about sustainability features.

4.2 The Delivery Channel (Group 2)

The middle part of the framework consists of real estate agents, property valuers and financiers. The effective incentives and/or rebate systems from the financiers could attract potential home buyers to purchase properties with sustainability features. Green loan which offers the premium interest for potential home buyers or home investors could be able to attract the interest of the market to look into properties with sustainability features. Financiers need to have factual data on the economic benefits of having favourable home mortgage rebates. Therefore, this requires property data collectors and property research analysts to continuously collecting information related to sustainability features and analysing the economic

benefits of having such features. Property valuers have to look into such data and propose for a change in their valuation methodologies to include sustainability features into the property value.

Real estate agents, property valuers, property research analysts and property data collectors are proposed to have professional education on sustainability issues. It is proposed for property research analysts to be involved in this knowledge diffusion strategy as they are involved in extracting the data from the real estate database and analysing the trend of current property price.

4.3 Market Adaptation (Group 3)

The last stage of the framework is related to the market adaptation. The government regulations and professional marketing strategies could not be feasible without the market adaptation. Property practitioners could help with market adaptation by paying attention to the housing terminology. For example, the real estate agents should be able to explain the true meaning of star ratings and put this in the more user-friendly words in the property advertisements. The property practitioners could deliver some of the knowledge on the benefits of sustainability features to their clients. The market demand on sustainable housing has to be increased together with the increasing property price of sustainable housing. By having the increased demand and increased property price of sustainable housing, the supply of sustainable housing could be increased to meet the demand. This in turn could potentially encourage the property practitioners to collect and distribute information about sustainability features to their clients.

5 Conclusion

This paper examines the perspective of real estate agents, property valuers and financiers, in relation to the importance of sustainability features, challenges and suggestions to manage information. Their perspectives could determine the market demand and indirectly affect industry motivation on sustainable housing development. By identifying the perspectives and practices of property practitioners and proposing the knowledge diffusions for property practitioners and home buyers, the home buyers could have greater chance to realise the importance of sustainability features and hence demand for sustainable housing. The proposed framework could allow for more effective information gathering and distribution from reflective learning of property practitioners' experiences and future management of information and collaboration with other stakeholders to provide more complete information from the supply to demand side agents. With more complete information, home buyers could make informed-decision during the property purchase selection phase.

References

1. Australian Bureau of Statistics (2016). 6202.0—Labour force, Australia, Nov 2016. [Online] Available at: http://www.abs.gov.au/ausstats/abs@.nsf/mf/6202.0
2. BIS Shrapnel (2016) Australian housing outlook 2016–2019. QBE Insurance (Australia) Ltd, Sydney, Australia
3. Ball A et al (2016) Australian energy update 2016. Department of Industry, Innovation and Science, Canberra, Australia
4. Blum A (2001) Building Passport—a tool for quality, environmental awareness and performance in the building sector. In: OECD/IEA joint workshop on the design of sustainable building policies—summary and conclusions and contributed papers, Paris
5. Coleman S (2016) Australia state of the environment 2016. Commonwealth of Australia, Canberra
6. CoreLogic (2017) Housing market and economic update March 2017. RP Data Pty Ltd, Australia
7. Jenwell C. (2013) LJ Hooker: how to convert real estate agents and start a revolution. [Online] Available at: http://www.thefifthestate.com.au/articles/lj-hooker-how-to-convert-real-estate-agents-and-start-a-revolution/57284
8. Liveability (2017) About us. [Online] Available at: https://liveability.com.au/about/
9. Miller W, Stenton J, Worsley H. Wuersching T (2014) Strategies and solutions for housing sustainability: Building information files and performance certificates. Queesnland University of Technology, Brisbane, QLD
10. Neuman W (2011) Social research methods: qualitative and quantitative approaches, 7 edn. Allyn & Bacon, Boston
11. Sekaran U, Bougie R (2013) Research methods for business. Wiley, United Kingdom
12. Wang Z, Sarkis J (2013) Investigating the relationship of sustainable supply chain management with corporate financial performance. Int J Prod Perform Manag 62(8):871–888
13. Wong S, Susilawati C, Miller W, Mardiasmo D (2016) Enhancing information about sustainability features for sustainable housing delivery. In: Koskinen K, et al (eds) Proceedings of the 10th world congress on engineering asset management (WCEAM 2015). Springer, Tampere, Finland, pp 407–414
14. Wong S, Susilawati C, Miller W, Mardiasmo D (2016) Understanding Australian real estate agent perspectives in promoting sustainability features in the residential property market. In: 7th International Conference on Energy and Environment of Residential Buildings, Brisbane
15. Zedan S, Miller W (2015) Using relationship mapping to understand sustainable housing stakeholders' actions. Carnegie Mellon University, PA, Pittsburgh

Modelling the Effect of Time-Dependent Covariates on the Failure Rate of Wind Turbines

Feixiang Wu, Yifan Zhou and Jingjing Liu

Abstract Wind power is one of the most mature renewable energy technologies. In practice, wind farms are widely distributed, and wind turbines often work under different environments. Therefore, it is important to investigate the effects of environmental covariates on the failure of wind turbines. This paper focuses on the modelling of the effect of time-varying environmental covariates on the failure rate of a component of wind turbines. The relations between the component failure rate and the environmental covariates are modelled using the proportional hazard model (PHM) based on a practical dataset. The modelling process can be used to evaluate the effects of environmental factors on the failure rate of other components.

1 Introduction

Wind power has become one of the most mature and fastest growing renewable energy technologies. Currently, the wind power industry faces the challenge of high maintenance cost, which is also the bottle neck of the further development of wind power industry. In reality, wind farms are widely distributed, and the environmental factors of different wind farms have significant variations. These variations can cause dissimilar failure rates of wind turbines. The mechanical components (e.g., bearings and gear boxes) in areas with high wind speed can have a large failure rate. On the other hand, the failure rate of electronic components of wind turbines is more sensitive to the temperature.

F. Wu (✉) · Y. Zhou (✉)
Mechanical Engineering School, Southeast University, Nanjing 211189, China
e-mail: 220160266@seu.edu.cn

Y. Zhou
e-mail: yifan.zhou@seu.edu.cn

J. Liu
Economic & Technological Development Zone, Xinjiang Goldwind Science & Technology Co., Ltd., No. 8 Bo Xing 1st Road, Beijing 100176, China

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_71

In reliability engineering, indicators extracted from the environmental factors that affect the failure rate are often referred to as covariates. Covariate models are used to describe the relationship between the failure rate and covariates. In covariate models, the failure rate is regarded as a function of component age and covariates. The PHM proposed by Cox [1] is a commonly used approach to model the relationship between covariates and failure rates. Parameter estimation and goodness of fit test approaches for the PHM with time independent covariates can be obtained from Ref. [8]. Liao et al. [4] introduced a parameter estimation method for the PHM with time-dependent covariates. To overcome the shortage of failure event data, an approach to incorporate expert knowledge into parameter estimation was proposed in Ref. [11]. An important assumption of the PHM is that the effects of covariates on the failure rate are time-independent, which is not always true. Kumar and Westberg [3] converted a covariate with time-dependent effects to several covariates with time independent effects that can be processed by the PHM. An alternative way to model the relationship between covariates and a failure rate process is the additive hazard model (AHM). The AHM does not have the proportional hazard rate assumption, which is therefore more flexible than the PHM. For some applications, the AHM has more plausible performance than the PHM [5]. In the original AHM, the regression coefficient vector is time-dependent, which makes the AHM more flexible. However, the number of covariates is limited due to the complexity of parameter estimation. Lin and Ying [5] treated the regression vector as time-independent to simplify the AHM. Subsequently, they could estimate model parameters using a partial likelihood function similar to the PHM. McKeague and Sasieni [7] proposed a partly parametric additive risk model to strike a balance between the flexibility and mathematical tractability. The partly parametric additive risk model assumed that only a part of elements in the regression vector were time dependent. Some research developed a hybrid model by combining the PHM and the AHM. Lin and Ying [6] investigated an additive-multiplicative hazard model and proposed an efficient parameter estimation method. Based on the hybrid model proposed by Lin and Ying, Torben and Thomas [9] treated the regression coefficients as time-dependent variables. A transformed hazard model was proposed as a unified formulation of the additive, multiplicative, and hybrid hazard model [10].

Among the above mentioned models, the PHM is the most widely used. In the PHM, the partial-likelihood method can be used to estimate the coefficients of environmental covariates without knowing the baseline failure rate function. The formulation of the partial-likelihood function is uncomplicated and its maximization process is robust. Therefore, the PHM is a suitable tool to process large datasets. Another significant advantage of the PHM is its ability to deal with time-varying environmental covariates. Therefore, this paper uses the PHM to analyse the influence of environmental factors on the failure rate of a component in wind turbines.

2 Introduction of the Dataset

The dataset used in this study is obtained from the failure and operational records of practical wind turbines over about five years. Because a wind turbine is a complex system and consists of a large number of components, this paper only focuses on components liable to fail. For confidentiality reasons, the component is named component M. In the dataset, a total of 6976 units of component M have the enough history of operating environment. According to the failure records, 424 of the 6976 units of component M failed, and the rest are still running at the end of the record duration. Environmental factors are divided into two categories: climatic factors and technical improvements. Climatic factors such as temperature and rainfall often change with regions and seasons. Three climatic factors are taken into consideration in the study. They are denoted as factor A, factor B, and factor C, because of confidentiality. The values of factors A and B are obtained from the SCADA system in wind turbines. Unfortunately, the history of factors A and B is less than two years. Subsequently, three indicators, the monthly mean, minimum and maximum values are extracted from the raw data of factors A and B. It is assumed that there is no significant difference in climatic indices in the same month of different years. Factor C is systematically measured during location selection of wind farms. Therefore, factor C is a constant at each wind farm. The original data are transformed in this paper for the confidential reasons. Component M in this dataset come from 70 different wind farms. These 70 sites cover a large number of different types of climatic patterns. Another type of environmental factors is technical improvements. The technicians continuously improve the design of component M according to the feedback from fields. These changes can have a significant impact on the lifetime of the component. Two technical improvements have been carried out to component M, which are denoted as improvements I and II. The beginning times of the two technical improvements at different wind farms are recorded. Finally, nine indicators listed in Table 1 are used as environmental covariates in this paper.

Table 1 The environment factors and covariates considered in this paper

Environment factor	Covariate
Factor A	Monthly mean value of factor A
Factor A	Monthly minimum value of factor A
Factor A	Monthly maximum value of factor A
Factor B	Monthly mean value of factor B
Factor B	Monthly minimum value of factor B
Factor B	Monthly maximum value of factor B
Factor C	Value of factor C
Improvement I	Whether improvement I is performed
Improvement II	Whether improvement II is performed

3 Data Analysis Using the PHM

In reliability engineering, the PHM is used to evaluate the influence of working environment on the failure rate of equipment. The formulation of the PHM is

$$h(t, \mathbf{X}(t)) = h_0(t)\exp\left(\mathbf{b}^{\mathbf{T}}\mathbf{X}(t)\right) = h_0(t)\exp\left(\sum_{m=1}^{M} b_m X_m(t)\right). \quad (1)$$

Here, $h(t, \mathbf{X}(t))$ represents the failure rate at time t under the influence of the environmental covariate vector $\mathbf{X}(t)$ that is also a function of time t. The function $h_0(t)$ is the baseline failure rate. The notation b_m $(m = 1, 2, \ldots, M)$ is the coefficient of covariate m. These coefficients quantify the effect of the covariates on the failure rate. The advantage of the PHM is that the covariate coefficients can be estimated by maximising the partial-likelihood function even when the baseline failure rate function is unknown. After the coefficients of covariates are estimated, the effect of each covariate on the failure rate is investigated by the Z test. Furthermore, the likelihood ratio test is employed to find the optimal combination of covariates.

3.1 Preliminary Analysis and Screening of Covariates

Before PHM modelling, the covariates are preliminarily analysed to screen covariates that do not affect the failure rate of component M. The values of monthly average of Factor A are largely between −50 and 100, and are concentrated in the range from −10 to 70. The monthly mean value of Factor A is then divided into six intervals as shown in the first row of Table 2. The ratio of the total number of failed components to the total number of running components in different time intervals is calculated, which is called the failure and running ratio (F-R ratio) in this paper. It is an indicator of the effect of a covariate on the failure rate. The F-R ratios of the six segments are demonstrated in the last row of Table 2. The F-R ratio has an obvious increasing trend with the rise of the mean value of Factor A, except when the mean value of Factor A is higher than 70. It can be caused by the relatively small population of component M that works under such a high Factor A.

Similar analysis is conducted to other climatic factors. Finally, the monthly average and maximum values of factor A, the monthly average of factor B, and the

Table 2 The failure and running ratio under different monthly mean values of factor A

	Below −10	−10 to 10	10 to 30	30 to 50	50 to 70	Above 70
Failure	49	50	64	130	122	9
Running	24,783	30,607	35,696	54,829	40,470	3661
F-R ratio(10^{-3})	1.977	1.634	1.793	2.371	3.015	2.458

value of factor C are selected as the candidate environmental covariates in the PHM. However, the monthly mean and maximum values of factor A have a highly linear positive correlation with a correlation coefficient of 0.9494. Therefore, it is not appropriate to simultaneously consider the two indicators in the PHM. This issue is further discussed in Sect. 3.2. In addition, technical improvements I and II reduce the F-R ratio significantly. Therefore, the two improvements are also modelled in PHM.

3.2 PHM Modelling and Analysis

PHM modelling is to find the optimal combination of covariates and estimate the model parameters. Firstly, the monthly mean and maximum values of factor A are both included in the PHM. The results of parameter estimation are shown in Table 3. The P values of the Z test about both the two covariates in Table 3 are more than $\alpha = 0.05$, which indicates that the effects of the two covariates on the failure rate are not significant. The reason of this conclusion is that the standard deviations S_b of the estimated coefficients are large, which shows that the parameter estimation is unreliable.

When only the monthly mean or maximum value of factor A is considered in the model, the parameter estimation result is in Table 4. The P-value and the standard deviation are small when the monthly mean and maximum value are considered separately in the PHM model. This example shows that the PHM including two highly correlated covariates can cause misleading conclusions. Consequently, only the monthly mean value of factor A is adopted in the PHM.

Tables 5 and 6 list the result of parameter estimation when the factor C or the monthly mean value of factor B are the only covariate of the PHM. The P values of the two Z tests indicate that both the monthly mean value of factor B and the value of factor C have significant effects on the failure rate. Similar analysis shows that both the technical improvements I and II reduce the failure rate obviously, while the effect of improvement I is more significant.

According to the above analysis, the monthly mean value of factor A, the mean value of factor B, the value of factor C, technical improvement I, and technical improvement II are used as the covariates in the PHM. The result of parameter estimation is shown in Table 7. Because the coefficient of the monthly mean value of factor B does not pass the Z test, the PHM without factor B is further

Table 3 The parameter estimation result of the PHM with the monthly mean and maximum values of factor A as covariates

Covariates	b	S_b	P-value
Average of A	0.003611	0.005891	0.539921
Maximum of A	0.005661	0.006289	0.368001

Table 4 The parameter estimation result of the PHM with only monthly mean or maximum value of factor A

Covariates	b	S_b	P	$\log(L)$
Average of A	0.008628	0.001938	8.48E−06	−3679.84
Maximum of A	0.009297	0.002072	7.23E−06	−3679.63

Table 5 The parameter estimation result of the PHM with the monthly mean value of factor B as the only covariate

Covariates	b	S_b	P-Value	$\log(L)$
Mean of factor B	0.03590	0.01156	1.90E−03	−3685.59

Table 6 The parameter estimation result of the PHM with the value of factor C as the only covariate

Covariates	b	S_b	P-Value	$\log(L)$
Factor C	0.008466	0.001416	2.23E−09	−3672.99

Table 7 The parameter estimation result of the PHM with three climatic indicators and two technical improvements

Covariates	b	S_b	P-Value
Mean value of factor A	0.007114	0.002024	3.63E−04
Mean value of factor B	0.021772	0.012725	8.71E−02
Factor C	0.011496	0.001501	1.87E−14
Improvement I	−1.44553	0.129337	5.32E−29
Improvement II	−0.56683	0.229621	1.36E−02

investigated. The result of parameter estimation is shown in Table 8. The likelihood ratio test is performed, which supports the conclusion that the PHM without the mean value of factor B fits the dataset better. Finally, the monthly mean value of factor A, the value of factor C, and two technical improvements are selected as the optimal combination of covariates. According to the point estimates of the covariates, following results are obtained:

Table 8 The parameter estimation result of the PHM after the monthly mean value of factor B is not included

Covariates	b	S_b	P-Value
Mean value of factor A	0.007215	0.002023	3.63E−04
Factor C	0.011915	0.001498	1.78E−15
Improvement I	−1.45415	0.129517	2.99E−29
Improvement II	−0.58678	0.229339	1.05E−02

- When the monthly mean value of factor A rises by 10, the failure rate of component M will increase by 7.5%.
- When the value of factor C rises by 10, the failure rate of component M will increase by 12.7%.
- The failure rate of component M decreases to 23.4% after technical improvement I.
- The failure rate of component M decreases to 55.6% after technical improvement II.

Although both factors A and C have obvious effects on the failure rate, the effects of factor C are more significant. In the obtained dataset, the value of C ranges from 80 to 200, and the failure rate can increase to 417.8% accordingly. The two technical improvements also enhance the reliability of component M significantly. When both the two improvements are conducted, the failure rate reduces to 13% of the original.

4 Conclusion

This paper studies the influence of time-varying environmental covariates on the failure rate of component M of wind turbines. The size of dataset is large, and the covariates change over time. Therefore, the PHM is selected as the modelling tool to quantitatively analyse the effects of environmental factors. An approach to primary selection of potential covariates are developed. Then, the modelling process including parameter estimation and significance test of coefficients is demonstrated. Furthermore, the likelihood ratio test is used to compare the fitness of different combinations of covariates. A practical dataset about component M of wind turbines is used to demonstrate the application of the modelling process. The modelling process can also be applied to other components working in time-varying environment.

Acknowledgements The research work is supported by the National Natural Science Foundation of China (Grant Nos. 71671041).

References

1. Cox DR (1972) Regression models and life-tables. J Roy Stat Soc Ser B (Methodol) 34 (2):187–220
2. Kravdal Ø (1997) The attractiveness of an additive hazard model: an example from medical demography. Eur J Popul 13(1):33–47
3. Kumar D, Westberg U (1996) Proportional hazards modeling of time-dependent covariates using linear regression: a case study [mine power cable reliability]. IEEE Trans Reliab 45 (3):386–392

4. 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, 2006. RAMS'06 annual, pp 127–132
5. Lin DY, Ying Z (1994) Semiparametric analysis of the additive risk model. Biometrika 81 (1):61–71
6. Lin DY, Ying Z (1995) Semiparametric analysis of general additive-multiplicative hazard models for counting processes. Ann Stat 23(5):1712–1734
7. McKeague IW, Sasieni PD (1994) A partly parametric additive risk model. Biometrika 81 (3):501–514
8. Prasad PVN, Rao KRM (2002) Reliability models of repairable systems considering the effect of operating conditions. In: Reliability and maintainability symposium, 2002. Proceedings Annual2002, pp 503–510
9. Torben M, Thomas HS (2002) A flexible additive multiplicative hazard model. Biometrika 89 (2):283
10. Zeng D, Yin G, Joseph GI (2005) Inference for a class of transformed hazards models. J Am Stat Assoc 100(471):1000
11. Zuashkiani A, Banjevic D, Jardine AKS (2006) Incorporating expert knowledge when estimating parameters of the proportional hazards model. In: Reliability and maintainability symposium, 2006. RAMS'06 Annual2006, pp 402–408

Configuring and Optimizing the Maintenance Support Resource Based on a Double Layer Algorithm

Xiwen Wu, Bo Guo, Ping Jiang and Shiyu Gong

Abstract Optimizing and configuring the maintenance support equipment is investigated and it is translated into RACP (resource availability cost problem) with uncertain activity durations. The problem aims to minimize the cost of the maintenance support system with the constraints of project finish time. Precedence relations among maintenance activities are coded by random key, and then a double layer optimization algorithm is used to solve the problem. The inner layer configures the amount of maintenance equipment, while the outer layer schedules the project execution procedure. Particle swarm optimization embedded with scatter search is applied to produce high quality solutions. Finally, comparative computational experiments are designed and run on benchmark datasets to test the efficiency and performance of methods. The computational results show that this optimization method has significant value to solve the resource configuration problem and the outcome has actual engineering significance.

1 Introduction

Maintenance resources are the physical foundation of maintenance support. The planning and scheduling of maintenance resources has a significant effect on the cost and efficiency of maintenance support. The investigation of maintenance resource includes predicting the resource demand, allocating and resource amount,

X. Wu (✉) · B. Guo · P. Jiang · S. Gong
College of Information System and Management, National University of Defense Technology, Changsha, China
e-mail: wuxiwen11@nudt.edu.cn

B. Guo
e-mail: boguo@nudt.edu.cn

P. Jiang
e-mail: jiangping@nudt.edu.cn

S. Gong
e-mail: sygong@nudt.edu.cn

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_72

and scheduling the activity execution [3, 8, 20]. This paper considers a maintenance support resource optimization problem for a combat unit that consists of several weapons. Each weapon will go through its maintenance activities and requires a certain amount of resources. The durations of activities are uncertain but subject to certain distributions. The problem aims to minimize the resource cost taking into account a project deadline and precedence relations among the activities.

In the case that all the activity durations can be determined, the problem can be regarded as a resource availability cost problem (RACP), which was first introduced by Möhring [13] as a NP-hard combinatorial optimization problem. Meng et al. [12] developed a tabued scatter search method to solve RACP. Afshar-Nadjafi [1] proposed a mixed integer programming formulation in this problem. Peteghem and Vanhoucke [15] applied Invasive Weed Optimization (IWO) to solve the RACP with tardiness.

When the problem is extended to activity durations that are uncertain, few direct solutions have been introduced. Yamashita et al. [19] formulated the problem within the robust optimization framework and heuristic strategy is developed. Others transform it to a series of Resource-Constrained Project Scheduling Problem (RCPSP). Artigues et al. [2] developed and implemented a scenario-relaxation algorithm to produce high-quality solutions. Many evolutionary algorithms have been applied to solve the problem. Golenko-Ginzburg and Gonik [7] developed a heuristic scheduling procedure. Shukla et al. [17] applied particle swarm optimization (PSO). Other local search algorithms were combined with PSO as improvement, including attribute reduction [4], genetic algorithm [18] and path relinking [5]. However, the methods were indirect and require a large amount of computation.

This paper develops an optimization model and a new double-layer algorithm is applied. In order to improve the efficiency, the outer schedules the execution of activities while the inner optimizes the resource allocation scheme.

2 Model

Take all the maintenance activities in the combat unit into consideration. Let c_k denote the cost of kth resource, and the amount allocated for kth resource is x_k. The number of total activities is N, the start time of the ith activity is s_i, the duration is T_i, the amount of resource required for the ith activity is $r_{k,i}$. The project must be finished within D_m with the probability more than θ_m. The model is as follows.

$$\min\left(\mathrm{f}(\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_2) = \sum_{\mathrm{k}=1}^{\mathrm{K}} \mathrm{c}_\mathrm{k}\mathrm{x}_\mathrm{k}\right)$$

S.T.

$$\begin{cases} s_j - s_i - T_i \geq 0 & (1) \\ x_k - \sum_{i \in A_i} r_{k,i} \geq 0, k \in R^+ & (2) \\ P(D - s_{N+1}) - \theta \geq 0 & (3) \\ T_i \sim f_i(t) & (4) \end{cases}$$

Equation 1 indicates that the ith activity is a predecessor of the jth activity. Equation 2 ensures that the resource occupied cannot overrun the resource available. Equation 3 shows that the project completion probability constraint. Equation 4 denotes the durations follow certain random distributions, which depends on the particular problem (Take normal distribution as an example, then $T_i \sim N(\mu_i, \sigma_i^2)$, where μ_i and σ_i^2 are the mean and variance of the ith activity.)

3 Solution Algorithm

3.1 Solution Representation and Decoding Schemes

Popular representation methods include serial schedule generation [6], limited constraint [14], activity list [11], random key [15] and shift vector [10]. Considering the uncertainty, random key is used in this paper.

Suppose the set of activities executing is Y^*, activities awaiting execution are H^*. The decoding process is to update Y^* and H^* with the constraints, until the operational support project is finished. The detail of decoding procedure is in Table 1.

Table 1 Pseudo code for decoding procedure

Set $Y^* = \emptyset$, $H^* = \emptyset$, Mark activities without predecessor activity and add it to set H^*
While($H^* \neq \emptyset$)
Traverse set Y^*, calculate the earliest finish time of activities in the set t_f; $t = t_f$;
Delete activity whose finish time is t_f from set $Y^*$and the release the resource;
B_*Begin*=true;
While($H^* \neq \emptyset$and B_*Begin*=true)
Traverse set H^*,Find activity j which has the largest key and the resource is free;
if(activity j exists and the resource requirement can be satisfied)
Delete activity j from set H^*;Add activity j to set Y^*, occupy the corresponding resource; $s_j = t$;Add the subsequence activity of activity j to set H^*;
else
B_*Begin*=false
End while
End while, Output the finish time of support project

3.2 Schedule Optimization Procedure

The schedule optimization procedure is the outer layer of the whole solution algorithm. Scatter search embedded Particle swarm optimization is applied in this paper as the schedule optimization procedure. Particle swarm optimization (PSO) was put forward originally by Kennedy and Eberhart in 1995 [9]. To improve the efficiency, scatter search is embedded in PSO. The algorithm framework is in Figs. 1 and 2. The sub procedures can be described as follows.

Step 1 Calculate the latest start time of activities. Generate the initial resource allocation scheme and the initial schedule solution;

Step 2 Allocate resources to the solutions above using resource allocation procedure and calculate the objective function. Build the reference set;

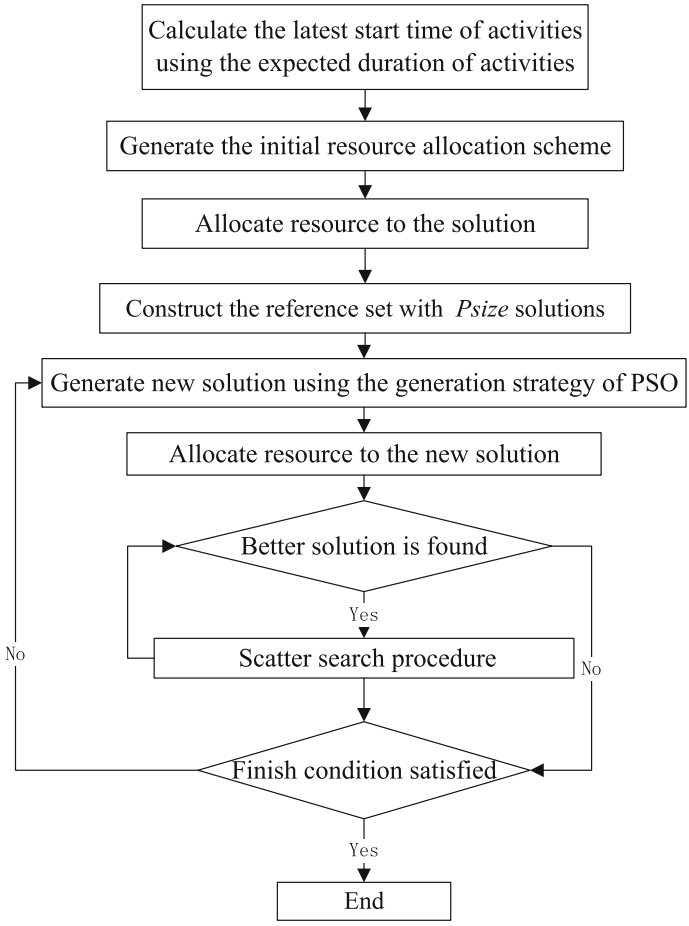


Fig. 1 Schedule optimization procedures

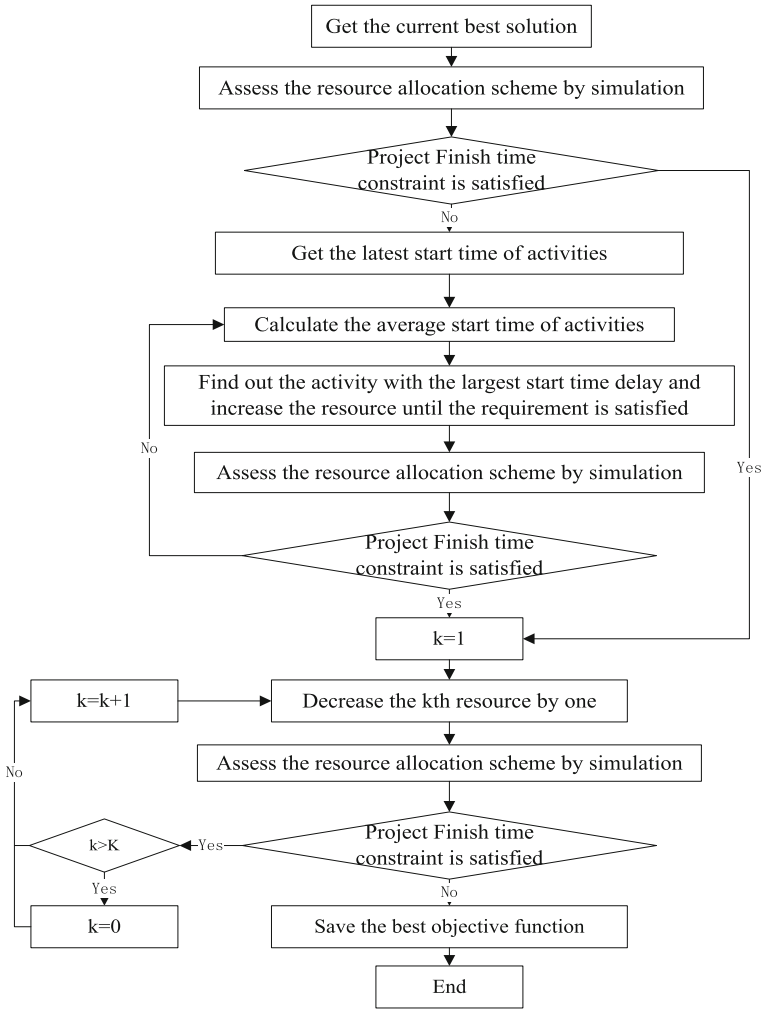


Fig. 2 Resource allocation procedures

Step 3 Generate new solutions using the generation strategy of PSO, calculate the objective function of new solution; If new current best solution is found, go to Step 4, else go to Step 5;

Step 4 Find better solution with scatter search procedure, go to Step 4. Suppose the number of solutions in reference set is P_{size}, the solution vector in reference set is X_i, the current best solution is X_b, the executing solution is X_e, the pseudo code of scatter search procedure is shown in Table 2.

Step 5 If the termination condition is satisfied, save the current best solution and the resource allocation scheme, otherwise go to Step 3.

Table 2 Pseudo code for scatter search procedure

```
For each j from j=2 to Psize
    X_e = X_j;
    For each i from i=1 to N
        if(X_b[i] ≠ X_e[i])
        Apply crossover on the ith position of X_b and X_e;
        Generate new solution New_1and New_2;
        Calculate the objective function of new solutions;
        End if
        if(New solution is better than the current best solution)
        Update the best solution, j=Psize;
        End if
        if(New solution is better than the worst solution in reference set)
        Update the reference set, replace the worst reference solution;
        End if
    End for
End for
```

3.3 *Resource Allocation Procedure*

Base on the random key coding and decoding method, resource allocation method should be applied on the current solution. Resource allocation is designed as a two-stage procedure. The detailed procedure is as in Fig. 3. The detail of sub procedures can be described as follows.

Step 1 Get the current best solution and assess the solution by simulation. If the success probability constraint is satisfied, go to Step 3, otherwise go to Step 2;

Step 2 Get the latest start time of support activities from the outer-layer algorithm and calculate the average start time of support activities. Find out the activity with the largest start time delay and increase the resource until the requirement is

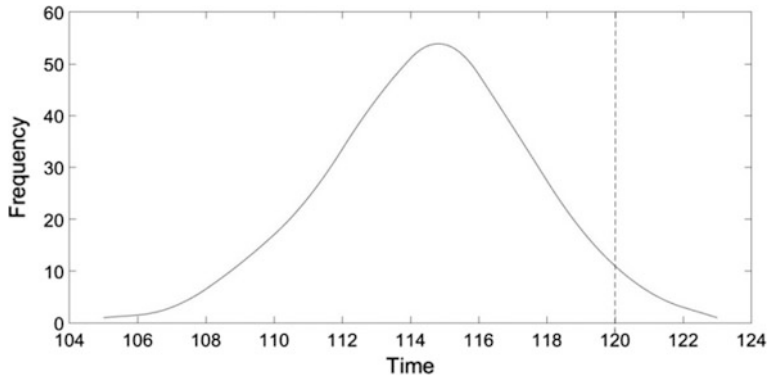


Fig. 3 Frequency distribution of project execution time

satisfied. Assess the solution again. If the project accomplishment probability constraint is satisfied, go to Step 3, otherwise repeat Step 2;

Step 3 Decrease the resource allocated from the first activity one by one and class by class. Assess the solution of resource allocation schemes and get the corresponding project accomplishment probability. If the current scheme cannot satisfy the project accomplishment probability constraint, go to Step 4;

Step 4 Save the resource allocation scheme which can satisfy the project accomplishment probability constraint and has the minimum cost.

4 Computational Example

Suppose a case that a combat unit consists of 12 weapons of the same type. Among all the weapons there exists no order for maintenance. The maintenance support activities for each weapon are set as in Table 3. The cost for maintenance resource is in Table 4. In a typical mission, the maintenance support of the combat unit should be finished in 120 min with success probability more than 0.9.

Using the solution algorithms mentioned above with the parameters that $c_1 = c_2 = 2$ and $\omega = 2$. The improved PSO contains 50 swarms and iterates for 200 times. The resource allocation scheme is as in Table 5. The total cost is 212, with the mission success probability of 0.948. The allocation scheme can ensure that at each moment the resource occupied is no more than what is allocated while the cost is the minimum. The allocated amount is not the sum of required as in Table 3, but it served as the constraint in the algorithm.

Figure 3 shows the frequency distribution of the project execution time with 200 instances, where the vertical dashed line denotes the project deadline. The efficiency of the algorithm is proved that the project finish time constraint can be satisfied successfully.

In order to evaluate the performance of the proposed algorithm, basic PSO and genetic algorithm (GA) are used as comparisons in this case. For GA, the initial population size is 50, and the mutation mechanism is applied to each solution with a

Table 3 The process of maintenance activity

Activity number	Activity duration distribution	Resource number (required amount)	Predecessor activity
1	N(4, 0.5)	1(4), 2(8)	–
2	N(5, 0.5)	2(2), 3(1), 4(1)	1
3	N(8, 1.5)	5(1), 6(1)	1
4	N(1, 7)	2(1), 3(1), 4(1)	2, 3
5	N(1, 8)	5(1), 6(1)	4
6	N(3, 0.5)	5(1), 6(1)	5
7	N(4, 0.5)	5(1), 6(1)	6

Table 4 The cost of maintenance resource

Resource number	Unit price
1	4
2	2
3	1
4	10
5	30
6	20

Table 5 Resource allocation scheme

Resource number	Allocated amounts
1	4
2	12
3	2
4	2
5	3
6	3

0.01 probability. To evaluate the performance of algorithms, algorithm execution time (seconds), total cost, mission success probability and average project execution time are used as indicators. The experiment result is shown in Table 6.

It is found that the improved PSO performed better than basic PSO and GA as the outer layer algorithm in that it accelerates the optimization process to a higher level. All algorithms can find the best allocation scheme ever found. But for the project execution, improved PSO works better than other algorithms that in the simulation samples, it has the largest mission success probability and the shortest average project execution time.

Table 6 Performance of different algorithms

Algorithm	Algorithm execution time	Total cost	Mission success probability	Average project execution time
Improved PSO	239.0	212	0.948	114.8
Basic PSO	257.4	212	0.925	117.6
GA	328.5	212	0.928	115.5

5 Conclusion

A double layer algorithm is presented to solve the maintenance support resource allocation problem. Computational experiments are applied to test the efficiency and performance of the proposed algorithm. The results show that the performance of the proposed algorithm is efficient. The average project execution time is 114.59, while the project finish probability is 0.95 for the 200 instances in experiment above, which is pretty satisfying. The double layer heuristic strategy is significant that it is able to find better solution than basic PSO and GA, the execution time for the proposed double layer algorithm is shortest as well.

References

1. Afshar-Nadjafi B (2014) Using grasp for resource availability cost problem with time dependent resource cost. Econ Comput Econ Cybern Stud Res 48(1):201–215
2. Artigues C, Leus R, Nobibon FT (2013) Robust optimization for resource-constrained project scheduling with uncertain activity durations. Flex Serv Manuf J 25(1):175–205
3. Baoli A (2010) Genetic and simulated annealing algorithm and its application to equipment maintenance resource optimization. Fire Control Command Control (1):144–145, 149
4. Ding W, Wang J, Guan Z (2012) Cooperative extended rough attribute reduction algorithm based on improved pso. J Syst Eng Electron 23(1):160–166
5. Duarte A, Martí R, Gortazar F (2011) Path relinking for large-scale global optimization. Soft Comput 15(11):2257–2273
6. Ghoddousi P, Eshtehardian E (2013) Multi-mode resource-constrained discrete time-cost-resource optimization in project scheduling using non-dominated sorting genetic algorithm. Autom Constr 30(30):216–227
7. Golenko-Ginzburg D, Gonik A (1997) Stochastic network project scheduling with non-consumable limited resources. Int J Prod Econ 48(1):29–37
8. Jing J (2013) Optimization techniques for equipment maintenance support resources in China: a literature review. Chin J Ship Res 8(4):116–121
9. Kennedy J, Eberhart R (1995) Particle swarm optimization. Proc IEEE Int Conf Neural Netw 4:1942–1948 (IEEE Xplore)
10. Lambrechts O, Demeulemeester E, Herroelen W (2008) A tabu search procedure for developing robust predictive project schedules. Int J Prod Econ 111(2):493–508
11. Lova A, Tormos MP, Cervantes M (2009) An efficient hybrid genetic algorithm for scheduling projects with resource constraints and multiple execution modes. Int J Prod Econ 117(2):302–316
12. Meng H, Wang B, Nie Y, Xia X, Zhang X (2016) A scatter search hybrid algorithm for resource availability cost problem. In: Harmony search algorithm. Springer, Berlin, Heidelberg
13. Möhring RH (1984) Minimizing costs of resource requirements in project networks subject to a fixed completion time. Oper Res 32(32):89–120
14. Özdamar L (1999) A genetic algorithm approach to a general category project scheduling problem. IEEE Trans Syst Man Cybern Part C Appl Rev 29(1):44–59
15. Peteghem VV, Vanhoucke M (2010) A genetic algorithm for the preemptive and non-preemptive multi-mode resource-constrained project scheduling problem. Eur J Oper Res 201(2):409–418

16. Peteghem VV, Vanhoucke M (2015) Heuristic methods for the resource availability cost problem
17. Shukla SK, Son YJ, Tiwari MK (2008) Fuzzy-based adaptive sample-sort simulated annealing for resource-constrained project scheduling. Int J Adv Manuf Technol 36(9):982–995
18. Wu L, Wang Y, Zhou S (2010) Improved differential evolution algorithm for resource-constrained project scheduling problem. J Syst Eng Electron 21(5):798–805
19. Yamashita DS, Armentano VA, Laguna M (2006) Scatter search for project scheduling with resource availability cost. Eur J Oper Res 169(2):623–637
20. Zhang W (2010) Optimal selection method of maintenance resources for weapon equipment. Mod Electron Tech 33(20):136–139

Bridge Management Integrating Big Data of Structural Health Monitoring

Yunxia Xia and Chunwei Zhang

Abstract In bridge condition assessment, structural health monitoring (SHM) data can reduce the level of uncertainty in the operational loadings and structural responses, which increases the reliability of the evaluation, and the effectiveness of the management activities. Therefore, it has a potential to decrease the life-cycle cost of bridges. To realize the economic benefit of the SHM systems, a bridge management system (BMS) integrating big data of SHM collected continuously from the bridges is proposed in this paper. The BMS includes four modules: structural inventory, inspection/SHM/finite-element model (FEM), bridge condition assessment, and maintenance decision. The SHM data is utilized in module of bridge condition assessment, which derives a three-dimensional bridge condition rating system. The proposed methodology is expected to assist priority ranking of bridge management activities throughout the bridge's life time.

1 Introduction

Bridges deserve special concern not only because of their vast investment, but also because of the cost of the traffic disruptions and even the catastrophic casualties due to their failure. However, by their very nature bridges are consistently subjected to traffic volumes and heavy truck loads as well as environmental stressors such as temperature, humidity, scour and chemical attack. Therefore, they are deteriorating under the combined actions of the above factors as well as the material ageing. To manage bridges effectively and economically, a management system is usually established.

Almost all the bridge management systems (BMSs) worldwide are based on visual inspections. However, the limitations of the visual inspection have long been recognized. For example, many of the early signs of deterioration or damage may be missed because of the ability of human eyes; the inspection results are subjec-

Y. Xia (✉) · C. Zhang
Qingdao University of Technology, Shandong, China
e-mail: xiayunxia@qut.edu.cn

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_73

tive, take a long time, and high manpower and cost are required. Inaccurate condition assessment has been recognized as the most critical barrier to the effective management of bridges.

Structural health monitoring (SHM) systems achieve site-specific data of operational loads and structural responses through a variety of sensors and sophisticated data acquisition and transmission systems [1, 2, 4, 5, 7]. The SHM data can reduce the uncertainty in the loadings and structural response, which increases the reliability of structural evaluation and management activities. As a result, the SHM technology has a potential to decrease the life-cycle cost of bridges. With a large amount of SHM data in hand now, the greatest and most urgent challenge in fulfilling the promise of SHM technology is how to integrate the big data into the BMS.

Stimulated by the challenging problems stated above, this paper proposes a framework to integrate the big data of SHM into the BMS. An SHM-integrated BMS including four modules: structural inventory, inspection/SHM/finite-element model (FEM), bridge condition assessment, and maintenance decision, is proposed first. The SHM data is taken into account in the module of bridge condition assessment that results in a three-dimensional bridge condition rating (BCR) system. Thus, it would play an important role in the maintenance decision and fulfil its economic promise.

The remainder of this paper is organized as follows. Section 2 describes the BMS incorporating SHM data. The procedure to implement the SHM-based BMS is illustrated in Sect. 3. Finally, Sect. 4 gives the conclusions.

2 Bridge Management System Integrating SHM Data

After reviewing the BMSs worldwide, a BMS composed of four modules: structural inventory, inspection/SHM/FEM, bridge condition assessment, and maintenance decision, as shown in Fig. 1, is proposed to integrate the SHM data into BMS. The systematic categorization of structural components is a fundamental element. The systematic categorization as well as the general information, including the location, geometric and material properties of the structural components are stored in the structural inventory. A well-devised structural inventory will minimize the time of searching, and speed up the efficiency of the bridge management activities.

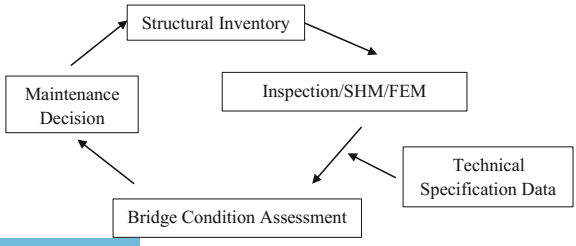


Fig. 1 Modules of SHM-based bridge management system

Providing a wealth of high-quality in situ data, SHM can effectively complement the traditional bridge condition assessment methods. The compelling advantages of SHM technologies include but not limited to updating existing models for the structural resistance and load effect, and showing performance variations over time and warning against threshold levels. It makes the bridge condition assessment more efficiently and simple. However, limited by practical problems such as budget and difficulty of deployment, only a necessary number of sensors are installed at discrete points of a bridge. Besides, though great progress has been made, it is still difficult for large-scale bridges to identify the damage or to determine the structural deterioration based on the SHM data. Therefore, it seems that the SHM data alone is not enough to evaluate the bridge condition.

A full 3D-FEM can reliably simulate the global and local structural characteristics of a bridge; thus, it is a valuable tool to provide structural information for bridge condition assessment. Although the visual inspection has inherent drawbacks, it has capabilities which SHM may not possess. For instance, some initial damage such as stripping of painting, spalling of concrete, visible cracks or wear, which cannot be sensed by an SHM system, may be discovered in visual inspections. Information such as the as-built report and maintenance record of the bridge is also essential in bridge condition assessment. Therefore, a comprehensive BCR system incorporating all the information from SHM, FEM, inspection observations, and technical specification data, is developed to provide more reliable bridge condition evaluation and more accurate baseline for bridge maintenance decision. Finally, bridge management decision can be optimized. By this means, the SHM data is integrated into the BMS.

3 Framework for Mining Big SHM Data in Bridge Management

The designed framework to exploit big data of SHM for bridge management is shown in Fig. 2. The structural reliability theory accommodates uncertainties in load effects and structural resistance, just as what described by the SHM data, to evaluate the structural serviceability or safety. Therefore, it may be the best tool to integrate the big SHM data into bridge condition assessment. Extreme value distributions (EVD) of the load effects obtained from the long-term SHM data are used to evaluate the structural reliability [8]. For those structural components without sensors, the structural reliability can be estimated with the help of the FEM. To get more accurate estimation, the site-specific live load model is desired. Subsequently, bridge condition is comprehensively evaluated and predicted based on the data-driven structural reliability. Based on this, the BCR can be conducted incorporating the FEM, inspection observations and other information such as the as-built report and maintenance records of the bridge.

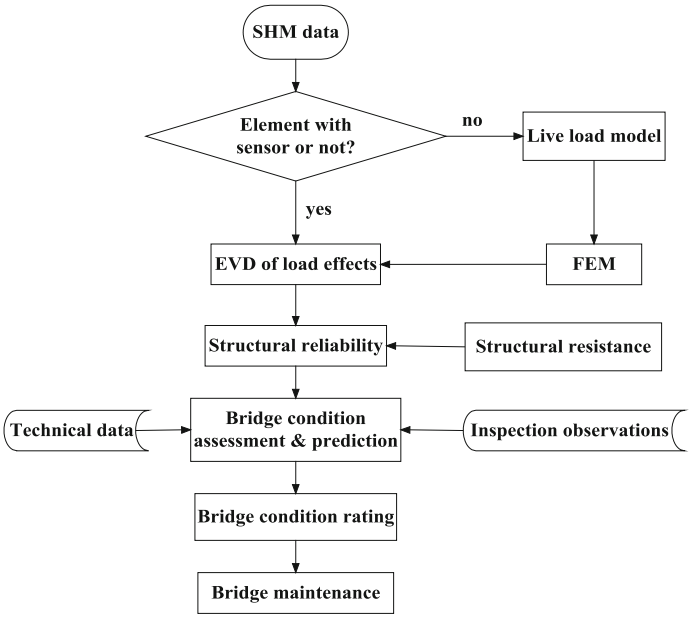


Fig. 2 Procedure for integration of big SHM data into BMS

3.1 Extreme Value Distributions of Load Demands

Among the tail fitting approaches based on extreme value theory (EVT), two are particularly popular: block maxima (BM), and peaks-over-threshold (POT) [3]. These two approaches have their own pros and cons. The BM approach has the advantage of time referencing, which is necessary in calculating probabilities of exceedance during the lifetime of a structure. However, only the maximum values in given blocks of time are taken into account in this approach. Even if several very large values were recorded, only one value in each time-block is considered, so a lot of useful data might be wasted. The POT approach accounting for the peaks which exceed a specified threshold can address the above issue, but its time reference is not as clear as the BM approach. The BM and POT data are in general fitted to the generalized extreme value (GEV) distribution and the generalized Pareto (GP) distribution, respectively.

The BM approach focuses on the statistical behavior of $M_n = \max(X_1, X_2, \ldots, X_n)$, where $(X_1, X_2, \ldots, X_n)$ is a sequence of independent random variables with a common distribution function F. The non-degenerate distribution function G is a distribution of the GEV family:

$$G(z) = \exp\left\{-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi}\right\} \tag{1}$$

defined on z such that $1 + \xi(z - \mu)/\sigma > 0$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$. The three parameters μ, σ, and ξ are location, scale and shape parameters, respectively. The cases $\xi = 0$, $\xi > 0$, and $\xi < 0$ are named the extreme value distribution (EVD) with types *I*, *II* and *III*, which are also widely known as Gumbel, Fréchet and Weibull families, respectively.

The POT approach studies those data that surpass a threshold level u. For a u large enough, when $M_n = \max\{X_1, X_2, \ldots, X_n\}$ satisfying the conditions to be approximated by a GEV, the distribution function of $(X - u)$ conditional on $X > u$, is approximately

$$H(y) = 1 - (1 + \xi y/\tilde{\sigma})^{-1/\xi} \tag{2}$$

defined on $y/y > 0$ and $(1 + \xi y/\tilde{\sigma}) > 0$, where

$$\tilde{\sigma} = \sigma + \xi(u - \mu) \tag{3}$$

The family of distributions defined by Eq. 2 is called the GP distribution. This theorem implies that if the block maxima of a dataset have an approximating GEV distribution, the threshold excesses can be fitted to a GP distribution with parameters associated with the GEV distribution. Particularly, the shape parameter ξ in Eq. 2 is equal to that of the corresponding GEV distribution.

3.2 Time-Dependent Structural Reliability

The load and structural resistance models contain many uncertainties. As a result, bridge condition assessment is an estimation, rather than an exact conclusion [6]. The structural reliability theory is perhaps the best tool to integrate the long-term SHM data into bridge condition assessment. The reason is that it has a capability to accommodate the uncertainties in load effects and resistance-related parameters, as reflected in long-term SHM data. Bridges deteriorate over time under aggressive environment and live loads, which leads to a reduction of the structural reliability. On the other hand, prior service loads reduce the uncertainty associated with bridge resistance and increase the reliability. These two contrary factors vary with time and may negate each other, so the reliability of a bridge is time-dependent. SHM data provide information about the bridge resistance, and continuously prove that the resistance is higher than the maximum load experienced already. In addition, SHM data provide useful information of site-specific loads, load effects and even resistance models. Consequently, the structural reliability of the bridge can be updated continuously.

3.3 Site-Specific Live Load Model

SHM data provide important information to develop live load models for bridge evaluation. For example, the weigh-in-motion (WIM) stations provide data such as the gross weight and axle distributions of the vehicles, and proportion of different vehicles on each lane. The CCTV shows the view of traffic situations on the deck, from which the distance between vehicles can be roughly judged. Data measured from strain gauges can be used to derive the information of traffic. With the live load models, the load-carrying capacity of those bridge components without sensors installed on can be analysed with the help of the full 3D FEM of the bridge. The detailed procedure for the development of the site-specific live load model can be found in [10].

3.4 Bridge Condition Rating System

To make it convenient for the bridge engineers and decision-makers to determine the maintenance priority and strategy, results of bridge condition assessment are preferred to be represented by indices in certain metric scales. A relatively accurate and comprehensive BCR system can help bridge managers to make effective budget allocation for bridge maintenance. To incorporate the SHM data into the bridge management system and provide guidelines for bridge maintenance and inspection, a three-dimensional BCR system was developed by the first author [9]. This system includes criticality rating (CR), vulnerability rating (VR), and inspection-based rating (IR). The criticality rating estimates the strength surplus of the structural components and the consequence of their failure, based on the SHM data and the FEM of the bridge. The live load models obtained based on long-term SHM data are applied to the FEM to analyse the structural performance. The vulnerability rating estimates the exposure of the structural components to the adverse effects, and the likelihood of detection in visual inspections. The as-built report and maintenance record are important information for vulnerability rating. The inspection-based rating assesses the current deterioration, crack and wear of the structural components according to the observations of inspections.

The results of CR, VR, and IR are transformed into percentage which ranges from 0 to 100%. Then the percentages are put into a 3D coordinate system: CR% as X-axis, VR% as Y-axis, and IR% as Z-axis, to form a rating cube. Setting 50% as the border, the cube is divided into eight quadrants: Quadrant 1 (Q1) to Quadrant 8 (Q8). This eight-quadrant cube is used as a guidance in the bridge management system to determine the priorities of the maintenance activities of bridge components. In this cube, CR takes the highest priority, and VR takes the lowest priority. Therefore, in the scheduling of inspection and maintenance for bridge components, the priority sequence is in the order of Q1–Q8.

4 Conclusions

To fulfill the economic potential of the long-term SHM systems, a framework to integrate big data of SHM into the bridge management is studied in this paper. A BMS composed of four modules: structural inventory, inspection/SHM/FEM, condition assessment, and maintenance optimization, is proposed to take advantages of the SHM data. The SHM data is utilized in the modules of condition assessment and prediction, which derives a three-dimensional BCR system. This make the SHM data play an important role in the bridge maintenance decisions and optimization. The management decisions made based on the BCR results should be a trade-off between structural safety and economy.

References

1. Abe M, Fujino Y (2009) Bridge monitoring in Japan. In: Encyclopedia of structural health monitoring. Wiley Online Library
2. Brownjohn JM (2007) Structural health monitoring of civil infrastructure. Philos Trans Royal Soc Lond A Math Phys Eng Sci 365(1851):589–622
3. Coles S (2001) An introduction to statistical modeling of extreme values. Springer
4. Ou JP, Li H (2009) Structural health monitoring research in China: trends and applications. In: Structural Health Monitoring of Civil Infrastructure Systems. Woodhead Publishing, Cambridge
5. Pines D, Aktan AE (2002) Status of structural health monitoring of long-span bridges in the United States. Prog Struct Mat Eng 4(4):372–380
6. Rakoczy AM, Nowak AS (2014) Reliability-based strength limit state for steel railway bridges. Struct Infrastruct Eng 10(9):1248–1261
7. Wong KY (2004) Instrumentation and health monitoring of cable-supported bridges. Struct Control Health Monit 11(2):91–124
8. Xia YX, Ni YQ (2016) Extrapolation of extreme traffic load effects on bridges based on long-term SHM data. Smart Struct Syst 17(6):995–1015
9. Xia YX, Ni YQ, Wong KY (2013) Development of a 3D bridge rating system incorporating structural health monitoring data. In: Proceedings of the 6th international conference on structural health monitoring of intelligent infrastructure (SHMII) (CD-ROM), Hong Kong, 9–11 December
10. Xia YX (2017) Integration of long-term SHM data into bridge condition assessment. Ph.D. thesis, Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong

A Modified Sideband Energy Ratio for Fault Detection of Planetary Gearboxes

Mian Zhang, Dongdong Wei, Kesheng Wang and Ming J. Zuo

Abstract The sideband energy ratio (SER) method has been proved effective for fixed-shaft gearbox fault detection. In this paper, we utilize a reported model of the vibration signal from a planetary gearbox and derive a modified sideband energy ratio (MSER) indicator for fault detection of planetary gearboxes. In addition, recognizing that measured speed variations are inevitable even a constant speed is set at a constant level, we have considered a bandwidth of frequencies instead of a single frequency when the MSER is computed. Lab experimental data analysis demonstrates that the proposed MSER is effective in separating the health conditions of healthy, tooth crack, and tooth pitting of the sun gear of a planetary gearbox. Future research directions include other faults of planetary gearboxes and dynamic speed conditions.

1 Introduction

Gearboxes are widely used to transmit power and motion. The commonly used configurations of gearboxes include fixed shaft gearboxes and planetary gearboxes. Due to their unique physical characteristics, planetary gearboxes are extensively used in aerospace, automotive, and heavy industry applications such as helicopters, wind turbines, heavy trucks and mining equipment [1, 3, 6, 10].

M. Zhang · D. Wei · K. Wang · M. J. Zuo
University of Electronic Science and Technology of China, Chengdu, China
e-mail: zoommian@foxmail.com

D. Wei
e-mail: redone17@163.com

K. Wang
e-mail: keshengwang@uestc.edu.cn

M. J. Zuo (✉)
Department of Mechancal Engineering, University of Alberta, Edmonton, Canada
e-mail: ming.zuo@ualberta.ca

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_74

Planetary gearboxes have a compact mechanical structure and provide a large transmission ratio. Therefore, they are often used in high output power requirement transmission systems [2]. Due to the heavy load and harsh working environment, key components of planetary gearbox, such as gears, may experience damages such as crack, spalling, and pitting. These damages may lead to undesirable dynamic behaviours such as intensive vibration, serious performance reductions, and even major breakdowns [5].

Vibration monitoring for early fault detection of both fixed-shaft and planetary gearboxes has attracted considerable research interest. When such gearboxes work under relatively constant speed conditions, frequency domain methods are widely used for fault detection. Due to the complicated motion dynamics of many moving parts in planetary gearboxes, early fault detection for planetary gearboxes is much more challenging than for fixed-shaft gearboxes.

For a fixed-shaft gearbox in a wind turbine, Hanna [4] used the Sideband Energy Ratio (SER) for gear fault detection. The method was demonstrated to be able to not only detect the presence of gear damage but also pin point to the gear experiencing the damage. Pattabiraman [9] used SER to monitor gear damage progression in a fixed shaft gearbox in a wind turbine. Their case study also demonstrated that SER was a reliable defect detection parameter for tracking defect progression of a gear in fixed shaft gearboxes.

However, for planetary gearboxes, there are several problems when applying SER for fault detection. First of all, the selection of sidebands is not as straightforward. The complicated frequency spectrum has several sidebands around the central meshing frequency. For example, the planet rotating frequency and the carrier rotating frequency don't exist in fixed-shaft gearboxes. Second, the physical meaning of SER for the vibration signal collected from a planetary gearbox needs to be provided. Last, a method for bandwidth selection of the central frequency and the sideband frequencies need to be developed. Considering practical speed variations even under relatively constant speed conditions, it is necessary to consider the energy contained in a bandwidth rather than only at a single peak point.

This paper aims to provide a theoretical explanation of the SER method when applied for planetary gear systems. With the sideband selection of the sun gear fault frequency, together with its harmonics, the SER could be considered as a sum of fault amplitude modulation (AM) indexes. In order to examine the effect of bandwidth selection of the central frequency and the sideband frequency, experimental data under three fault types with rotation speed of 30 Hz are analysed. The results show that the modified SER (denoted by MSER) calculated with a well selected bandwidth is much better than with a single peak point. The schematic of MSER is given in Fig. 1.

The remainder of this paper is organized as follows. In Sect. 2, the SER method will be first interpreted in terms of the sun gear local fault frequency using a reported signal model of a planetary gearbox. The experimental study for validation of the modified SER method is given in Sect. 3. Summary and conclusions are given in Sect. 4.

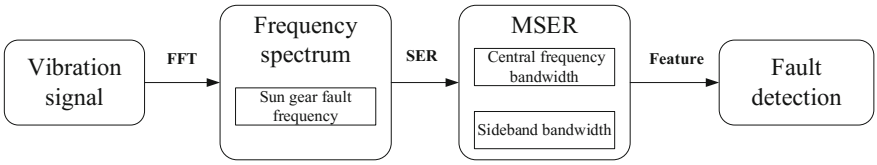


Fig. 1 The schematic of MSER for fault detection

2 Theoretical Analysis of SER

2.1 *Definition of SER*

The reported sideband energy ratio, that is, SER, is defined as the ratio between the sideband peaks and the gear mesh centre frequency peak. First six pairs of sidebands have been used in its calculation for fixed-shaft gearboxes. It has been found to be effective for gear damage detection for fixed-shaft gearboxes. In Eq. 1, SER is expressed as the ratio between the sum of the amplitudes of six sidebands and the amplitude of the centre mesh frequency [4].

$$\text{SER} = \frac{\sum_{i=1}^{6} \text{Sideband Ampitude}_{\text{i}}}{\text{Center mesh frequency amplitude}} \tag{1}$$

According to Eq. 1, the SER method sums the amplitudes of the first six sideband peaks on each side of the centre mesh frequency and divides this sum by the amplitude of the central mesh frequency. Before using this method for fault detection of planetary gear system, a theoretical analysis of this SER measure is done using a reported signal model for a planetary gearbox.

2.2 *Signal Model of Sun Gear Local Fault*

When there is a local sun gear fault, according to Feng [2], the vibration of a planetary gearbox consists of both amplitude modulation (AM) and frequency modulation (FM). The sun gear local fault causes modulation of the gear meshing vibration and the frequency of this modulation, in fact, gives the characteristic fault frequency of the sun gear. In addition, the gear meshing frequency will also be modulated by the sun gear rotational speed. As a result, vibration signals to be perceived by a sensor mounted on the gearbox casing can be modeled as [2]:

$$x(t) = \underbrace{[1 - \cos(2\pi f_{shaft} t)]}_{AMbyshaftrotation} \underbrace{[1 + A\cos(2\pi f_{sun} t + \phi)]}_{AMbysungearfaultfrequency} \cos[2\pi f_{mesh} t + \underbrace{B\sin(2\pi f_{sun} t + \varphi)}_{FMbysungearfaultfrequency} + \theta] \tag{2}$$

Table 1 Characteristic frequencies of a planetary gearbox

Frequency	f_{shaft}	f_{mesh}	f_{sun}	$f_{carrier}$
Expression	n	$\frac{Z_s Z_r}{Z_s+Z_r} * n$	$\frac{NZ_r}{Z_s+Z_r} * n$	$\frac{Z_s}{Z_s+Z_r} * n$

where f_{shaft} is the shaft rotational frequency, f_{sun} is the sun gear fault characteristic frequency, f_{mesh} is the gear meshing frequency, A and B are the modulation coefficients of AM and FM, respectively, and ϕ and φ are the initial phases of AM and FM, respectively. The relationships among the main frequency components of a planetary gearbox are given in Table 1, where $f_{carrier}$ is the carrier rotation frequency, N is the number of equally spaced planet gears, n is the sun gear rotation frequency, Z_s is the number of teeth of the sun gear, and Z_r is the number of teeth of the ring gear tooth.

According to Eq. 2, the fault frequency induced by sun gear fault will directly lead to the following sideband components: $f_{mesh} \pm f_{sun}$. During the engagement of the faulty tooth, the meshing excitation magnitude caused by different severity levels of the fault may be different. The amplitudes of the sidebands $f_{mesh} \pm f_{sun}$ in the frequency spectrum may be an effective fault severity indication. If the data is long enough, we don't believe that the phase factor will influence these amplitudes. We will not consider the FM part either in this paper. Considering the transfer path effect [7, 8] of the signal, the signal model given in Eq. 2 is simplified to the following in Eq. 3.

$$x(t) = A_0(1 - \cos\,2\pi f_{shaft}t)[1 + \beta_1\,\cos\,(2\pi f_{sun}t)]\,\cos\,(2\pi f_{mesh}t) \tag{3}$$

where A_0 is the AM index of the vibration signal incorporating the transfer path effect, and β_1 is the AM index of the sun gear fault.

Expanding Eq. 3 gives:

$$\begin{aligned} x(t) = {} & A_0\,\cos\,(2\pi f_{mesh}t) - A_0\,\cos\,(2\pi f_{shaft}t)\,\cos\,(2\pi f_{mesh}t) \\ & + A_0\beta_1\,\cos\,(2\pi f_{sun}t)\,\cos\,(2\pi f_{mesh}t) \\ & - A_0\beta_1\,\cos\,(2\pi f_{sun}t)\,\cos\,(2\pi f_{shaft}t)\,\cos\,(2\pi f_{mesh}t) \end{aligned} \tag{4}$$

Based on Eq. 4, the frequency contents of the resultant frequency spectrum will have $f_{mesh}, f_{mesh} \pm f_{sun}, f_{mesh} \pm f_{shaft}$ and $f_{mesh} \pm f_{shaft} \pm f_{sun}$. Note that for each of the resultant frequencies, the amplitude values depend on the two modulation indexes A_0 and β_1. The sideband components $f_{mesh} \pm f_{sun}$ only come from the part of Eq. 4 marked in a rectangle. This means that β_1 is of great importance for sun gear local fault detection. It will be further discussed below.

The central frequency as well as the sidebands are the key components in the SER method for fixed-shaft gearboxes. We now discuss them for a planetary

gearbox. Generally, the vibration induced by a gear fault may be reflected in a series of harmonic sidebands. According to the rectangle marked in Eq. 4 and considering the harmonic AM effects of f_{sun}, the vibration model can be modified as,

$$x(t) = A_0 \left[1 + \sum_{m=-\infty}^{+\infty} \beta_m \cos\left(2\pi m f_{sun} t\right) \right] \cos\left(2\pi f_{mesh} t\right) \tag{5}$$

For the sidebands of the sun gear fault frequency, according to Eq. 1, we can use Eq. 5 and modify the SER of Eq. 1 to the following,

$$MSER = \frac{\frac{1}{2}A_0 \sum_{m=1}^{6} \beta_m}{\frac{1}{2}A_0} = \sum_{m=1}^{6} \beta_m \tag{6}$$

The modified SER expression given in Eq. 6 establishes a relationship between the central frequency and the sun gear fault induced sidebands, eliminates the effect of A_0, and retains only the local fault sideband amplitudes, namely $\sum_{m=1}^{6} \beta_m$. For detection of the sun gear tooth fault, we can use the modified SER expression given in Eq. 6 under constant rotation speeds.

2.3 *Bandwidth Selection*

From the theoretical explanation given in Sect. 2.2, the MSER is expected to be a good feature for sun gear fault detection. Actually, another important factor should be noticed that, the bandwidth selection of the sideband and the central frequency is also important. Due to practical fluctuations of the rotation speed, though it is set at a constant value, the energy contained in a bandwidth for MSER application is expected to be better than the instant peak frequency value along. We will conduct bandwidth selection in later data analysis.

3 Experimental Studies

In the following, experimental data were acquired from a lab planetary gearbox test rig. The experimental set-up consisted of a spur gearbox and a one stage planetary gearbox driven by a 2.24 kW three phase electrical motor controlled by a motor speed controller. A vibration accelerometer (with sensitivity of 100 mV/g and frequency range 0–10 kHz) was used for capturing vibration data. Data lasting 4.8 s in time were collected for each controlled condition. The sampling frequency was set to be 10240 Hz. For each sun gear fault scenario, 10 groups of data were collected under the constant sun gear rotational frequency of 30 Hz. The test rig

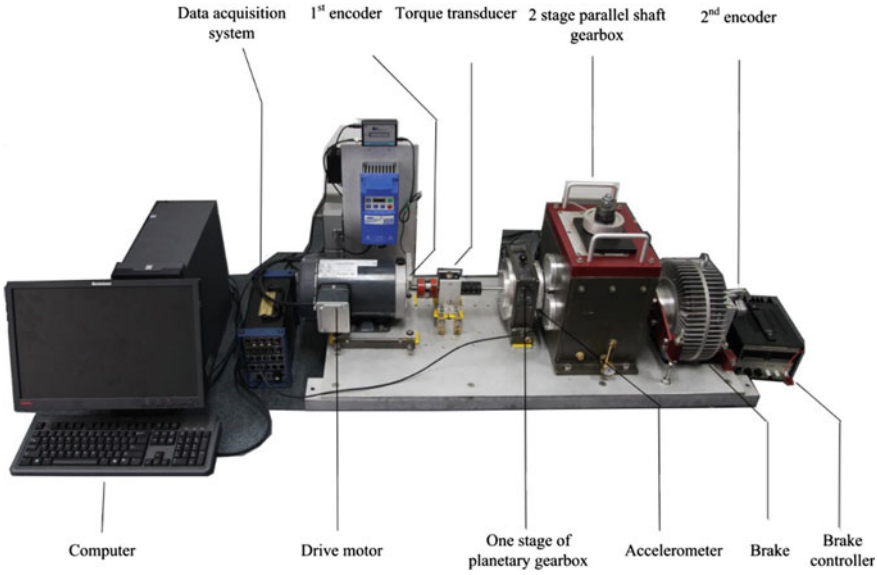


Fig. 2 Planetary gearbox test rig

Table 2 Test rig parameters

Item:	# of teeth (sun gear)	# of teeth (planet gear)	# of teeth (ring gear)	# of planets
Value:	28	36	100	4

configuration is shown in Fig. 2 and the key parameters of the planetary gearbox are listed in Table 2.

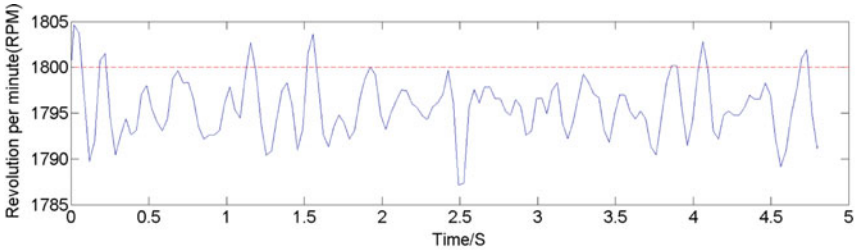
Three different sun gear health scenarios were considered for study. They were the healthy condition, the tooth pitting condition, and the tooth crack condition. Figure 3 illustrates these three health conditions.

(a) Healthy (b) Tooth pitting (c) Tooth crack

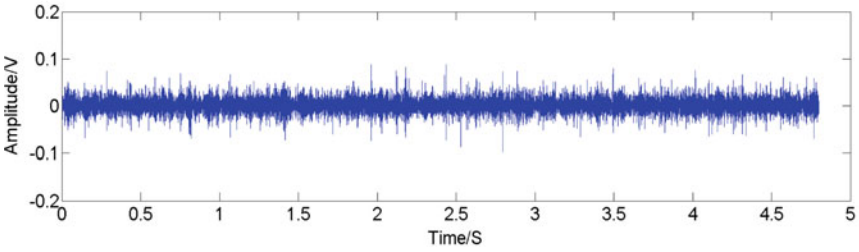
Fig. 3 The three considered health conditions of the sun gear

Take the healthy sun gear scenario when the controlled rotation speed was 30 Hz as an example. We have plotted the measured rotation speed, the time domain vibration data, and the frequency spectrum of the vibration data in Fig. 4. We can see clear variations in the measured rotating speed even the controller speed was set at 30 Hz (or 1800 RPM).

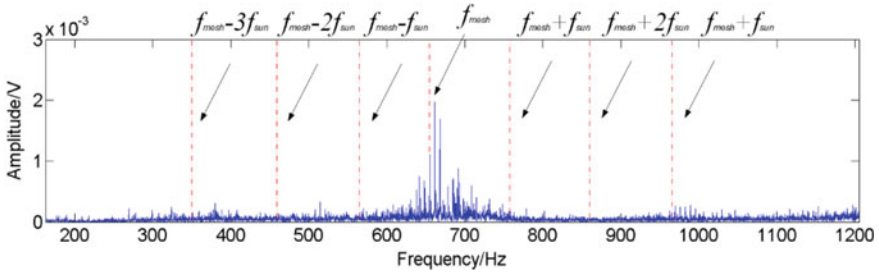
Due to the natural speed fluctuations that exists in the collected data, the bandwidth selected makes a difference in data analysis and fault detection. In Fig. 5, we have plotted the MSER calculated under two different bandwidth treatments, (a) when a bandwidth of 2 Hz was used around the central frequency and (b) when the single frequency identified at the peak amplitude was used. Data



(a) Rotation speed

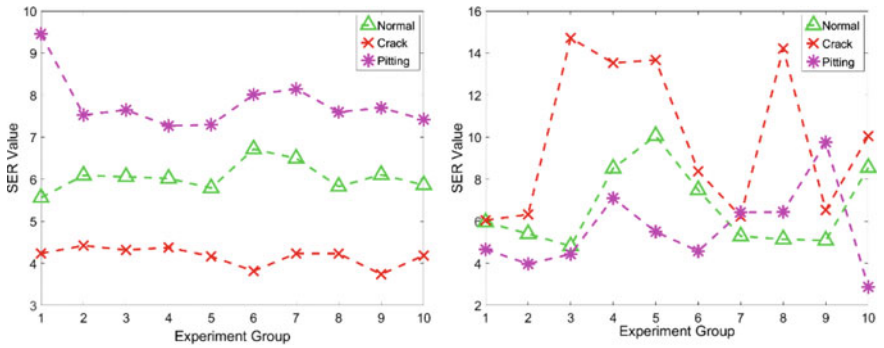


(b) Time domain vibration signal



(c) Frequency spectrum of the vibration data

Fig. 4 Measured rotation speed (**a**), time (**b**), and frequency domain signals (**c**) when the sun gear was healthy and the speed was set at 1800 RPM



(a) MSER calculated with a bandwidth (b) MSER with a single frequency

Fig. 5 Effect of bandwidth selection on MSER when rotation speed = 30 Hz

analysed were collected under a constant rotation speed of 30 Hz. The MSER values under all three health scenarios of the sun gear are presented in Fig. 5.

From Fig. 5, when the rotation speed is 30 Hz, the fault detection ability of MSER with a proper bandwidth setting (2 Hz) is much better than the MSER calculated with a single frequency at the peak. It can be seen that all three health scenarios are clearly separated in the plotted MSER value trends in Fig. 5(a). However, in Fig. 5(b), the three health scenarios are mixed together and thus are difficult to detect with the SER trend.

4 Summary and Conclusion

The original SER method, which has been proved effective for fixed-shaft gearboxes, is studied and modified in this paper. For planetary gearboxes, the modified SER indicator called MSER is developed utilizing a reported signal model of planetary gearbox signals. Added by the proper sideband selection, the proposed MSER is demonstrated to be an effective indicator of sun gear health condition. Future research includes consideration of other faults in planetary gearboxes and dynamic operating speed conditions.

Acknowledgements This research is supported by the National Key Research and Development Program of China (2016YFB1200401) and Natural Sciences and Engineering Research Council of Canada (Grant #RGPIN-2015-04897).

References

1. Feng K, Wang KS, Zhang M, et al (2016) A diagnostic signal selection scheme for planetary gearbox vibration monitoring under non-stationary operational conditions. Meas Sci Technol 28(3)
2. Feng Z, Zuo MJ (2012) Vibration signal models for fault diagnosis of planetary gearboxes. J Sound Vib 331(22):4919–4939
3. Guo Y, Parker RG (2011) Analytical determination of mesh phase relations in general compound planetary gears. Mech Mach Theory 46(12):1869–1887
4. Hanna J, Hatch C, Kalb M, et al (2011) Detection of wind turbine gear tooth defects using sideband energy ratio™. China Wind Power (Beijing, China, 19–21 Oct 2011)
5. Lei Y, Lin J, He Z et al (2012) A method based on multi-sensor data fusion for fault detection of planetary gearboxes. Sensors 12(2):2005–2017
6. Lei Y, Lin J, Zuo MJ, et al (2014) Condition monitoring and fault diagnosis of planetary gearboxes: a review. Measurement 48(1):292–305
7. Liang X, Zuo MJ, Hoseini MR (2015) Vibration signal modeling of a planetary gear set for tooth crack detection. Eng Fail Anal 48:185–200
8. Liu L, Liang X, Zuo MJ (2016) Vibration signal modeling of a planetary gear set with transmission path effect analysis. Measurement 85:20–31
9. Pattabiraman TR, Srinivasan K, Malarmohan K (2015) Assessment of sideband energy ratio technique in detection of wind turbine gear defects. Case Stud Mech Syst Signal Process 2015 (2):1–11
10. Singleton K (2006) Analysis of two stage planetary gearbox vibration. KSC Consulting LLC, Case Study

Use of Cyclostationarity to Detect Changes in Gear Surface Roughness Using Vibration Measurements

Xihao Zhang, Wade A. Smith, Pietro Borghesani, Zhongxiao Peng and Robert B. Randall

Abstract Wear in gears can usually be detected from the vibration signal once the wear has reached a 'macro' level—millimetre-scale variations from the original involute profile. Macro level wear is often preceded and accompanied by micro-level surface roughness changes (micrometre scale), arising from either abrasive wear or contact fatigue pitting. These micro- and macro-level phenomena interact with one another, and so knowledge of surface roughness is needed to be able to predict macro-level wear. It was recently suggested that it may be possible to use the cyclostationary properties of the vibration signal as an indicator of gear surface roughness, and the present paper examines this possibility further. It is thought that roughness information is carried by random high frequency vibrations that are modulated by the gearmesh cycle, and so any speed changes in the system should change both the carrier and modulating frequencies. The paper tests this hypothesis by studying laboratory measurements from a spur gearbox running at different speeds and with gears of different roughnesses. The findings will be very important for the further development of gear prognostics methods.

1 Introduction

Wear is one of the most common failure modes of gears, with abrasive wear and fatigue being the main wear mechanisms. Gear wear is a very complex process that is influenced by many factors, such as lubrication, temperature, operating conditions, material properties and tooth geometry. It has been established that vibration-based techniques can be used to detect macro-level wear (millimetre scale

X. Zhang · W. A. Smith (✉) · Z. Peng · R. B. Randall
School of Mechanical and Manufacturing Engineering, University of New South Wales (UNSW), Sydney, Australia
e-mail: wade.smith@unsw.edu.au

P. Borghesani
Queensland University of Technology, Brisbane, QLD, Australia
e-mail: p.borghesani@unsw.edu.au

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_75

variations from the original involute profile), through changes in the amplitudes of gearmesh harmonics, in particular the second harmonic [7, 9]. However, most of the techniques were developed for fault detection and diagnosis, rather than for prognosis and remaining useful life (RUL) prediction. Predicting RUL is the least developed aspect of gear condition monitoring, but it carries perhaps the greatest potential benefits, both economic and safety related. The wear and lifespan of gears is related to tooth surface roughness [6]; however, information about tooth surface roughness is very difficult to obtain without stopping the machine and taking detailed measurements. As such, establishing a relationship between surface roughness and vibration would provide a very valuable tool for the diagnosis of the gear wear state and for the calculation of remaining useful life.

It was recently proposed that such a relationship could be established through the use of cyclostationary (CS) signal analysis [10], and the present paper explores this possibility further. It is widely known that most of the power in gear vibration signals resides in the periodic components, such as the gearmesh frequency harmonics, but it is thought that important information is also carried in the second-order cyclostationary (CS2) components—in this case random vibrations that are modulated by the gear meshing process [4, 5]. One source of these random vibrations in gears is from the friction and asperity contacts that occur between the sliding surfaces of the meshing teeth, and it is thought that the strength of these vibrations would be closely related to tooth surface roughness. This paper investigates the relationship between gear surface roughness and cyclostationarity using data obtained from a laboratory spur gearbox test rig. In particular, the effect of running speed on the frequency distribution of the cyclostationary content is examined for a number of different surface roughness levels. This will allow for more targeted metrics of surface roughness to be developed in the future.

2 Gear Surface Roughness and Cyclostationarity

2.1 Cyclostationary Signals and Their Indicators

A signal is defined to be cyclostationary at the nth-order if its nth-order statistical properties are periodic with respect to time [4]. The most relevant case here is that of a second-order CS signal, which has a periodic autocorrelation function (or variance), a typical example of which is a cyclic repetition of random bursts. Note that CS2 signals are random signals, such that their cyclic structure (i.e. periodic variance) is not apparent in the ordinary frequency spectrum. For such cases, cyclostationary signal processing techniques must be applied to uncover the underlying cyclic nature of the signal.

The degree of cyclostationarity in a signal can be measured by established indicators [8]. The indicator of second-order cyclostationarity (ICS2) is defined as:

$$\text{ICS2} = \sum_{k \in Z, k \neq 0} \frac{\left| C_{2x}^{k\alpha_0}(0) \right|^2}{\left| C_{2x}^{0}(0) \right|^2} \tag{1}$$

where $C_{2x}^{0}(0)$ is the mean square power of the 'centred signal' x_c, obtained by subtracting the synchronous average from the raw signal. $C_{2x}^{k\alpha_0}$ is the second-order cyclic cumulant for the set of all cyclic frequencies α ($= k\alpha_0$). According to Raad et al. [8], a consistent estimator of $C_{2x}^{\alpha}(0)$ for a discrete signal $x(n)$ with length N is given by the components at frequencies α in its envelope spectrum, in this case using the squared signal as an approximation for the squared envelope:

$$C_{2x}^{\alpha}(0) \approx N^{-1} DFT\left\{x_c^2(n)\right\}(\alpha) \tag{2}$$

where $DFT\{x(n)\}(\alpha)$ stands for the N-point discrete Fourier transform of signal $x(n)$ calculated at frequencies α.

While ICS2 gives an indicator of the overall level of second-order CS content in a signal, it does not indicate its spectral frequency distribution. For this, one must employ more comprehensive tools such as the Spectral Correlation (SC), $S_x(\alpha,f)$, defined as the double discrete Fourier transform of the instantaneous autocorrelation function (itself a function of both time and time lag for CS signals) [2]. The Spectral Correlation indicates the power distribution of the signal with respect to both the spectral frequency f and the cyclic frequency α [1]. In this paper we employ a normalised version of the SC, the Spectral Coherence, defined as [2]:

$$\gamma_x(\alpha,f) = \frac{S_x(\alpha,f)}{\sqrt{S_x(0,f)S_x(0,f-\alpha)}} \tag{3}$$

where $S_x(0,f)$ represents the ordinary power spectral density at frequency f. I.e., the CS content at frequency f is normalised by the power at frequencies f and f–α in the stationary part of the signal. The Spectral Coherence can also be interpreted as the SC of the whitened signal, which tends to magnify weak cyclostationary signals [2].

2.2 *Effect of Gear Surface Roughness on Vibration Signals*

As explained in [10], the rationale for employing cyclostationary analysis in this context arises from the periodic variation in sliding velocity (and contact forces) inherent in meshing gears. The variation in the number of meshing tooth pairs and the cyclic trend in the sliding velocity of the meshing pairs as a function of gear rotation angle result in possible CS2 components generated from the random asperity contacts on the gear tooth faces. As shown in the same reference, a periodicity in the sliding velocity is produced corresponding to the gearmesh frequency. It is likely that the strength of the vibrations arising from random asperity contacts

would be closely tied to sliding velocity (and level of roughness), leading to the amplitude modulation effect experienced in [10]. Note that even though this vibration (the carrier signal) is random, with unknown frequency content, it can be separated from other parts of the signal using CS tools because the modulation signal is both known and periodic, with a cyclic frequency equal to gearmesh frequency.

3 Methodology

3.1 Experimental Setup

Data was collected from a number of tests conducted on the UNSW gearbox test rig, shown in Fig. 1. The gearbox consists of a single parallel gear stage and is powered by an induction motor connected to a variable frequency drive (VFD). A water pump is connected to the output shaft to apply a torque load, and a tacho (twice-per rev) is connected to the free end of the output shaft. The drive and driven gears are KHK mild steel spur gears, with 46 and 25 teeth, respectively.

The vibration signal was collected using a B&K 4370 accelerometer stud-mounted on the top of the gearbox casing. The signal was recorded using a National Instruments CompactDAQ with an NI-9234 module. The sampling frequency was 51 kHz.

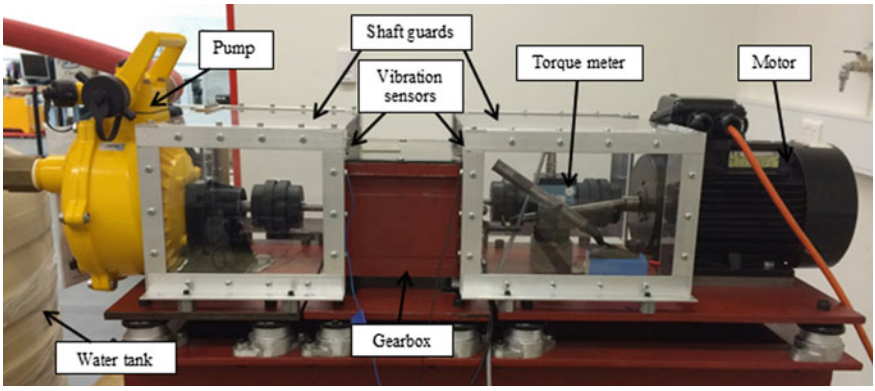


Fig. 1 UNSW gearbox rig

3.2 Test Program and Surface Roughness Measurement

Three sets of tests were conducted, each with different gear pairs. Before some tests the meshing surfaces of the gears were roughened manually (see below). Each test consisted of a long running period (88, 24 and 45 h for Tests 1, 2 and 3, respectively), in which the tooth surfaces were allowed to evolve (generally smoothen) naturally over time, and periodically the vibration signals were recorded and the tooth surfaces measured again to obtain the roughness level corresponding to each vibration signal. In total, 15 measurement points were used: two in Test 1 (points 1A and 1B), four in Test 2 (2A-2D) and nine in Test 3 (3A-3I). The rig was run with an input shaft speed f_{in} of 23 Hz and a torque load of 14 Nm, but at each measurement point signals were also recorded at 15 Hz/7 Nm.

Sandpaper (grit 120, 220 and 320) was used to roughen the gears at the start of Tests 2 and 3. (In Test 1 the gears were as-supplied.) The gears in Tests 1, 2 and 3 had initial surface roughnesses (R_a values) of 0.6, 1.5 and 3.2 µm, respectively, and these roughnesses reduced to 0.5, 1.1 and 1.5 µm over the duration of the tests.

Surface roughness was measured using a Mahr M1 Perthometer (cut-off wavelength 0.8 mm), with measurements taken along a number of teeth (randomly selected), and multiple measurements taken along the same tooth. The measured R_a values were then averaged to give the reported value. At the start of Tests 2 and 3, with manually roughened surfaces, a consistent roughness was achieved by ensuring that the R_a values of all measurements fell within ±0.1 µm of the mean value.

3.3 Signal Processing

As explained in Sect. 2.2, it is hypothesised here that surface roughness information would be carried in the random part of the vibration signal (and modulated by a periodic function associated with gearmesh), so order tracking and time synchronous averaging were applied to remove deterministic components synchronous with both shafts. The residual signal (obtained by subtracting the two TSA signals from the order-tracked signal) was then used to calculate ICS2 according to Eqs. (1) and (2), with the cyclic frequency set to the gearmesh frequency ($\alpha = f_{\mathrm{gm}}$). The residual signal was also used to calculate the Spectral Coherence (again at $\alpha = f_{\mathrm{gm}}$), as defined in Eq. (3), using a new fast computation code developed by Antoni et al. [2, 3].

4 Results and Discussion

4.1 Correlation of CS Indicator and Surface Roughness

Figure 2 shows the plots of ICS2 vs surface roughness for all the measurement points for both input shaft speeds (15 and 23 Hz). The data from Tests 1 and 2 in the high speed case was in fact presented in [10], where a very good correlation between ICS2 and roughness was observed. It was explained in that paper that point 2B was thought to be unreliable because looseness had developed in the rig, so that point was discarded in the regression analysis. The plots shown here, with a greater range of roughness levels and a new speed/load, suggest a more complex relationship. Certainly, the monotonic trend previously observed in the high speed case (right plot) seems to break down around R_a = 2 μm. It is possible that for the outlying measurement points at higher roughness levels (the first four measurements from Test 3) there is a fundamental difference in the interaction of the contacting surfaces. For example, a higher R_a value—obtained using coarser grit sandpaper—may mean a larger average wavelength in the surface asperities, and beyond a certain point this could lead to a reduction in the rate of asperity breakage and deformation, and hence a lower ICS2 value. However, this explanation is at odds with the fact that throughout Test 3 (45 h), the average surface roughness of the gears did change considerably (from R_a = 3.2 to 1.5 μm). Implicit in this point is that R_a alone is not an adequate metric to completely characterise a random surface, and this is one area that will be examined in future analyses on the cyclostationarity/roughness relationship.

In the case of low speed/load (left plot), there is no observable trend, even from the Test 1 and 2 data points. A likely reason for the poorer correlation in the low speed case is due to the very low torque load (7 Nm), which could not be controlled

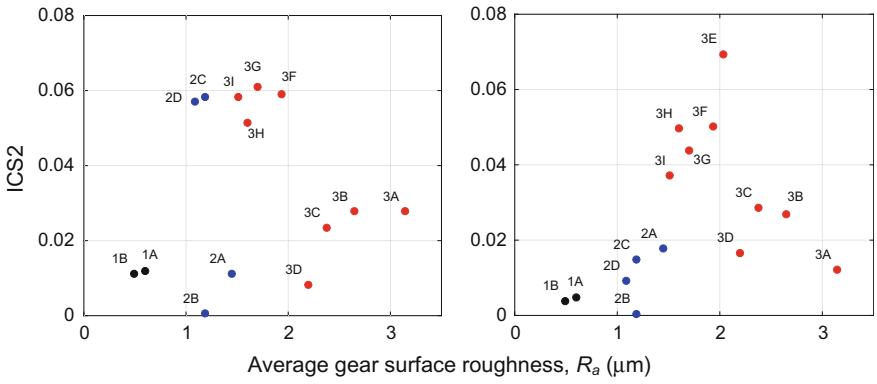


Fig. 2 ICS2 (calculated at $\alpha = f_{gm}$) versus gear surface roughness; left: f_{in} = 15 Hz; right: f_{in} = 23 Hz

independently of speed with the rig setup at that time. Future testing will address this issue.

4.2 *Effect of Running Speed on Cyclostationarity*

One way of improving on ICS2 as a potential roughness indicator is to develop a more targeted metric that examines the spectral frequency range carrying the most roughness information. This section contributes to that by studying the effect on CS content of the machine running speed. The rationale for using CS processing tools to examine surface roughness (see Fig. 1) was based on variations in sliding velocity between the contacting surfaces throughout the mesh cycle, and so it seems likely that running speed (which dictates the range of sliding velocities) would have a direct effect on the frequency of the CS content. This CS spectral content was investigated using the Spectral Coherence defined in Eq. (3), expressed as $|\gamma_x|^2$ to give values ranging from 0 to 1. Four measurement points were chosen for analysis —1B, 2A, 3A and 3F—representing a range of roughness and ICS2 levels. Figure 3 shows the Spectral Coherence for these points over the 0–25 kHz spectral frequency range. This represents the power of the signal content modulated at the gearmesh frequency normalised by the power in the stationary part of the signal. To give a clearer indication of the spectral distribution, the first order spectral moment (or 'mean frequency') for each coherence plot is shown as a dotted line on the graphs, and the results for the two speed cases are plotted on the same axes.

It is clear in every plot that the average frequency of the CS content increases with the running speed, almost in direct proportion, with the average change in mean frequency found to be about 1.4, while the speed ratio was around 1.5. That is, not only does the cyclic frequency change with running speed, but so does the range of dominant carrier frequencies. This indicates that the main carrier signals are indeed tied to sliding velocity and the rate of interaction of the asperities, and not, for example, fixed frequency resonances such as might be excited in the presence of a bearing fault. Note that this analysis is complicated by the fact that the machine speed only changes the range (or average) of the sliding velocity, i.e. the sliding velocity still spans a range from zero to the maximum.

To develop a more effective roughness indicator further work is needed to establish the frequency range over which most roughness information is carried, over a wider range of speeds, roughnesses and torque loads. However, the finding that the 'mean' carrier frequency seems closely related to sliding velocity helps in the development of more sophisticated tools, for example to isolate frequency modulation effects.

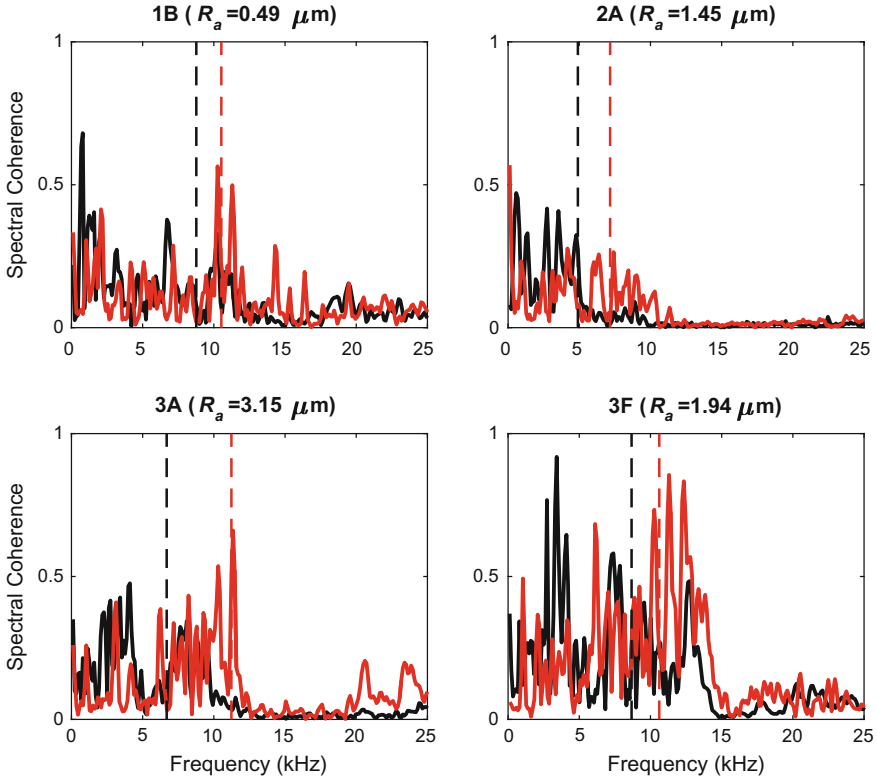


Fig. 3 Spectral coherence $\left|\gamma_x\left(\alpha = f_{gm}, f\right)\right|^2$ obtained at different operating speeds corresponding to measurement points 1B (top-left), 2A (top-right), 3A (bottom-left) and 3F (bottom-right). Black lines: f_{in} = 15 Hz; red lines: f_{in} = 23 Hz. Dashed lines represent first-order spectral moments

5 Conclusion

This paper investigated the possibility of using the cyclostationary (CS) properties of vibration signals to detect changes in gear tooth surface roughness. It was previously proposed that information about the surface roughness of meshing gears would be carried in the random vibrations generated from asperity impacts, and these vibrations would be modulated by the gear meshing cycle, which is entirely deterministic. This creates a second-order cyclostationary (CS2) signal, and a previous study found a strong correlation between CS2 content and surface roughness. The present paper, using new data covering a much larger range of roughness levels, found this relationship to be more complex than originally thought, with a negative correlation between these two variables observed beyond a certain roughness level (R_a = 2 μm). More work is needed to understand the physics underlying this observation.

The paper also used Spectral Coherence to study the frequency content of that part of the vibration signals modulated at the gearmesh frequency, and how it is affected by running speed. For the four measurement points studied, covering a range of surface roughnesses, it was found that the mean frequency of the coherence varied almost in direct proportion to the running speed, indicating that the main carrier signals in the CS2 content were based on sliding velocity, such as impacting asperities, and not on fixed frequency resonances. This provides useful information for the development of more targeted roughness metrics in the future.

Acknowledgements The authors would like to thank Mr. Yunzhe Yang and Dr. Chongqing Hu for their help in the experiments. This research was supported by the Australian Research Council, through Discovery Project [DP160103501].

References

1. Antoni J (2007) Cyclic spectral analysis in practice. Mech Syst Signal Process 21:597–630
2. Antoni J, Xin G, Hamzaoui N (2017) Fast computation of the spectral correlation. Mech Syst Signal Process 92:248–277
3. Antoni J (2016) Fast_SC program, mathworks file exchange. Accessed 3 March 2017. https://au.mathworks.com/matlabcentral/fileexchange/60561-fast-sc-x-nw-alpha-max-fs-opt-?focused=7143132&tab=function
4. Antoni J, Bonnardot F, Raad A, El Badaoui M (2004) Cyclostationary modelling of rotating machine vibration signals. Mech Syst Signal Process 18:1285–314
5. Capdessus C, Sidahmed M, Lacoume JL (2000) Cyclostationary processes: application in gear faults early diagnosis. Mech Syst Signal Process 14:371–385
6. Krantz T (2005) The influence of roughness on gear surface fatigue. NASA technical report NASA/TM-2005-213958
7. Mark WD (1978) Analysis of the vibratory excitation of gear systems: basic theory. J Acoust Soc Am 63:1409–1430
8. Raad A, Antoni J, Sidahmed M (2008) Indicators of cyclostationarity: theory and application to gear fault monitoring. Mech Syst Signal Process 22:574–587
9. Randall RB (1982) A new method of modeling gear faults. J Mech Des 104:259–267
10. Yang Y, Smith WA, Borghesani P, Peng Z, Randall RB (2015) Detecting changes in gear surface roughness using vibration signals. In: Acoustics 2015. Hunter Valley, Australia

A Data-Driven Decision Model: A Case Study on Drawworks in Offshore Oil & Gas Industry

Pengyu Zhu

Abstract Offshore installations are complex and need to be maintained properly to sustain expected performance. Critical failures on these installations could pose great threats to productivity, personnel safety, and the environment. The research is designed to suggest some practical solutions for improving decision quality and reliability. During operation and maintenance (O&M) activities, much data are collected, and it is believed that making full use of them has great potential for improving production efficiency, as well as for reducing risks. Drawworks is studied to elaborate a data-driven methodology. The research suggests some practices for identifying and using critical data sets as a driving force to improve decision-making. Both qualitative and quantitative analysis are used during the study. Competence management is also studied as a necessary part of the data-driven decision-making setting.

1 Introduction

Offshore oil & gas (O&G) operations are expensive and generally have higher safety requirements compared to onshore operations. Production safety and availability are largely influenced by the performances and conditions of key equipment of offshore installations. Human decisions play an important role in the process. However, in traditional terms, decisions related to operation and maintenance activities are often made based on experiences. The consistency and quality of experience-based decision practices are questionable [17].

It has been widely acknowledged that data & information have great potentials for improving the quality of decision-making [1, 4, 12]. With the development of sensor technology and digitalization, more and more data are collected and need to be integrated into decision making, but the definition of the most critical data sets, and the integration process are still heavily experience-based. The paper promotes a

P. Zhu (✉)
Cluster on Industrial Asset Management, University of Stavanger, Stavanger, Norway
e-mail: pengyu.zhu@uis.no

J. Mathew et al. (eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Lecture Notes in Mechanical Engineering, https://doi.org/10.1007/978-3-319-95711-1_76

normative approach to help decision-makers gain more control of data and explore the maximum value of them in a systematic way. A critical equipment in offshore drilling is selected and described in Sect. 2. In Sect. 3, the methodology of the data-driven approach is explained. In Sect. 4, the failure mechanism of the selected equipment is studied, and how to establish the contextual data architecture for decision-making is explained.

2 Drawworks

Faller [7] listed the most critical equipment for offshore drilling facility as including top drive (crown block), drawworks, and mud pumps. Drawworks was selected in this study as its criticality in offshore drilling was sometimes overlooked. An operator from North Sea recoded all maintenance activities related to the drawworks between 2004 and 2014. Records showed that 10% of total maintenance activities were corrective maintenance, while the rest was planned maintenance. From the oil operator's point of view, the drawworks was quite reliable. However, a drawwork incident several years back induced a production shutdown for more than two weeks, which caused great economic loss. Drawworks can, in many ways, be the reason for production loss or catastrophic failures that threaten the environment and human lives. Its possible impacts should not be neglected. A system of systems diagram of drawworks is depicted in Fig. 1.

Some essential functions that drawworks perform include: dragging the drill string/casing out of the hole; controlling the speed of drill string/casing; controlling the weight that is applied on the drilling bit; providing a power take-off for the chain-driven rotary table if no other hoisting equipment is installed [6, 14].

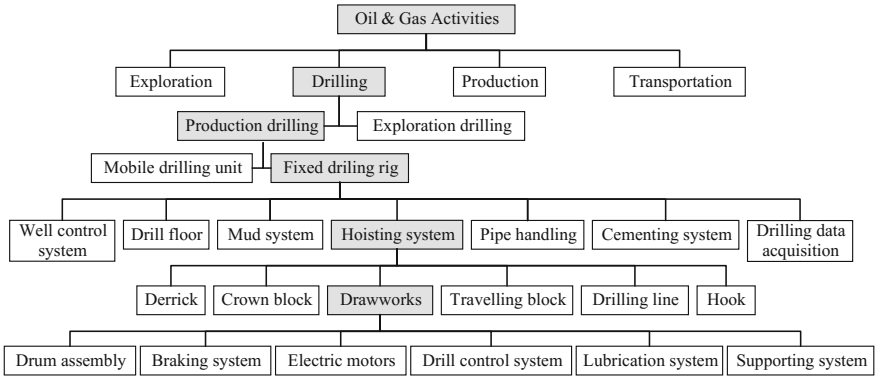


Fig. 1 System of systems. Adapted from [3, 15, 19]

3 Methodology

New sources of data are generated and collected with the development of advanced monitoring technologies and with the increasing requirement on safety. The paper promotes a normative data-driven approach to help improve the contextual awareness and assist decision-makers to define and make use of the right data in decision-making processes.

The study began with the identification of critical failure modes and mechanisms of drawworks that had the severest impacts on production, safety, or the environment. Failure modes and symptoms analysis (FMSA) and fault tree analysis (FTA) were implemented to study the failure logic and symptoms. This process helped define the critical parts/systems that the data architecture could be built around, and provided a reasoning logic for decision-making, both qualitatively and quantitatively. The definition and collection of various data sets and dissemination of data were explained with a case study. In the end, the paper studied the role and competence of personnel in the new decision setting.

4 Case Study Results and Analysis

4.1 Failure Modes and Symptoms of Drawworks

Failure modes and symptom analysis (FMSA) is a systematic risk analysis tool, where failure modes, causes, local effect, and system effect are identified, following the guidelines defined by ISO [11]. FMSA analysis was performed at the component level of drawworks, and more than 80 failure modes were analysed [19]. The highest-ranking failure modes are listed in Table 1.

In the analysis, shutdowns of the whole system or threats to human safety or environment were all considered unacceptable. The analysis used the monitoring priority number (MPN) as an indicator to help prioritize the monitoring sequence of failure modes. The formula for calculating MPN is [10]:

$$\text{MPN} = \text{DET} \times \text{SEV} \times \text{DGN} \times \text{PGN} \tag{1}$$

where, DET refers to probability of detection, SEV is severity of failure, DGN is diagnosis confidence, and PGN is prognosis confidence. The numbers for DET, SEV, DGN and PGN in the table are assigned subjectively from 1 to 5, which represents the degree of confidence from weak to strong. According to the definition of MPN, the higher its value is, the higher is its monitoring priority [10].

Table 1 Part of FMSA analysis.

FMSA of drawworks										
System	Failure symptoms	Failure mode	Cause of failure	Local effect	System effect	DET	SEV	DGN	PGN	MPN
Drum assembly	Breakage of wireline	Structure deficiency	Off-design service: fatigue, or over holding weight	Possible drop of load, failure of drawworks	Drop of load	4	5	4	4	320
Braking system	Fast wearing of brake discs	Load drop	Maintenance error	Insufficient braking	Insufficient braking, drop of load	4	5	4	4	320
Braking system	Breakage	Structure deficiency	Off-design service: fatigue, or over holding weight	Possible drop of load, faliure of drawworks	Drop of load	4	5	4	4	320
Braking system	Brake power is not enough	Failure to stop on demand	Operation error: disconnection, loose items	Insufficient braking	Insufficient braking, drop of load	4	5	4	4	320
Braking system	Brake functions improperly	Failure to stop on demand	Maintenance error: pneumatic system error	Unexpected braking, failure to brake	Delayed drilling, drop of load	4	5	4	4	320
Electric motors	Unstable power output	Erratic output	Maintenance error	Decreased capacity of power output	Threat of ho isting/ lowering safety	4	5	4	4	320

(continued)

Table 1 (continued)

FMSA of drawworks										
System	Failure symptoms	Failure mode	Cause of failure	Local effect	System effect	DET	SEV	DGN	PGN	MPN
Electric motors	Unstable power output	High output	Management error	Higher speed than expected	Threat of hoisting/ lowering safety	4	5	4	4	320
Drill control system	Switch failure	Failure to function on demand	Operating error: short/open circuit, or earth isolation problem	Loss of control of drawworks drilline/ braking system	Emergency braking triggered unexpected during ho isting/ lowering operation	4	5	4	4	320
Drill control system	Failure to respond on command	Failure to function on demand	Operating error	Loss of control of drawworks drilline/ braking system	Drop of load/ travelling block	4	5	4	4	320
Lubrication system	No rise in pressure	Internal leakage	Maintenance error	Loss of function of pumping	Stop of production due to automatic utilization of emergency braking	4	5	4	4	320

Adapted from Zhu [19]

4.2 Fault Tree and Reliability

In this paper, both qualitative and quantitative reliability analysis were needed for decision-making. Due to the limited reliability data from OREDA [16], electric motor was used to illustrate the process. There are two redundant electric motors in drawworks. The fault tree analysis (FTA) is depicted in Fig. 2.

Reliability and sensitivity analysis was implemented based on the FTA. Reliability data of electric motor was found in OREDA [16]. Failures of different parts was assumed to be exponentially distributed with parameter λ. Mean time to failure (MTTF) equals $1/\lambda$. The reliability of component i is calculated:

$$p_i = \frac{MTTF_i}{MTTR_i + MTTF_i} = \frac{MTTF_i}{MTTR_i + (1/\lambda_i)} \tag{2}$$

Terje Aven [2] suggested that Birnbaum's measurement could be appropriate in the context of oil production, as small changes in operation and maintenance performance might induce a large change in system reliability. The use of Birnbaum's measurement has been elaborated by some researchers [9, 18]. The importance of component i is calculated [5]:

$$I_i^B = h(1_i, p) - h(0_i, p) \tag{3}$$

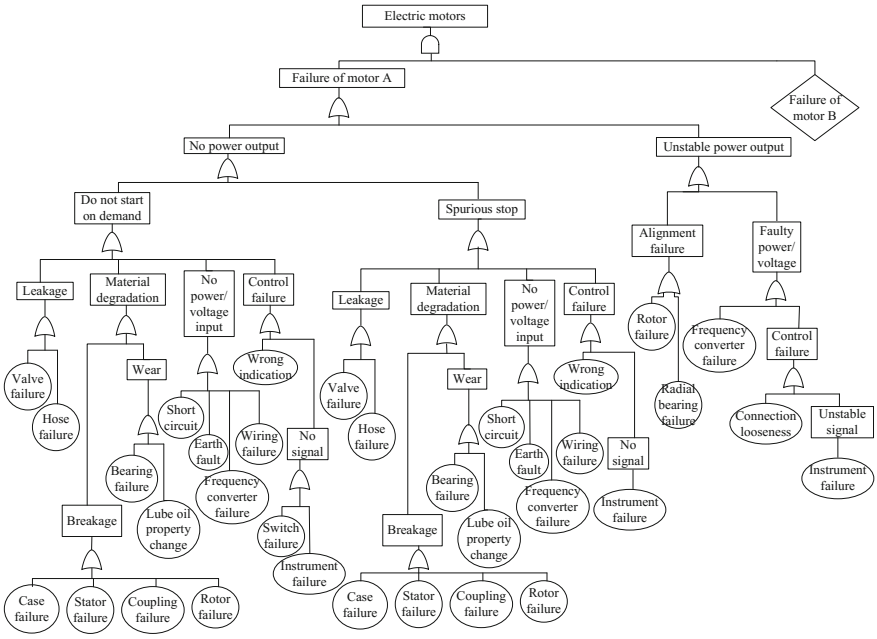


Fig. 2 FTA of electric motor in drawworks system [19]

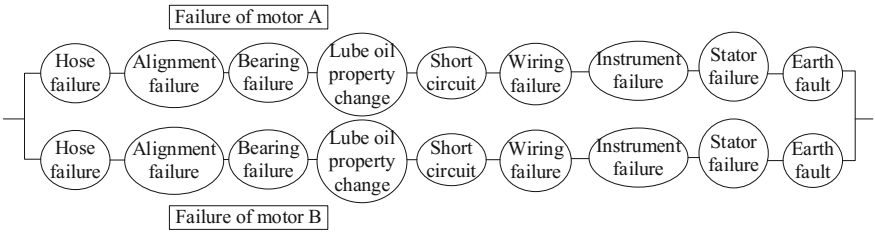


Fig. 3 Reliability diagram for electric motors. Adapted from Zhu [19]

where I^B is the measurement of importance for component i, $h(1_i, p)$ is the overall system reliability when component i is in its best condition, and $h(0_i, p)$ is the overall system reliability when component i fails.

In this case, the basic events in the fault tree were considered independent from each other, and the state of each component was set to be binary (either fail or function). A simplified reliability block diagram (shown in Fig. 3) was depicted based on the FTA, which were later used for reliability calculation. For the parts that were not recorded in the OREDA handbook, the reliability numbers were rounded up to 1 for simplification.

The results of Birnbaum's importance were shown in Table 2. The highlighted component (instrument failure) was the one that had the highest potential influence on the system reliability of the electric motor. It was thus suggested that maintenance in relation to the instrument parts be prioritized for continuous monitoring and preventive maintenance planning.

From the reliability analysis, it is noticeable that the difference of importance between the components/parts are small, but the difference can be big in other systems. Operators are recommended to establish and use their own databases to serve the purpose in addition to the reliability data from industry data base.

Table 2 Reliability calculation results with Birnbaum's importance measurement

Component	Failure rate (λ)	MTTF	MTTR	I^B
Bearing	4.128	242,247	1.4	0.000308867
Hose/pipe	0.458	2,183,329	2.6	0.000308865
Short circuit	0.916	1,091,665	7.5	0.000308867
Open circuit	0.458	2,183,329	7.3	0.000308866
Earth fault	0.916	1,091,665	6.7	0.000308867
Instrument failure	**10.094**	**99,069**	**15.3**	**0.000308913**
Alignment failure	0.916	1,091,665	5.1	0.000308866
Oil property change	0.916	1,091,665	4.2	0.000308866
Wiring failure	3.67	272,480	29.9	0.000308899
Stator failure	1.374	727,777	9.5	0.000308869

Adapted from Zhu [19]

4.3 Development of Data Architecture

Various data sources are created during the design, operation and maintenance of systems. How to manage data efficiently and effectively needs to be addressed. Different decision-makers have different needs and requirements on data. An integrated architecture to identify and access key data sets is a key element of data-driven decision-making. Liyanage [13] introduced a conceptual data architecture, where he discussed how key data sets should be defined in a holistic manner and how data should be disseminated depending on the specific context. In this section, the establishment of the data architecture of drawworks is explained in two main parts, which are data definition and data dissemination.

Definition of Data Sets

Project data include the system, location and field in which the drawworks is operated. ISO [11] defines the characteristics and attributes in defining and collecting equipment data, failure data, and maintenance data. Equipment data include the manufacturing information of the drawworks, its operating mode, operating power, the location of its applications and so on. Failure data needs to be recorded with the correct identification and with a time tag. Failure modes, possible causes, and potential consequences could be identified with the help of FTA and FMSA. In addition, it is suggested that failure data from both the operator and drawworks' suppliers are recorded and shared for a better understanding of the system. As proposed by ISO [11], maintenance data should cover the category of maintenance activity, the item maintained, the resources and manpower used, the impact from maintenance, downtime and so on. There can be overlappings of events between failure data and maintenance data, but they are stored with different criteria and serve different purposes. Reliability analysis, as explained in Sect. 4.2, uses failure history or industry database as the basis for calculation.

Building a contextual database is the key part of the data architecture and plays a critical role in O&M decision-making. The use of FMSA analysis explains how key monitoring parameters are defined, as discussed in Sect. 4.1. Table 3 shows the use of FMSA as a platform for discussing the technical options of gaining condition data. Alternative methods may exist to collect specific data, and the decision as to which method is more efficient and cost-effective belongs to the operator. The identifications of failure symptoms and failure mechanisms is helpful in implementing diagnosis and prognosis of the system.

Dissemination of Data

The basic principle of data dissemination is to make sure that the right people have access to the right data at the right time [8]. In the data architecture, data are used for both common and specific purposes. Common use data include project data, equipment data, industry database, and so on. These kinds of data provide the basic information of drawworks' operating context and environment and should be available and accessible for all personnel involved. Specific data include failure

Table 3 Approaches to gain condition awareness, based on FMSA analysis

System	Drum assembly	Braking system	Braking system	Braking system	Braking system	Electric motors	Electric motors	Drill control system	Drill control system	Lubrication system
Failure symptoms	Breakage of wireline	Fast wearing of brake discs	Breakage	Brake power is not enough	Brake functions improperly	Unstable power output	Unstable power output	Switch failure	Failure to respond on command	No rise in pressure
Condition awareness	Visual inspection	Visual inspection	Visual inspection	Test	Test	Torque monitoring	Torque monitoring	Visual inspection	Test	Pressure monitoring

data, maintenance data, reliability data, and contextual data of the drawworks. These data sets are directly connected to the history, condition and performance of the drawworks. Not all personnel need to access and understand these data sets, but these data are critical for maintenance engineers, drilling engineers, and so on, to gain context awareness, to optimize maintenance plans, and to implemented critical tasks.

In the case, condition monitoring system has not been applied with full scale on drawworks. One challenge was the lack of definition of critical parameters to monitor. Another major challenge was the lack of platform to integrate these data into decision-making. With the help of FMSA, the first challenge can be easily solved by analysing the components/sub-systems with the highest value on MPN and Birnbaum's importance measurement. For the second challenge, FTA results can be used as the platform for decision-making, both logically and with quantitative analysis.

4.4 Extended Competence Management

Three disciplines are involved in the O&M activities regarding drawworks, including drilling, maintenance, and safety. Based on the observations, employees with roles and responsibilities related to drawworks have been working in their positions for many years. The knowledge and competence of each discipline were observed to be specialized but comprehensive related to the tasks they performed.

However, the complexity of the conditions in which the drawworks is operated still keeps growing, even though the field has been developed for a long time. New sensors and equipment have been added, and new levels of safety and production performance are expected. The interconnections of different data sets, the knowledge from different domain experts and the decision alternatives are increasing. Making informed decisions in the way they used to do is becoming more and more challenging. There is a trend of different disciplines being expected to become multi-disciplinary to improve work performance [13]. In the new context, drilling engineers may be expected to understand safety concerns, as well as financial costs, from example. Companies need to realize that extended competence can play a key role in realizing the full value of the new data architecture, as well as in improving decision-making quality.

5 Discussion and Conclusion

The industry is moving towards an era with great appreciation of data and information. Decision makers are expected to continuously define critical data sets and to make efficient and informed decisions with the right data. A pressure vessel

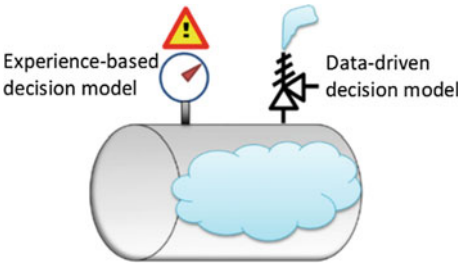


Fig. 4 Pressure vessel model for experience-based and data-driven decision model

model is used to illustrate the value of the data-driven decision method compared to traditional experience-based decision model, as shown in Fig. 4.

In the pressure vessel model in Fig. 4, either pressure gauge or pressure relief valve can be installed to respond to a possible overpressure scenario. The traditional experience-based decision model is just like the pressure gauge in the model, which only means something when someone with the right experience happens to notice it and knows how to react to it. The suggested data-driven approach is like the pressure relief valve, which comes with a whole solution for decision-makers, with the context awareness, failure modes and mechanisms, diagnosis, and consequences displayed and explained.

The paper promotes a normative approach to help decision-makers define the scope of key data sets as well as the process to establish a contextual data architecture for decision-making. The analysis tools that are used in the approach are known to most decision-makers in the offshore O&G industry sector, which makes the approach relatively simpler to be implemented. The method aims to provide some practices to help decision-makers gain control and confidence when too much data can become overwhelming and confusing. The approach can be used to assess other critical equipment or systems with simple adaptations.

Acknowledgements The paper is mainly derived from the author's master thesis [19]. The patience of Sukhvir Panesar and Rajesh Kumar, regarding discussions of work processes related to drawworks, is gratefully acknowledged, as is the help of Prof. J. P. Liyanage in defining and establishing the research methodology.

References

1. Athey S, Levin J (2017) The value of information in monotone decision problems. Res Econ
2. Aven T (2012) Reliability and risk analysis. Springer Science & Business Media
3. Baker R (1998) A primer of offshore operations, 3rd edn. University of Texas at Austin, Houston
4. Berger JO (2013) Statistical decision theory and Bayesian analysis. Springer Science & Business Media
5. Birnbaum ZW (1968) On the importance of different components in a multicomponent system. Seattle lab of statistical research. University of Washington. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/670563.pdf

6. Cayeux E, Daireaux B, Dvergsnes E (2011) Automation of drawworks and topdrive management to minimize swab/surge and poor-downhole-condition effect. SPE Drill Complet 26(04):557–568
7. Faller K (2008) Combining condition monitoring and predictive modeling to improve equipment uptime on drilling rigs. Paper presented at the SPE Gulf Coast Section 2008 digital energy conference and exhibition
8. Fischer G (2012) Context-aware systems: the 'right' information, at the 'right' time, in the 'right' place, in the 'right' way, to the 'right' person. Paper presented at the proceedings of the international working conference on advanced visual interfaces
9. Hajian-Hoseinabadi H (2013) Reliability and component importance analysis of substation automation systems. Int J Electr Power Energy Syst 49:455–463
10. ISO (2002) ISO13379, condition monitoring and diagnostics of machines—data interpretation and diagnostics techniques which use information and data related to the condition of machines—general guidelines
11. ISO (2016) ISO/TC 67, ISO 14224:2016, petroleum, petrochemical and natural gas industries—collection and exchange of reliability and maintenance data for equipment (3rd ed). International Organization for Standardization
12. Larson JR Jr, Christensen C, Franz TM, Abbott AS (1998) Diagnosing groups: the pooling, management, and impact of shared and unshared case information in team-based medical decision making. J Pers Soc Psychol 75(1):93–108
13. Liyanage JP (2003) Operations and maintenance performance in oil and gas production assets: theoretical architecture and capital value theory in perspective. Ph.D. Thesis, Norwegian University of Science & Technology, Norway
14. Nguyen TC (2012) Drilling engineering: rotary drilling system. Lecture notes. The New Mexico Institute of Mining and Technology, New Mexico
15. Norsok (2012). D-001 NORSOK STANDARD, drilling facilities, D-001, Rev3. Norway
16. OREDA (2009) OREDA handbook. 1-topside equipment. DNV, Norway
17. Rakow T, Newell BR (2010) Degrees of uncertainty: an overview and framework for future research on experience-based choice. J Behav Decis Mak 23(1):1–14
18. Si S, Dui H, Zhao X, Zhang S, Sun S (2012) Integrated importance measure of component states based on loss of system performance. IEEE Trans Reliab 61(1):192–202
19. Zhu P (2015) Data-driven decision-making practice in response with drawworks maintenance notifications. Master of Science, University of Stavanger, Stavanger, Norway. Retrieved from https://brage.bibsys.no/xmlui/handle/11250/2353394