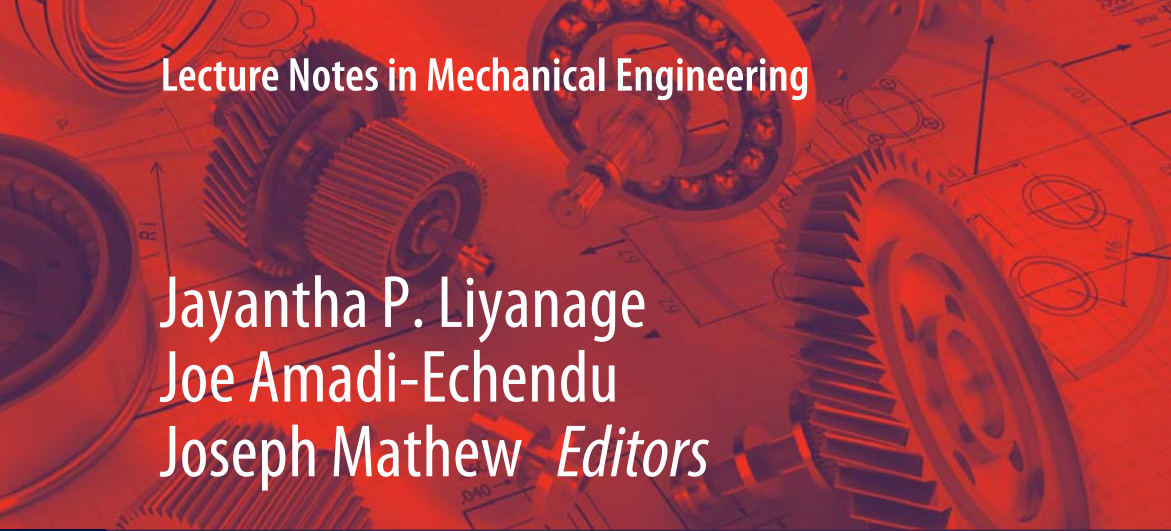


Lecture Notes in Mechanical Engineering

Jayantha P. Liyanage
Joe Amadi-Echendu
Joseph Mathew *Editors*

Engineering Assets and Public Infrastructures in the Age of Digitalization

Proceedings of the 13th World Congress on Engineering Asset Management

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 the Springer Editor of your location:

China: Dr. Mengchu Huang at mengchu.huang@springer.com
India: Priya Vyas at priya.vyas@springer.com
Rest of Asia, Australia, New Zealand: Swati Meherishi at swati.meherishi@springer.com
All other countries: Dr. Leontina Di Cecco at Leontina.dicecco@springer.com

To submit a proposal for a monograph, please check our Springer Tracts in Mechanical Engineering at http://www.springer.com/series/11693 or contact Leontina.dicecco@springer.com

Indexed by SCOPUS. The books of the series are submitted for indexing to Web of Science.

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

Jayantha P. Liyanage · Joe Amadi-Echendu ·
Joseph Mathew
Editors

Engineering Assets and Public Infrastructures in the Age of Digitalization

Proceedings of the 13th World Congress on Engineering Asset Management

Springer

Editors
Jayantha P. Liyanage
Faculty of Science and Technology
University of Stavanger
Stavanger, Norway

Joe Amadi-Echendu
Graduate School of Technology
Management
University of Pretoria
Hatfield, South Africa

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

ISSN 2195-4356 ISSN 2195-4364 (electronic)
Lecture Notes in Mechanical Engineering
ISBN 978-3-030-48020-2 ISBN 978-3-030-48021-9 (eBook)
https://doi.org/10.1007/978-3-030-48021-9

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

Preface

The 13th World Congress on Engineering Asset Management (WCEAM-2018) was held on 24–26 September 2018 in Stavanger, Norway. It was organized and hosted collaboratively by the Cluster on Industrial Asset Management (CIAM) of University of Stavanger and the International Society of Engineering Asset Management (ISEAM).

The brief idea for organizing the World Congress in Stavanger emerged in early 2017. It gradually became a major undertaking through various dedicated tasks in 2017 and 2018. This remarkable journey took place through a demanding terrain relying much on active support, engagement, and encouragement from many sources and champions. It brought together minds and hearts of both local and global engineering asset management communities to create a meaningful and a professional event.

The World Congress was organized in 2018 under the theme of "Engineering Assets and Public Infrastructures in the Age of Digitalization", aiming to establish a strategic fusion between leading research and modern industrial practices. It engaged domain experts from all parts of the world. It is a triumph of true passion, creative thinking, and joint commitment.

WCEAM-2018 constituted 10 keynote speakers, 25 parallel technical sessions, 106 research and industrial presentations from over 20 countries, 4 technical safaris to selected engineering sites in the region, and 3 on-site operational demonstrations. It attracted over 200 delegates from 25+ countries. Among the most popular themes of the World Congress included:

- Asset economics and decision analysis
- Asset management in Industry 4.0
- Co-value creation, organizational strategy, and new business models
- Data science, asset health, and predictive analytics
- Human capital and organizational development
- Performance measurement and management
- Modern digital applications
- Critical asset processes and process efficiency

- Dynamic modelling, simulation, and visualization
- Smarter and safer assets
- Sustainable assets and processes.

The Organizing Committee of WCEAM-2018 comprised Prof. Jayantha P. Liyanage (University of Stavanger), Fride Solbakken (Stavanger Maritime Forum), Frode Berge (Stavanger Chamber of Commerce), Per Morten Haar (Region Stavanger), Alexander Landsnes (Stavanger Forum), Ole Trætteberg (Gassco), Øyvind Rudolf Lea (Equinor), and Birger Haraldseid (Greater Stavanger). Their involvement and ideas have been very valuable all the times through the initial concept to the final event.

We gratefully acknowledge generous contributions from the primary sponsors (Rogaland county, Stavanger municipality, and Sola municipality), silver sponsor (Apply Sørco), and additional sponsor (Petroleum Safety Authority of Norway). We also received continuous support from all industrial partners of CIAM that included: Equinor, ConocoPhillips, Aker BP, DNV GL, Apply Sørco, Oceaneering, Gassco, Petroleum Safety Authority of Norway, Kverneland Group, Petrolink, Jotne, Ernst & Young, PricewaterhouseCoopers, Visco AS, VIS-M As, Stinger, and ORS Consulting.

The inaugural and keynote speakers of the event shed some light on the current importance and future prospects of various disciplines of engineering asset management. These specially invited speakers included:

- Christine Sagen Helgø (The Honourable Mayor of Stavanger)
- Prof. Marit Boysen (Rector of the University of Stavanger)
- Line Heldal Bakkevig (Vice-President, Equinor, “Digitalization towards Field of the Future”)
- Michael Campbell (Senior Director, Children’s Health Queensland Hospital and Health Service, Australia, “Innovative Health Infrastructure and Medical service solutions in the Digital age”)
- Adrian Park (Vice-President, Intergraph Norway/Hexagon PPM, USA, “Digital transformation and leveraging the Digital Twin, through the Asset life cycle”)
- Tom Eystø (CEO, Massterly, Norway, “Towards a New Maritime Adventure with Autonomous shipping and Ferry services”)
- John Woodhouse (Managing Director, Woodhouse Partnership, UK, “Don’t forget human psychology in asset management decisions – it’s not all about data and analytics”)
- Tor Kristian Gyland (CEO, Green Mountain AS, Norway, “#TheNaturalChoice: Modern Data centers for Critical assets from inside and out”)
- Anne-Lene Festervoll (Director, PricewaterhouseCoopers, Norway, “Workforce of the future” – and the Global future of HR”)
- Sverre Alvik (Director, DNV GL, Norway, “Global Energy Transition Outlook for Engineering Assets and Public Infrastructures”)
- Cato Vevatne (CEO, KVS Technologies AS, Norway, “Industrial Robotics: The Robotic Ecosystem – smart, distributed and scalable”)

- Knut Roar Wiig (CEO, Nordic Unmanned AS, Norway, "Industrial Drones: Unexplored possibilities and Industrial disruption").

The event constituted 4 technical safaris hosted by Road Authority of Norway (Ryfast, the world's longest underwater tunnel), IRIS Research Institute (robotic rigs and automated drilling), Oceaneering AS (eROV and integrated operations), and Green Mountain AS (data inside a mountain). The on-site demonstrations performed by KVS Technologies (industrial robotics for operational safety), Nordic Unmanned AS (industrial drones), and Teknova (Machine diagnostic) added meaningful technical insights into these new industrial technologies.

All the authors contributed professional papers with novel ideas, concepts and solutions across a range of disciplines that had scientific and pragmatic value. A specific group of experts also organized special sessions, which included: Nuclear Power Plant Monitoring and Long-Term Asset Management (by Prof. Belle Upadhyaya, University of Tennessee, USA, and Jon Kvalem, OECD Halden Reactor Project, Norway), Asset Management for Hydro Power Sector (by Dragan Komljenovic, Hydro-Quebec, Canada), Advances in Equipment Condition Monitoring (by Thomas J. J. Meyer, Teknova, Norway), Applied Data Science (by Assoc. Prof. Antorweep Chakravorty, University of Stavanger, Norway), Macro ergonomics and Organizational issues for Human performance and Workplace safety (by Prof. Seppo Vayrynen, University of Oulu, Finland), Organizational Strategy and Entrepreneurship (by Prof. Tatiana Iakovleva, University of Stavanger, Norway), Service Innovation in Maintenance through Industry 4.0 (by IFAC TC 5.1 WG AMEST (Advanced Maintenance Engineering, Services and Technology) and IFIP WG 5.7 SIG in Service Engineering, coordinated by Marco Macchi of Politecnico di Milano, Italy; Paolo Gaiardelli of Università degli Studi di Bergamo, Italy; and Christos Emmanouilidis of Cranfield University, UK), Tunnel Safety (by Prof. Ove Njå of University of Stavanger, Norway, and Helen Roth of Tunnel Safety Cluster, Norway), and Regulations and Audits of Late life Engineering assets (by Ole Jørgen Melleby, Petroleum Safety Authority of Norway). All papers were double-blind peer-reviewed by a dedicated International Scientific and Review Panel that comprised global experts with domain expertise related to engineering, operations, and management of engineering assets and public infrastructures. Their commitment as members of the International Scientific and Review Panel was instrumental to the quality and high standard of the event.

The WCEAM-2018 administrative team made a remarkable contribution in various capacities to ensure the success of the event. The team comprised Andrew Kilmartin, Vanessa Grace Booc, Assoc. Prof. Idriss El-Thalji, Joanne Therese Taylor, Shasheema Hewapathirana, Tone Dalaker, David Martin Bø, Yasar Denis Cam, Odd Arne Nordmark, and Arun Sindre Joshi, all from University of Stavanger, Norway. In addition, professional support was also provided by Elin Paulsen of Region Stavanger, Sindre Haaland of Conventor, the Clarion Energy Hotel (World Congress venue), and Event Eye (developer of the WCEAM-2018 app). On behalf of ISEAM, additional administrative assistance was generously received from Betty Goh, Executive Assistant of the Asset Institute, Australia.

We also gratefully acknowledge the unconditional engagement and continuous support from the 2017 and 2018 Board of Directors of CIAM, and the Board of Directors of ISEAM.

In essence, all contributors and delegates of WCEAM-2018 World Congress were a part of a noble purpose. The very success of the World Congress in 2018 is truly attributable to their professional involvement and generous commitments.

It indeed was a privilege to plan, organize, and deliver such an interdisciplinary global event focussing on the emerging domain of engineering asset management, where both academic and industrial communities are challenged to think different, broaden their perspectives, and work together for the betterment of public and private sectors. This is a part of a collective professional effort to resolve industrial complexities, economic chaos, and environmental impact to make the world a more sustainable place to live.

Truly, this is all about us and our common future.

Thank you all very much !!!

Jayantha P. Liyanage
Chair, WCEAM-2018 World Congress, Former Chair of CIAM

Joe Amadi-Echendu
General Chair, WCEAM-2018 World Congress, Chair ISEAM

Joe Mathew
Co-chair, WCEAM-2018 World Congress, Founding Chair ISEAM, and CEO, Asset Institute, Australia

Contents

Part I: Advances In Modern Engineering Assets and Infrastructure Management

The Implementation of ISO 55000 in Small and Medium Enterprises: Requirements and Constraints

Ibifuro Ihemegbulem(✉) and David Baglee

University of Sunderland, Sunderland, UK
bg49te@research.sunderland.ac.uk,
david.baglee@sunderland.ac.uk

Abstract. In 2014 the International Standardisation Organisation (ISO) responded to the increasing demand from asset managers to provide a structured and comprehensive standard to improve the effectiveness of assets within an organisation. The new standard ISO 55000 is comprehensive and detailed with capability for supporting the coordinated activities of an organisation in its quest to realise value from assets through, among others, improved financial performance and improved maintenance strategies. This allows the organisation to improve its decision-making, balance costs, risks and performance. The aim of this paper is to examine the compositions of ISO 55000 and identify the key elements necessary for its successful and economically feasible implementation. The study, is the last stage of a three year PhD programme, that consist of part of the research analysed data with asset managers and asset maintenance personal within SMEs to obtain a set of key issues which need to be addressed. In addition, the paper highlights the need for the development of an effective and efficient approach that makes the adoption and implementation of ISO 55000 suitable for small to medium enterprises. This new approach addresses several inherent constraints often faced by small companies when adopting new manufacturing and maintenance initiatives.

Keywords: ISO 55000 · Small to Medium Sized Enterprises (SMEs) · Asset Management (AM)

1 Introduction

SMEs are essential to the economy of any country whether developed or developing economy (Kapurubandara and Lawson 2006). The importance of SMEs depends upon their role in evolution at diverse phases of economic development. SMEs contribute to the output and attract significant foreign reserves into the country and provide employment. According to (Al Mahrouq 2010) this makes them the backbone of the private sector all over the globe. The majority of organisations are considered SMEs, playing a significant role in the world economy (Islam et al. 2011). SMEs represent more than 95% of private sector enterprises, more than half of total employment and create one third of investment0 (Cansız 2008). The performance of SMEs has a major impact on the performance of the wider economy. The capacity of manufacturing

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 3–14, 2020.
https://doi.org/10.1007/978-3-030-48021-9_1

SMEs to maintain reliable and continuously improving the manufacturing processes can be considered key for ensuring long-term sustainability.

SMEs should be able to respond quickly to indications from the market (Revilla and Fernandez 2012; Thomä and Bizer 2013) to ensure their continuous survival. The main challenges faced by SMEs are linked to their developing phases, including lack of managerial competences, poor decision-making, and scarcity in finance, personnel resources and deficiency in essential proficiency. The lack of human resources in SMEs drives them to outsource to progress their organisations. These intensify the need for value creation among SMEs to guarantee business sustainability. The management process in SMEs is characterised by the extremely personalised choice, biases, and attitudes of the entrepreneurs and owners (Beaver and Jennings 1995). Managers must deal with a high form of intricacy that affects several of their companies' assets and product line. The difficulty in overcoming obstacles found in areas of competition, technology, markets, production capacity, product range, level of entrepreneurs' skills and product quality assurance contributes to the SME mortality rate (Murphy 1996).

Several SMEs have chosen to embark on enhancement programs that have produced positive outcomes in bigger companies, for example the ISO standards. Enhancement programs will need to align with the company's strategic focus, therefore it is necessary to ensure SMEs are able to design, develop and implement the necessary tools and techniques in ISO 55000 to remain competitive and sustainable. The ISO 55000 standard is seen as an improvement tool for large companies and not a system, which could be easily adopted by SMEs due to the complexity and large number of requirements. It is in this context that, the Implementation of ISO 55000 in Small and Medium Enterprises (SMEs) assumes significance. The overall question therefore is how a new appropriate approach to ISO55000 that will address issues specific to SMEs can be developed. While this is study is part of a larger doctorial study, certain questions were selected to address question raised in this paper such as constraints to SMEs and ranking the criticality of ISO 5500 element in the order of importance to SMEs. This will help in selecting the proper approach for implementation of ISO 55000 that addresses both the constraints and needs of SMEs so they benefit from the advantages enjoyed by larger firms.

2 SMEs Constraints

Despite incentives, policies, programmes and support designed at overhauling SMEs, most perform below expectation. Organisations and operators have several reasons as to why SMEs have not been able to live up to expectancy, an average operator always centre failure on lack of access to finance, others consider inappropriate management skills, difficulty in accessing global market, lack of entrepreneurial skills, inconsistency in monetary, fiscal and industrial policies, multiple taxation and levies, inadequate infrastructure and lack of modern technology etc. are mainly responsible. Harsh economic conditions and operating environment affect the development of SMEs in most countries, thereby afflicting business environment. Discussing some of this challenges are;

- Finance is mostly considered key constraint of SMEs, though this may be true but survey evidence show that finance contributes only 25% to the success of SMEs. In developing countries SMEs suffer from access to appropriate funds from both banks and capital markets; due to the opinion of higher risks resulting in high mortality rate of the business, information irregularity, inadequate collateral, absence of, or unverifiable history of past credit(s) obtained and lack of adequate historical records of the company's transaction (Oteh 2010). Most SMEs rely heavily on savings or bank loans for expansion capital (KPMG 2011). SMEs attempt to finance their fixed assets with long-term debt, and their current assets with short-term debt; SMEs with low asset structure have greater difficulty accessing long-term debt, the only option is to fall on short- term debt finance. Without enough and sustainable finance SME will not grow (SME-RC 2012).
- Entrepreneurial skills: SMEs operator must possess the capacity to manage and acquire basic skill of planning, organizing, coordinating, leadership and communication. Education is speculated to affect the revenue of businesses positively; lack of essential skills causes high failure rate. Inadequate investment in human capital hampers the SMEs growth because of the scarcity of skilled workers, managers and entrepreneurs (Tendler and Amorim 1996). Organisation must consider creative and innovative skills, work experience, technical and managerial skills to achieve necessary results.
- Inadequate infrastructure and institutional support: lack or weak infrastructure such as power, feeder roads, telecommunication, technologies etc. are still the bane of SMEs growth. Lack of access to appropriate technology as well as near absence of research and development
- Harsh conditions: fiscal and monetary measure, high interest rate and unstable foreign exchange as well as high inflation rates. These can weaken the economy and reduce the competitive abilities of SMEs and their existence becomes a struggle.
- High dependence on imported raw materials with the attendant high foreign exchange cost and scarcity at times.

3 ISO 55000

ISO 55000 is a written set of rules published by International Organization for Standardization (ISO 55000 2014). The rules describe practices that are universally recognized and accepted for guaranteeing that organizations constantly understand and meet the demand of their clients. ISO 55000 assists organisations to proactively manage the lifecycle of their assets, from acquisition to disposal. ISO 55000 is designed to apply to any asset type, though the main focus is on physical asset management (Woodhouse 2012). ISO 55000 permits a company to identify assets that are critical to fulfilling its investor requirements, business goals through a set of rules which supports the management of assets. It helps to manage the risks and costs related with possessing assets in an organized, efficient method that supports frequent improvement and continuous value creation. The aim is to use ISO 55000 to specify a set of requirements for establishing, implementing, maintaining and improving an organization's asset management system, which can be applied to any company where physical assets are central to their business. It is vital for a company to efficiently manage the wide range of assets on which it depends.

The ISO 55000 family of standards comprises three documents: 55000; provides critical overview, concepts and terminology;

55001; specifies the requirements for an effective Asset Management System; and

55002; offers interpretation and guidance for such a system to be implemented.

ISO 55000 series of standards is built on the general Plan, Do, Check, Act (PDCA) framework, which is recognized as the basis for continuous improvement in management systems (Moen and Norman 2010). Implementation of the standard boosts proactive maintenance of assets, which in turn should lead to reduced failure and waste as well as improved services. ISO 55000 helps organizations establish an Asset Management System (AMS) for optimizing assets. AMS communicates with elements that produce policy, objectives and procedures to accomplish an organization's objectives. Observing these requirements (ISO 55001) allows for consistent decision-making on activities that impact asset-related risks, performance, and cost profiles. This indicates that management should be equipped to make objective, predictable, and consistent decisions that involve trade-offs between short- and long-term effects, and optimal combinations of interrelated and conflicting benefits. ISO 55000 consist of seven main elements:

- Organizational Context
- Leadership
- Planning
- Support
- Operation
- Performance Evaluations
- Improvements.

Implementing these elements establishes a system specification for the management and care of assets subsequent to acquisition. Numerous organizations that have embarked on AM programs and failed, often exhibit a lack of understanding of the essential concepts they are trying to implement. It is vital that organizations understand, define and communicate the objectives of their AM program, taking into consideration different functional groups, such as maintenance or engineering.

4 Implementation Requirements

Implementation of ISO 55000 standards requires new information, analytical tools, new ways of communication in the organizations and new management practices. Introduction of ISO 55000 begins by organizational reform of policy, targets and outcomes in agreement with a tactical plan of the organization using steps as follows (Mikolaj et al. 2012):

- Authorization of asset structure and accountability for the assets;
- Monitoring of current state of the assets, potential necessities for the assets, approval of a developmental program and required resources for administration;
- Formulation of strategies for asset maintenance describing programs for achievement of the goals of the organisation and measurement of performance of the services;
- Selection of an appropriate method for introduction of the AM plan, (process of providing services);
- Inspection of asset information and data aimed at efficient decision making or efficient outputs of the AM plans;

ISO 55002 states that when the intended AMS is established or reviewed, it is important to ensure the approach is consistent and aligned with internal and external context of the organization. It is important since it can influence the scope and design of the AMS. Therefore, the following implementation objectives are considered:

- **Creating value:** The ISO 55000 series identifies that managing the asset will create value to the organization. Thus it is required of the organization to make a clear statement how the AM objectives will align with the organization as well as establish a decision-making process that reflects stakeholder needs and defines value.
- **Creating alignments:** AM related decisions enabling organization to achieve its objective. Organizations should implement a risk-based, information-driven, decision-making and planning process to transform the organizational objectives into AM plans. Organizations must endeavor to integrate the AM processes with the functional management processes.
- **Establishing leadership**: Organizations are required to establish leadership and constructive workplace culture, commitment from all managerial levels to successfully operate and improve AM within the organization.
- **Assuring that asset will perform optimally:** ISO 55000 series of standards assure that assets will fulfill their required purpose. Organizations should develop and implement processes that connect performance and purpose of assets to the organizational objectives. These processes assure capability across the life cycle of assets, providing monitoring and continuous improvement and providing necessary resources and competent personnel to demonstrate assurance by commissioning AM activities.

5 Methodology

To answer the question of how a new appropriate approach to ISO 55000 that will address issues specific to SMEs can be developed, a total of 197 sets of questionnaires were distributed to selected respondents SMEs and 128 questionnaires were collected back indicating a good response rate of 65%. A simple frequency and average analysis of respondent's responses to questions bother on constraints to SMEs and elements in ISO 55000 SMEs finds significant to business success is presented in the chart below (Fig. 1);

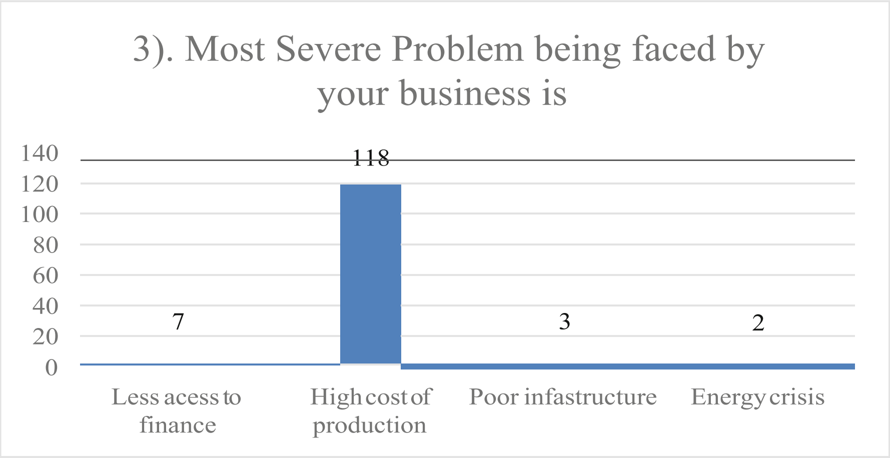


Fig. 1. Problem faced by SMEs

118 respondents said high cost of production were severe problem their business face, this study deduced that access to finance is not the main problem but high cost of production is the most severe problem faced by SMEs. Production costs are expenses, such as materials and labour that company incurs in the course of producing the product that they sell to consumers. The lower production cost, the higher profit; however, low production costs do not necessarily guarantee a high profit. Equipment's that lack maintenance and fails frequently lack precision and tend to produce defects at the same time increasing production cost which amounts to less profitability, thereby putting SMEs existence at risk.

7 respondents indicated that less access to finance was another severe problem faced by SMEs, this can be traceable to the unwillingness of banks to extend credit, inadequate documentation of business proposals, lack of appropriate and adequate collateral, high cost of administration and management of small loans as well as high interest rates. A prior study by Stella (2011), suggest that SMEs require adequate financing to meet their needs at each stage of their life cycle, from creation through operation, development, expansion, and beyond. Financing is necessary to help them set up and enhance their operations, develop new products, and invest in new staff or production facilities (Stella 2011). Less access to finance have been major constraints to SMEs hence the need to explore alternatives (Stella 2011). According to Shah et al. (2013), financial institutions behave more cautiously when providing loans to SMEs, and SMEs are usually charged comparatively high interest, high collateral and loan guarantees (Fig. 2).

Fig. 2. Success of organizations

124 respondents said planning, monitoring and control contribute to the success and optimizing business functions. Organisation needs to develop plans that focused on ensuring assets continue to support or link physical resources to service delivery programs and business objectives. These plans should include capital investment, maintenance, disposal, risk management of specific asset. These plans should cover all assets including land, buildings, information technology, infrastructure, collections, equipment and fleet, owned or controlled by an agency. Small or medium-sized businesses that do not have developed a specialized business plan face many chances of failure. Condition and levels of service “health” of an asset should be monitored. Measuring the condition of an asset involves systematically examining components and systems and documenting their condition according to the relevant standards for each element. Condition of the asset can include pipe breaks, pavement wear, roof leaks, foundation cracks, equipment malfunctions and failures, organisation processes. An effective control (KPI) structure should be establish and transmit asset policies and procedures using information system that provides reliable, relevant and timely data with which to make informed asset management decisions.

124 respondents said risk management contributes to the success of the organisation. Risk management is a organized way to identify and analyse possible risk, and formulate and implement appropriate responses according to classes of risks. The objective of the risk management is to make informed decisions on asset priorities across a very complex and diverse portfolio of assets and to maximize the return on investments. These responses may include risk prevention, risk transfer, minimising the impact or acceptance of risk. A combination of these strategies may apply to manage different individual risks within a particular activity or project. Management must consider the agency’s exposure to risk throughout the asset management process.

96 respondent said resource allocation contributes to their; Resource allocation in the context of human capital, the management team must be equipped with skills such

as finance management and operations management skills to ensure efficiency in the allocation of limited resources and strategic planning to ensure the firm's development (Hove and Tarisai 2013) (Fig. 3).

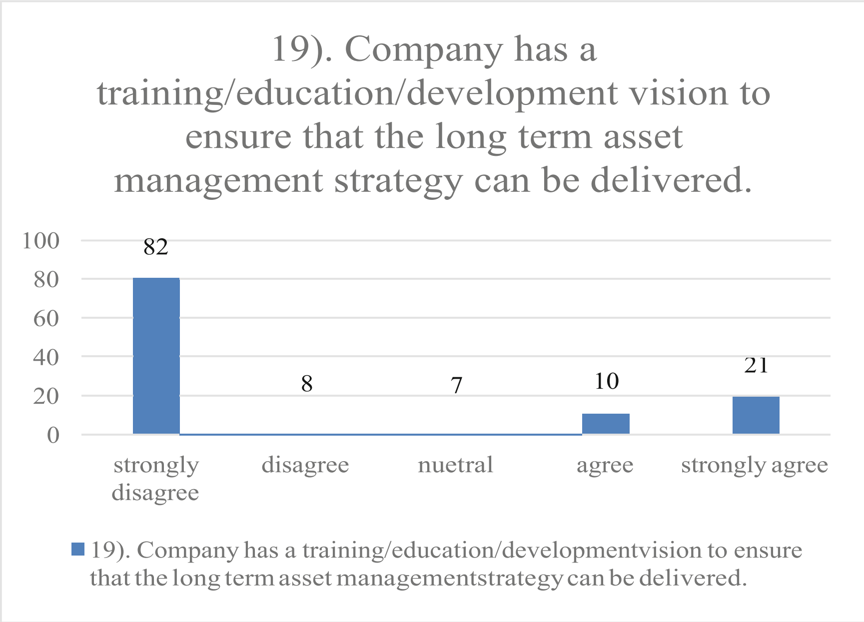


Fig. 3. Company vision

82 respondents strongly disagree that there organization trains/educate or have vision to ensure long-term asset management. Organizations must ensure that all persons who use work equipment have adequate training especially for of health and safety purposes; trainings must include methods to be adopted when using work equipment, risks associated to the equipment and precautions to be taken. Brainstorming for ideas can create vision after a training or workshop on successful practices (Campbell and Reyes-Picknell 2006). Vision is the conclusion of what to achieve, a realistic picture of a future state, which is desired for the organization; it is important to be innovative and encourage the employees to think new and big, and the vision should be understandable to everyone (Campbell and Reyes-Picknell 2006; Bergman and Klefsjö 2010; Thomas 2005) (Fig. 4).

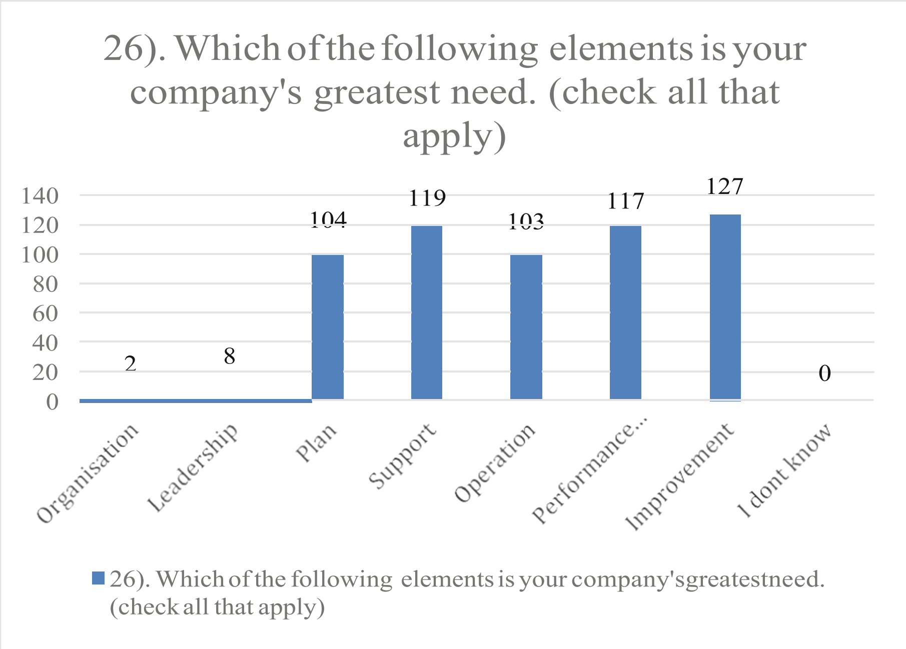


Fig. 4. Elements of ISO 55000 mostly needed

Respondents List What Their Greatest Need in the Elements of ISO 55000, Ranging from Highest Value

- 127 respondents said improvement is their greatest need; asset management always aims for continual improvement. If the performance of the AMS fails to meet its objectives, opportunities for improvement should be examined. It is important to check if there is improvement to a system to prevent event from reoccurring (ISO 55000 2014).
- 119 respondent said support is their greatest need within the elements of ISO 55000; there should be collaboration of all departments and staff and it is important that relevant stakeholders are aware of asset management requirements and expectations, and how implementation and changes to AMS can impact stakeholders. Organization should determine information needs related to assets, asset management and AMS. According to ISO 55002 (2014) the organization should define its period of responsibility for every asset, and requirements for maintaining documented information beyond disposal of assets. Resource allocation that wont generate repercussions on other parts of the organization to gain and/or to give additional resources.
- 117 respondents said Performance Evaluation remained their greatest need; the process for monitoring and control of operations should be established. Management review(s) provides top management with an opportunity to establish new and/or updated asset management objectives based on the evaluation of suitability, adequacy and effectiveness of assets, asset management and AMS. Frequent

internal audits should be implemented and maintained to ensure the effectiveness and suitability of AMS towards the requirements of the organization and the ISO 55000 suite.

- 104 respondents said Plan remained their greatest need; actions, responsibilities, resources and timescales intended to implement the strategic asset management plan and deliver the asset management objective. There should be documentation in place in order to prove that risk management processes have been executed as planned. Contingency planning should be established for significant events to ensure crucial operations (business continuity) and containment of damage, injury or loss of personnel and property.
- 104 respondents said operation remained their greatest need; Operation is the implementation, conduction and control of processes and activities previously planned. One need to:
  - Establish criteria. One need to define who is responsible, who is doing what, why it is done, determine procedures and ensure that a possible resource gap is dealt with.
  - Implement control mechanisms such as performance measures, risk management and scheduling reviews. This is of Importance in order two monitor and treat risk and deviations in the attempt to achieve balance between cost, risk and performance.

6 Benefits of Implementing ISO 55000

ISO 55000 provides organization the ability to determine themselves how the minimum set of requirements for an operative AMS can be implemented to fulfill their needs (Minnaar et al. 2013). The implementation of ISO 55000 will improve maintenance practices. These benefits include:

- Substantial improvement of asset reliability,
- Lower costs of servicing assets,
- Greater uptime and availability,
- Less downtimes and outages,
- Higher return on assets as well as in invested capital,
- More efficient and effective training: better-defined procedures and documentation facilitates employee training and knowledge transfer.

7 Conclusions and Recommendation

The aim of study is to contribute to help managers to detect focus areas in order to implement ISO 55000 in SMEs based on the requirements found in the ISO 55000 suite. The survey process was conducted to map the current status of organizations and elements in ISO 55000 that respondent believe they need in their organization. Results were analyzed, compared and integrated as positive trends and areas of improvements for each clause within the ISO 55000 suite. It should be brought to the reader's attention that these results are subjective and based on the author's interpretation of both requirements and interview research.

From findings **high cost of Production and less access to finance** were the most severe problems, Organizations will need to address issues resulting these factors. Organisation are to implement or improve on **Risk management, planning, monitoring and control, resource allocation,** and **senior management support** as they contribute to the success of their organisation as revealed from findings. Furthermore, it is identified by survey that the greatest needs of SMEs in ISO55000 elements in order of importance are: **Organizational Improvement, Organizational Support, Organizational Performance Evaluation, Organizational planning, Organisational operation,** and **Organizational leadership,** the implementation of these element will boost return on investment. Implementation of these indicated elements rather than the whole element of ISO55000 will increase the ease at which SMEs will adopt ISO55000 in their Asset management process, would see them achieved better performance and improved competitive advantages. Focusing on these elements will create pivotal point for business cutting away distractions and driving its own strategy in line with organisational objectives. Meeting the requirements in ISO 55000 does not guarantee that business performance will improve; organizations need to understand and embrace changes for a successful asset management strategy to be realized regardless of the accreditation status.

Organisations must also focus their efforts in the right place at the right time, have a clear vision with manageable components in place that can be implemented and tracked on an annual basis. Prioritizing and optimizing these efforts demands a thorough decision making process that leads to improve effectiveness and guides decisions. An asset manager advocate must be in place that will be able to communicate value, influence others, and can obtain the required organizational support. ISO 55000 structures must be incorporated within the organization's existing goals, structure, culture and their real benefits will be realized (Reyes-Picknell 2014). Leaders with asset knowledge and core maintenance strategy are important factors for a successful asset management.

References

Al Mahrouq, M.: Success factors of small and medium – sized enterprises (SMEs). The case of Jordan. Anadolu Univ. J. Soc. Sci. **10**(1), 89–90 (2010)

Beaver, G., Jennings, P.: Picking winners: the art of identifying successful small firms. Int. Rev. Strategic Manag. **6**(4), 91–106 (1995)

Campbell, J.D., Reyes-Picknell, J.V.: Uptime: Strategies for Excellence in Maintenance Management, 2nd edn. New York (2006)

Hove, P., Tarisai, C.: Internal factors affecting the successful growth and survival of small and micro agri-business firms in alice communal area. J. Econ. **4**(1), 57–67 (2013)

Islam, M., Khan, M., Obaidullah, A., Alam, M.: Effect of entrepreneur and firm characteristics on the business success of small and medium enterprises (SMEs) in Bangladesh. Int. J. Bus. Manag. **6**(3), 289 (2011)

ISO 55000. Asset management: Overview, principles and terminology (2014). www.standard.no. Accessed 10 Jan 2015

Kapurubandara, M., Lawson, R.: Barriers adopting ICT and E-commerce with SMEs in developing countries. Exploratory study in Sri Lanka, Collecter, Adelaide (2006)

Mikolaj, J., Trojanová, M., Pepucha, L.: Global trends on road administration based on Asset management. In: IFME Word Congress on Municipal Engineering (2012)

Murphy, M.: Small Business Management. Pitman s.l., London (1996)

Revilla, A.J., Fernandez, Z.: The relation between firm size and R&D productivity in different technological regimes. Technovation **32**, 609–623 (2012)

Reyes-Picknell, J.: ISO 55000 WHAT YOU NEED TO KNOW. s.l.:s.n (2014)

Thomä, J., Bizer, K.: To protect or not to protect? Modes of appropriability in the small enterprise sector. Res. Policy **42**, 35–49 (2013)

TWPL, T.W.: ISO 55000 Standards for Asset Management (2014). http://www.assetmanagementstandards.com/ISO55000.html

Woodhouse, J.: ISO 55000 Standards for asset management. s.l.:s.n (2012)

Successful Asset Management Strategy Implementation of Cyber-Physical Systems

Harald Rødseth[1(✉)] and Ragnhild J. Eleftheriadis[2]

[1] Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway
Harald.Rodseth@ntnu.no
[2] Sintef Raufoss Manufacturing AS, Product and Production Development, Trondheim, Norway
Ragnhild.Eleftheriadis@sintef.no

Abstract. With the onset of the breakthrough innovations in Industry 4.0, digitalisation is expected to improve the value creation from industrial assets. A fundamental in Industry 4.0 is to virtually represent technical objects such as machines and production plants in cyber-physical systems (CPS). As a concrete Norwegian initiative based upon these opportunities, the on-going research project CPS Plant is expected to contribute to improved production performance both in manufacturing and production companies. CPS Plant will develop and implement enabling technologies and methods for Norwegian Industries where CPS will integrate the virtual world with the physical world. To support this ambition for the Norwegian Industry, the development of a successful asset management strategy implementation in CPS Plant should be regarded as relevant. With a sound asset management strategy implemented, coordinated activities in the organisation should realize value from assets. Today an own framework for CPS Plant has been developed and is expected to be fundamental for further research activities in CPS Plant. The aim in this article is to propose an asset management strategy implementation that will support implementation of CPS Plant in Norwegian manufacturing and production companies. The asset management strategy will be based on fundamental organisation theories as well as experiences from the project. The article concludes that all though the asset management strategy Implementation will be tested in Norwegian companies, it remains to evaluate how this strategy should be "adjusted" for similar implementation in other countries.

1 Introduction

With the opportunities in the fourth industrial revolution, also known as Industry 4.0, breakthrough innovations in Industry 4.0, will enable digitalisation in terms of cyber-physical systems (CPS) (Kagermann et al. 2013). In manufacturing this development can be observed as the fourth stage of industrialisation. The first stage occurred at the end of eighteenth century and includes introduction of water-and steam- powered mechanical manufacturing facilities. The second stage includes electrically-powered mass production at the start of twentieth century whereas the third stage included use of electronics and IT to improve automation of manufacturing at the start of 1970s. Today,

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 15–22, 2020.
https://doi.org/10.1007/978-3-030-48021-9_2

we envision a fourth stage where several innovations of digital technologies are combined and exploited as an industrial internet. This includes innovations such as cloud computing, artificial intelligence technology, augmented reality and big data technology. In Norway the on-going research project CPS Plant is expected to apply these innovations and establish a CPS framework. Also, the project is expected to contribute to improved production performance both in manufacturing and production companies.

To support the ambition in Industry 4.0, the development of a successful asset management strategy implementation in CPS Plant should be of relevance. The importance of asset is also outlined in existing architectures for CPS such as Reference Architecture Model Industry 4.0 (RAMI 4.0) (IEC 2017). Also, similar architectures have been tested for maintenance application where sensors can support in better maintenance decisions (Lee and Bagheri 2015). Although these architectures are both described in standards and tested with case studies, it is still unclear how they relates to an asset management strategy.

It remains to investigate how CPS architectures can be build up with the asset management strategy based upon the definition described in ISO 55000 (ISO 2014). It seems that a successful asset management strategy should have aspects both as a plan and as a system. For the plan aspect, a roadmap for CPS should be further investigated. For the system aspect, elements from agile manufacturing (Pullan 2014) should be investigated due to the relevance as virtual enterprise.

The aim in this article is therefore to propose an asset management strategy implementation that will support implementation of CPS Plant in Norwegian manufacturing and production companies.

The future structure in this article is as follows: Sect. 2 presents relevant architectures for CPS and relevant application in maintenance. Section 3 elaborates foundations in asset management strategy and propose contribution to both the aspects of plan and systems in strategies. Concluding remarks are finally made in Sect. 4.

2 Cyber-Physical Systems and Maintenance

Cyber-physical systems (CPS) can be understood as "*integrations of computaions with physical processes.*" (Lee, 2008) To shape an enterprise model based on this definition a meta-model has been used such as RAMI 4.0 (IEC 2017) and has typical similarities with the meta-model GERAM (Myklebust 2002): Visualization in three dimensions and involving the life cycle of the enterprise. Also some frameworks have only considered two dimensions where the vertical and horizontal value chain is integrated (Geissbauer et al. 2014).

In the maintenance field, a one-dimensional architecture has been proposed with successful demonstrations (Lee and Bagheri, 2015, Rødseth et al. 2016). This architecture is denoted as the 5C architecture and is a step-by-step guideline for developing and deploying a CPS for manufacturing. Figure 1 illustrates this architecture which is relevant for maintenance. It starts with specifying the connection level and ends up with configuration level with self-maintenance and digital advices. For each level in the 5C architecture, description of predictive maintenance attributes is provided.

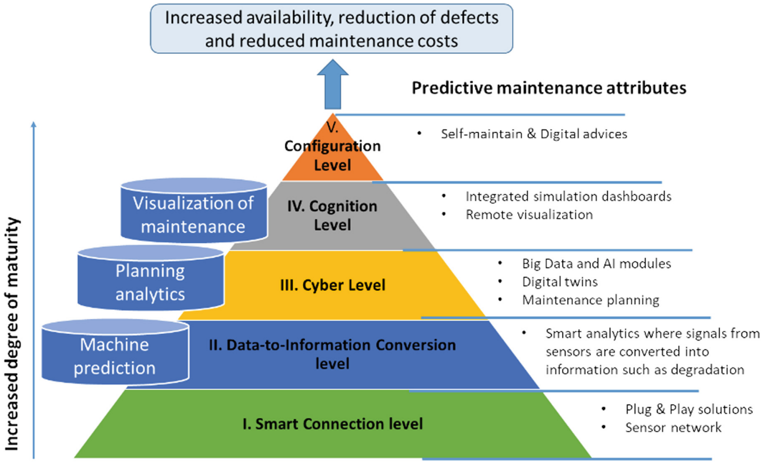


Fig. 1. CPS architecture proposed by Lee, adapted from (Lee and Bagheri 2015, Rødseth et al. 2016)

3 Asset Management Strategy Implementation

3.1 Asset Management Strategy

The term *strategy* stems from the Greek word *strategos* and means general in command of an army (Ronda-Pupo and Guerras-Martin 2012). In organisation theory there are several understanding of the term strategy. Yet, a definition of strategy has been proposed by Mintzberg et al. (1998) to be "*the pattern or plan that integrates an organization's major goals, policies and action sequences into a cohesive whole.*" It is also possible to explain a strategy in terms of a pattern where it is different forms of strategy. From ISO 55000, the notion strategic asset management plan is used (ISO 2014): *"Documented information that specifies how organisational objectives are to be converted into asset management objectives, the approach for developing* ***asset management plans****, and the role of the* ***asset management system*** *in supporting achievement of the asset management objectives."*

In this definition, strategy is considered both to be a plan and has a role as a system. Further in this section it is proposed how CPS both can has the role as a roadmap as well as a system in asset management strategy.

3.2 Asset Management Strategy as a Roadmap

Many different models of guidelines and roadmaps are already created for strategic use of CPS, it can be as follows:

1. Analyse your business, how far have you come in implementation and use of data. Which processes use data or provide data?

2. How far is digitalization achieved by data and ICT/infrastructure today? Your need to measure and map all areas of the organization – all the business layers. What kind of data do you have and what do you measure?
3. What is needed before everything will be digitized. Make an overview of what is happening, outline a "digital twin" (a shadow of architecture and systems) of your own organization.
4. Get understanding and culture for digital changes in your own organization. Identify and stimulate future scenarios, the skills and willingness to change.
5. Do we have the capacity and knowledge of this in our own organization? Look at the future opportunities and find the gap.
6. Predicting what's missing.
7. Fill the gaps, this will make the organization flexible for quick improvements.
8. Goal improvements will be achieved when you can use data from digital twins (shadows) in decision-making that is implemented and digitized.

The layers in RAMI 4.0 is the visual pictures of the machine or the systems of the assets in an organisation. In order to develop a digital model to be a digital twin it is important that it describes the physical product in total. A digital twin must contain a full description of other subsystems such as software and electronics. Based on the requirement specification, a functional model of the product is developed. With these things in place, you will have a description of how to make sure that the physical product will be. Some of the most important elements to call it a digital twin is the possibility to be able to analyse these models. The functions that is describing the model for how the product will look like in a 3D and how the physical functionality will be. These data will also be a part of the assets for the companies. We will therefore in the future factories see functional and dynamical assembly lines move more towards ad-hoc production network.

For successful asset management strategy implementation, some overall barriers should be controlled in the roadmap based on (Crombie 2016), see Table 1.

3.3 Asset Management Strategy as a System

Asset management has addressed a need for being agile due to turbulent environments faced by today's enterprise (Harris and Carapiet 2006). The need for agility has been classified into several categories such as marketplace, competition, customer requirements, technology and social factors, suppliers and internal complexity. The Agile methodology based on Scrum Master has been evaluated with respect to ISO 55000 where successful implementation is indicated to have benefits in terms of increased productivity and reduced costs, improved employee involvement, reduced time to market, higher quality, as well as improved satisfaction from stakeholders and employee (Crombie 2016).

Transferring the agility into the manufacturing domain leads to the concept agile manufacturing where industry is required to be able to produce products efficiently and respond to swift changes in market. With the focus on product development agile manufacturing is defined as "*the capability of an organisation, by proactively establishing virtual manufacturing with an efficient product development system…*" (Pullan

Table 1. Barriers that must be controlled in the roadmap based on (Crombie 2016)

Topics relevant for the road map	Possible barriers
Planning and decision-making level	○ Lack of strategic plans, knowledge about the stakeholder's needs and control procedures ○ Inconsistent decision criteria by the decision makers
Managerial and organizational	○ Poor management expertise in digitalisation ○ Lack of top management support in implementing CPS ○ Lack of communication channels within organisations and departments about future CPS plant ○ Lack of scope and job description for developing CPS
Information to/from resources	○ Lack of standard architecture for CPS ○ Lack of data about the implemented phases of the plans for CPS ○ Lack of shared knowledge in CPS implementation
Human resources	○ Lack of knowledge transfer between different demonstrators in CPS ○ No trained staff or lack of human resources in CPS
Social aspects	○ Departments unwilling to submit to CPS, but rather focus on traditional manufacturing systems
Financial resources and investment	○ Lack of recognising budget constraints for investing in CPS ○ Shortage of financial resources in investing in CPS

2014). It is further proposed by Pullan (2014) to have a coupling mechanism between modelling of the product itself and modelling of product development as process.

All though product development will have an important role in a virtual manufacturing, the whole value chain and the life cycle must be evaluated in a digital enterprise. In particular, both production and maintenance is specified in RAMI 4.0

(IEC 2017). The interaction between production and maintenance has also been elaborated in EN 16646 "Maintenance within physical asset management" (CEN 2014).

Figure 2 presents the proposed contribution towards a CPS based asset management system. It is inspired by the elements in agile manufacturing (Pullan, 2014) where the interaction (coupling) mechanism is between production and maintenance. In addition, this system includes the digitalisation layers in RAMI 4.0 which is connected to the physical asset. The chosen KPIs is based on earlier evaluations of CPS (Rødseth et al. 2016, Eleftheradis and Myklebust 2016) as well as the need for having an overall dashboard of KPIs that measuring the coupling mechanism between production and maintenance. This interaction is also denoted as integrated planning. As an example; overall equipment effectiveness (OEE) has been evaluated for how it is altered when maintenance ignores production and production ignores the equipment (Zuashkiani et al. 2011).

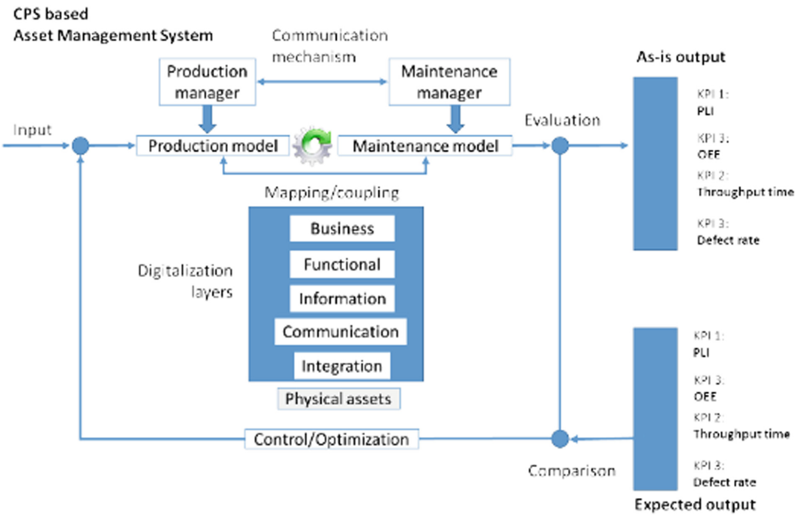


Fig. 2. Proposed contribution towards a CPS based asset management system inspired by (Pullan 2014)

Figure 2 could in CPS Plant be tested in the smelting industry where a specific equipment is evaluated with respect to predictive maintenance. In this demonstration, it is expected that information from sensors will support the computerized maintenance management system in the company dashboard with key performance indicators such as OEE.

4 Conclusions

The aim in this article has been to propose an asset management strategy implementation that will support implementation of CPS plant in Norwegian manufacturing and production companies. The strategy proposed in this article comprise both a roadmap as well as a contribution towards CPS based asset management System. This article concludes that all though the asset management strategy Implementation will be tested in Norwegian companies, it remains to evaluate how this strategy should be "adjusted" for similar implementation in other countries. All though this asset management strategy will support Norwegian manufacturing, further research should include other countries as well with relevant case studies in manufacturing and "adjust" it according to the specific country and industrial context. Further research in CPS Plant will require more testing of the strategy of several specific demonstrators in the project.

Acknowledgements. The authors wish to thank for valuable input from the research project CPS-plant.

The Research Council of Norway is funding CPS-plant.

References

CEN. EN 16646: Maintenance - Maintenance within physical asset management. Brussels (2014)

Crombie, A.C.: Agile asset management. In: Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015). Springer, Cham (2016)

Eleftheradis, R., Myklebust, O.: A guideline of quality steps towards zero defect manufacturing in industry. In: Proceedings of the 2016 International Conference on Industrial Engineering and Operations Management (2016)

Geissbauer, R., Schrauf, S., Koch, V., Kuge, S.: Industry 4.0 - Opportunities and Challenges of the Industrial Internet (2014)

Harris, H., Carapiet, S.: The need for agility in asset management. In: Proceedings of the 1st World Congress on Engineering Asset Management (WCEAM), 11–14 July 2006. Springer, London (2006)

IEC. IEC PAS 63088: Smart manufacturing - Reference architecture model industry 4.0 (RAMI4.0). Publicly Available Specification Pre-Standard. Swit- zerland, The International Electrotechnical Commission (2017)

ISO. ISO 55000 Asset management - Overview principles and terminology. Switzerland (2014)

Kagermann, H., Whalster, W., Helbig, J.: Recommendations for implementing the strategic initiative INDUSTRIE 4.0., Frankfurt, acatech – National Academy of Science and Engineering (2013)

Lee, E.A.: Cyber physical systems: design challenges. In: Proceedings - 11th IEEE Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing, ISORC 2008 (2008)

Lee, J., Bagheri, B.: Cyber-physical systems in future maintenance. Lecture Notes in Control and Information Sciences (2015)

Mintzberg, H., Quinn, J.B., Goshal, S.: The Strategy Process. Prentice Hall, London (1998)

Myklebust, O.: Enterprise Modelling Supported by Manufacturing Systems Theory. NTH, Trondheim (2002)

Pullan, T.T.: Decision support tool using concurrent engineering framework for agile manufacturing. Int. J. Agile Syst. Manag. **7**, 132–154 (2014)

Ronda-Pupo, G.A., Guerras-Martin, L.A.: Dynamics of the evolution of the strategy concept 1962–2008: a co-word analysis. Strateg. Manag. J. **33**, 162–188 (2012)

Rødseth, H., Schjølberg, P., Larsen, L.T.: Industrie 4.0 - a new trend in predictive maintenance and maintenance management. In: EuroMaintenance 2016 - Conference Proceedings. Artion Conferences & Events (2016)

Zuashkiani, A., Rahmandad, H., Andrew, K.S.J.: Mapping the dynamics of overall equipment effectiveness to enhance asset management practices. J. Qual. Maintenance Eng. **17**, 74–92 (2011)

Drilling Performance Management Through Reliability-Based Optimization

Omer Faruk Ugurlu(✉), Burak Ozdemir, and Mustafa Kumral

McGill University, Montreal, QC, Canada
{omer.ugurlu, burak.ozdemir}@mail.mcgill.ca,
mustafa.kumral@mcgill.ca

Abstract. Mineral commodity prices have declined in recent years. This decline forces mining companies to find effective cost management strategies to sustain their operations. Otherwise, many operations will be ceased or suspended. Effective equipment utilization has strong potential to reduce operational costs. Furthermore, unexpected events or failures during the operation may not properly be considered in production scheduling. These contingencies affect the subsequent production process and cause operational delays; hence, the operation cost increases. Equipment condition is a key element to reach the desired production rate. This paper focuses on optimization of reliability parameters to improve the performance of the machine associated with its condition and calculation of drill bit consumption accordingly. Multiple input factors, such as operating parameters, operation time and maintenance time, were considered and controlled simultaneously to simulate drilling operation by stochastic modeling technique, using historical data. A case study was carried out using discrete event simulation (DES). Multiple simulations were used to quantify risk. The research outcomes show that the proposed approach can be used as a tool to assist in production scheduling and asset management.

1 Introduction

The operating and the financial successes of a mining company are extremely dependent on the condition of its essential assets such as trucks, loaders and drilling machines. These critical assets have a huge impact on capital investment, operating cost, maintenance, risk and performance of the operation (Hastings et al. 2010). Since asset management has great importance for mining activity, it should be taken into consideration as one of the main concerns for a mining company to reach desired economic and technical goals.

Drilling is the primary operation for mining activities. Since it can be used for a variety of purposes such as exploration, evaluation, and extraction of valuable materials in earth's crust, it is a significant component of the operating cost. The main part of drilling operation cost comprises drill bit consumption. Hence, drill bit monitoring and asset management of drilling equipment are crucial aspects of mining activity, and efficient equipment utilization has strong potential to reduce cost and improve the operation. More information about drilling operation and operational parameters can be found at Ugurlu and Kumral (2016b, 2020a, 2020b and 2020c).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 23–29, 2020.
https://doi.org/10.1007/978-3-030-48021-9_3

The condition of the assets can be monitored by reliability analysis which mostly depends on the system failure data. Reliability analysis is essential to (1) ensure the condition of the equipment for effective equipment utilization and (2) forecast unexpected delays. Drilling machines are repairable systems which can be restored after a failure for satisfactory operation. Therefore, the system is repaired after any failure to perform desired performance. Times between failures and times to repairs are needed to characterize system reliability. More information about reliability analysis can be found at (Ozdemir and Kumral 2017, 2019) and (Ugurlu and Kumral 2016a).

Different strategies can be constructed to maximize equipment reliability and extend equipment life at the lowest cost. Simulation techniques can be used to imitate future events and show possible operating conditions. Since the mining environment is highly uncertain and operational conditions continuously change, deterministic solutions provide insufficient and misleading information. To overcome the hurdle of stochasticity of mining problems, discrete event simulation (DES) technique can be used. It is a stochastic mathematical modeling technique for discrete and probabilistic conditions (Shqair et al. 2014) and drill- ing operation includes discrete events such as moving from one drill hole to another and leveling the machine on the ground for stability while drilling. DES has been used by different researchers for a variety of purposes. For instance, Yarmuch et al. (2017) used DES to construct a production schedule. Also, Yuriy and Vayenas (2008) researched to investigate maintenance action of mining operations by DES model. Furthermore, Ozdemir and Kumral (2018) developed an agent-based simulation model for open pit mines to verify production targets regarding tonnage and grade. Moreover, Botín et al. (2015) carried out a study to minimize the highest risk parameters in a block caving project.

This paper presents a simulation-based optimization approach to improve drilling machine performance associated with its condition. Historical data was used to analyze the reliability of the machine. Reliability parameters (λ and β) were optimized by performing multiple simulations in Arena Software. The actual and the optimum condition were compared. The results show that proposed approach can be used to improve the quality of the maintenance action and asset management strategy.

2 Research Methods

The performance of the drilling equipment can be measured by reliability analysis. Since the drilling machines are repairable systems, the mean time between failure (MTBF) follows a Non-Homogeneous Poisson Process. Hence, the relation between failures can be modeled by power-law relation as it is shown in Eq. 1 (Tobias and Trindade 2011).

$$M(t) = \lambda * t^{\beta} \tag{1}$$

where M is the cumulative number of failures, and λ and β are the shape and the scale parameters, respectively. More information can be found at (Ugurlu and Kumral 2020b) and (Ozdemir and Kumral 2019).

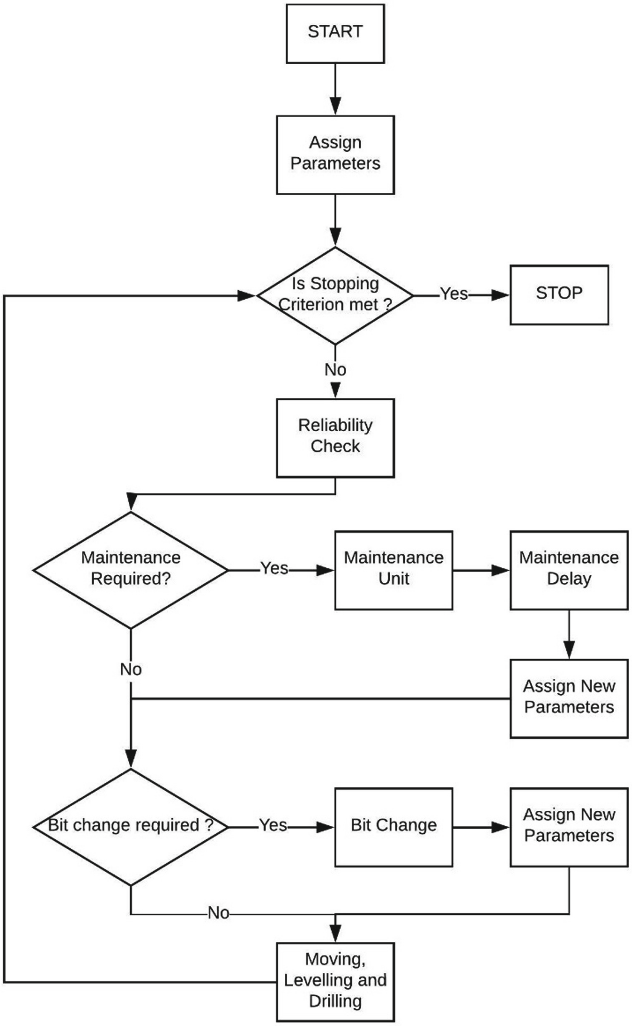


Fig. 1. Simulation model of drilling operation

As seen, the λ and β parameters have a direct effect on the availability of the machine which decides the number of the drilled hole. Moreover, these parameters can be controlled by the quality of the maintenance activity.

The drilling operation consists of a combination of discrete events (such as moving, leveling, drilling, etc.) and includes many uncertain variables (such as drilling parameters, drilling time and maintenance activity, etc.). Thus, a simulation-based optimization technique is used to optimize reliability parameters.

The simulation model of the drilling operation is represented in Fig. 1. The parameters are assigned to the equipment at the beginning of the simulation. When the reliability of the equipment is assessed, it is decided whether the equipment needs maintenance or not. If the equipment needs maintenance, it is sent to the maintenance unit. Hence, a delay is occurred due to the maintenance activity. After the maintenance, new parameters are assigned to the equipment. Then, the bit is checked if it is needed to be changed. The historical bit usage times were analysed and represented in a probability distribution. After each bit change, a bit usage time is assigned to the bit. If the time passed after the last bit change time is bigger than the generated bit usage time, the bit is changed and there is a delay in the bit change. This delay is also generated from a probability distribution based on field data. Otherwise, it directly moves to the drill hole, and start the operation with leveling and drilling. This cycle continues during the drilling operation. The drilling parameters (such as bit wear, weight on the bit, rotation speed and bailing air pressure) are randomly generated from a probability distribution based on field data.

3 Case Study

Throughout the research, a case study was conducted in an open-pit mine. The drilling operation of the mine is modelled in Arena simulation software. In the simulation model, the drilling time of one drill hole is calculated as a function of bit wear (*BW*), rotation speed (*RPM*), weight on the bit (*WOB*), and bailing air pressure (*BAP*). It was formulated by regression analysis of field data. As a result of regression analysis, the drilling time (*d*) of one drill hole can be calculated by Eq. 2 (Ugurlu and Kumral 2020c). The equation shows that bit deterioration increases the time required to drill a hole. On the other hand, as the machine operating parameters increase, the time required to drill a hole decreases. To quantify the influence of bit wear, time series regression analysis was conducted on time series data sets with machine operating parameters. More detailed information can be found in (Ugurlu and Kumral 2020c). The interactions among parameters are not statistically significant. Hence, they are ignored.

$$d = 129.68 + 1.55 \times BW - 0.78 \times RPM - 0.12 \times WOB - 0.33 \times BAP \quad (2)$$

This equation was used in the simulation model to calculate the drilling time of one drill hole. Since the parameters are uncertain, a probability distribution was selected for each parameter depending on the field data. Table 1 shows the selected distributions for each parameter. In each hole, random numbers were generated from the selected distributions and drilling times were calculated from the regression equation (Eq. 2).

Table 1. Distributions of uncertain variables

Parameters	Distribution
RPM (rpm)	PERT (64,66,72)
WOB (MPa)	PERT (405,415,440)
BAP (bar)	PERT (365,390,415)
Moving (min)	Normal (1.08,0.43)
Leveling (min)	Lognormal (1.18,0.45)

In addition to drilling parameters, the moving and leveling times are uncertain for each hole. Similarly, probability distributions were selected for these times. The field data show that the moving and leveling times follow normal and lognormal distributions, respectively.

The discrete event simulation model is formed, and all distributions are selected in the simulation model. The initial reliability parameters were obtained from the reliability analysis of the equipment by using field data. The parameters were used as input in the simulation model to estimate the number of drilled hole and bit usage in an uncertain mining environment. Moreover, an optimization problem was formulated in OptQuest which is the optimization module of Arena simulation soft- ware. The module performed a simulation-based optimization by changing the reliability parameters (λ and β) to maximize the number of drilled holes while minimizing the drill bit consumption. Once 100 scenarios were generated for a drill- ing machine, the number of drill holes and the number of usable bits were calculated. Stopping criteria were assigned as 90 days because of eliminating the seasonal effect. The initial and optimum values are given in Table 2. The simulation results show that the optimum reliability parameters increased the number of drilled hole by 234 while they decreased the number of the used drill bit by two. In other words, the number of drilled holes per bit increased from 43 to 49 with the optimum parameters.

Table 2. Initial and optimized values

Parameters	Initial values	Optimum values
λ	0.008712	0.000583
β	1.112	1.028
The number of drilled hole	2,125	2,359
The number of used bits	50	48
The number of the drilled hole per bit	43	49

Figure 2 presents the difference between the actual and optimum reliability curves. As seen, the reliability much decreases slower in the optimum reliability curve which provides higher performance for the drilling machine. The optimum reliability parameters can be further enhanced by increasing the quality of the maintenance activity.

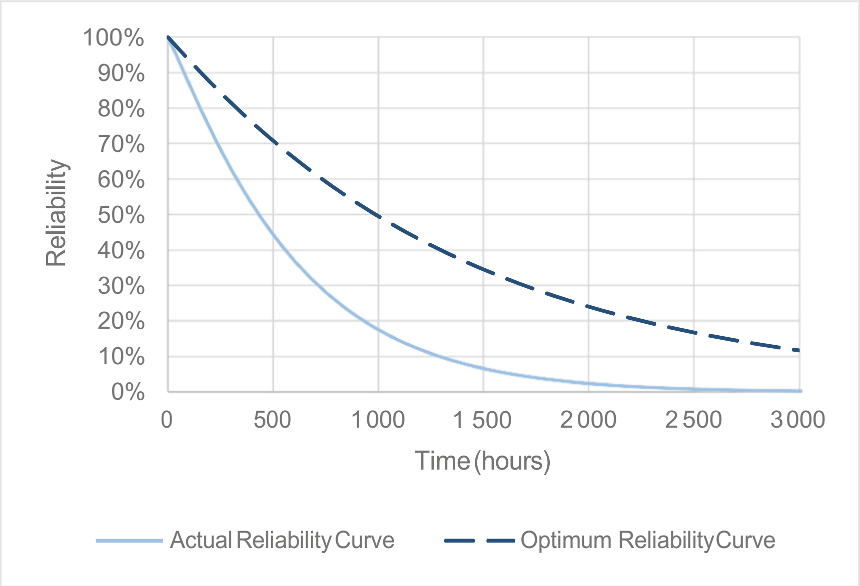


Fig. 2. Actual and optimum reliability curve

4 Conclusion

This paper presents a simulation-based optimization technique to optimize reliability parameters of drilling asset. Reliability analysis was performed to characterize the behavior of a drilling machine by a power-law relation with historical data. Also, simulation parameters and drilling time equation were determined based on field data. Once the simulation model was completed, it was run 100 times for three months of operation. It means that 100 different observations were obtained about the number of drilled hole and drill bit usage under uncertain mine environment. As a result of 100 simulations, 2,125 drill holes can be drilled by using 50 drill bits. Moreover, the reliability parameters were optimized by OptQuest optimization module of Arena simulation software. The module found the optimum reliability parameters which make the drilled hole maximum and used drill bit minimum by changing reliability parameters in each simulation. The drilling machine can drill 234 drill holes more by using two drill bits less with the optimum reliability parameters. In other words, the number of the drilled hole per bit was increased by 14%. The results show that the performance of a drilling machine can be improved while the consumption of the assets (drill bits) can be decreased by the operator experience and maintenance quality. Maintenance quality and the experience of the operators can be improved by training to boost the condition of the drilling assets. The principles of Reliability Centered Maintenance (RCM) can be used to manage the risk of failure of the assets at the right frequency. Also, interdepartmental relations between operation, maintenance, and other areas should be improved.

References

Arena 15.10. Rockwell Automation Inc, USA

Botín, J.A., Campbell, A.N., Guzmán, R.: A discrete-event simulation tool for real-time management of pre-production development fleets in a block-caving project. Int. J. Min. Reclam. Environ. **29**(5), 347–356 (2015)

Hastings, N.A.J.: Physical Asset Management, vol. 2. Springer, Heidelberg (2010)

Ozdemir, B., Kumral, M.: Stochastic assessment of the material haulage efficiency in the earthmoving industry. J. Constr. Eng. Manag. **143**(8) (2017)

Ozdemir, B., Kumral, M.: Analyzing human effect on the reliability of mining equipment. Int. J. Heavy Veh. Syst. **26**(6), 872–887 (2019)

Ozdemir, B., Kumral, M.: Appraising production targets through agent-based Petri net simulation of material handling systems in open-pit mines. Simul. Model. Pract. Theory **87**, 138–154 (2018)

Shqair, M., Altarazi, S., Al-Shihabi, S.: A statistical study employing agent-based modeling to estimate the effects of different warehouse parameters on the distance traveled in warehouses. Simul. Model. Practice Theory **49**, 122–135 (2014)

Tobias, P.A., David, T.: Applied Reliability. CRC Press, Boca Raton (2011)

Ugurlu, O.F., Kumral, M.: Analyzing drilling machine in aid to improve open pit mining operations. Rock Mechanics and Rock Engineering: From the Past to the Future, p. 171 (2016a)

Ugurlu, O.F., Kumral, M.: Optimization of drill bit replacement time in bench drilling. In: 24th World Mining Congress Rio de Janeiro, Brazil, 18–21 October (2016b)

Ugurlu, O.F., Kumral, M.: Cost optimization of drilling operations in open-pit mines through parameter tuning. Qual. Technol. Quant. Manage. **17**(2), 173–185 (2020a)

Ugurlu, O.F., Kumral, M.: Reliability-based performance analysis of mining drilling operations through Markov chain Monte Carlo and mean-reverting process simulations. Simulation (2020b). https://doi.org/10.1177/0037549720923751

Ugurlu, O.F., Kumral, M.: Management of drilling operations in surface mines using reliability analysis and discrete event simulation. J. Fail. Anal. Prev. (2020c, in press)

Yarmuch, J., Epstein, R., Cancino, R., Peña, J.C.: Evaluating crusher system location in an open pit mine using Markov chains. Int. J. Min. Reclam. Environ. **31**(1), 24–37 (2017)

Yuriy, G., Vayenas, N.: Discrete-event simulation of mine equipment systems combined with a reliability assessment model based on genetic algorithms. Int. J. Min. Reclam. Environ. **22**(1), 70–83 (2008)

An Integrated Asset Management Strategy for Transmission Line Conductors, Based on Robotics, Sensors, and Aging Modelling

Nicolas Pouliot[1(✉)], Gilles Rousseau[1], Daniel Gagnon[1], Denis Valiquette[1], André Leblond[2], Serge Montambault[1], and Dragan Komljenovic[1(✉)]

[1] Institut de recherche d'Hydro-Québec (IREQ), Hydro-Québec, Varennes, Canada
{Pouliot.Nicolas, Rousseau.Gilles2, Gagnon.Daniel, Valiquette.Denis, Montambault.Serge, Komljenovic.Dragan}@ireq.ca

[2] TransÉnergie, Hydro-Québec, Montreal, Canada
Leblond.Andre.2@hydro.qc.ca

Abstract. Powerline aging and associated refurbishments are undoubtedly an upcoming challenge for a lot of utilities. At Hydro-Québec, this motivated the launch of an integrated strategy aiming at providing the tools, technologies, and know-how to efficiently determine a health index for conductors without relying heavily on field sampling methods. The paper explains the strategy, which unfolds in three different projects. At the core of the strategy are several robotic platforms, including drones, which were developed at IREQ to reach and travel along the conductors. Mounted on the device is a portfolio of Non-Destructive Testing (NDT) technologies, developed to detect symptoms of prevalent aging mechanisms. Secondly, the conductor degradation and aging modelling project is aimed at determining the actual health index of the conductors, the rate of degradation, and their remaining useful life (RUL). Finally, a database tool is to be developed to provide seamless integration with enterprise databases. Vertical integration will thus be achieved to provide tools for supporting decision-making in asset management.

Keywords: Asset management · Non-Destructive Testing (NDT) · Robotics · Power transmission lines · Degradation · Aging model

1 Introduction

Power grids are the backbone of most economies and greatly contribute to our quality of life. However, overhead transmission lines (OHTLs) are aging, while availability and reliability requirements are growing more and more stringent. OHTL aging, along with the associated refurbishment programs, is undoubtedly an upcoming challenge that will require proper condition assessment and the definition of relevant end- of-life criteria. At Hydro-Québec TransÉnergie (HQT), the transmission division of Hydro-Québec, the conductor is considered the most important line component, ac- counting

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 30–38, 2020.
https://doi.org/10.1007/978-3-030-48021-9_4

for about 30% of transmission line reconstruction cost. Its actual condition is essential for following reasons:

- It is a key parameter for optimal maintenance decision-making;
- It drives the decision to build a new OHTL or to invest further in the old one.

HQT operates many lines aged over 60 years in which the most common type of conductor is the Aluminum Conductor Steel Reinforced (ACSR). Conductor condition assessment still remains a challenge because of uncertainties related to the characterization and understanding of degradation mechanisms, such as corrosion, fatigue, thermal aging, etc. There is a need to define the remaining tensile strength of conductors and to diagnose their exact condition.

For these reasons, HQT needs methods, tools and technologies for diagnosing conductor condition from sensor measurements using non-destructive testing (NDT). The long-term objective is to predict the conductor's remaining useful life (RUL) without having to take field samples. Assessment of aging OHTLs requires expertise, know- how and experience in a number of fields and technologies in order to provide accurate and reliable results. One innovative approach that is starting to make inroads in providing results is the application of robotics together with NDT technologies.

At Hydro-Québec, the above context motivated the launch of an integrated asset management strategy aiming at providing the tools, technologies, and know-how to efficiently determine conductor health index without relying heavily on conventional field sampling. The paper aims to describe this asset management strategy, which unfolds in three different projects and draws on the specific expertise of Hydro-Québec's research institute, IREQ.

2 An Integrated Asset Management Strategy

The asset management strategy incorporates fundamental and applied research in the field of transmission lines, and has three main objectives:

- Objective 1: Diagnosis of the conductor's condition and prognosis of its RUL without taking samples;
- Objective 2: Diagnosis of the RUL of other transmission line components;
- Objective 3: Holistic asset and overall performance management.

The third objective also includes improving the existing decision support tools. The relationship between strategy activities and objectives is shown in Fig. 1.

3 Conductor Aging Modelling

The OHTL degradation and aging modelling project is aimed at determining the health condition and RUL of conductors. Most overhead conductors on the Hydro- Quebec transmission grid are ACSR (Aluminum Conductor Steel Reinforced); a typical cross-section can be seen in Fig. 2.

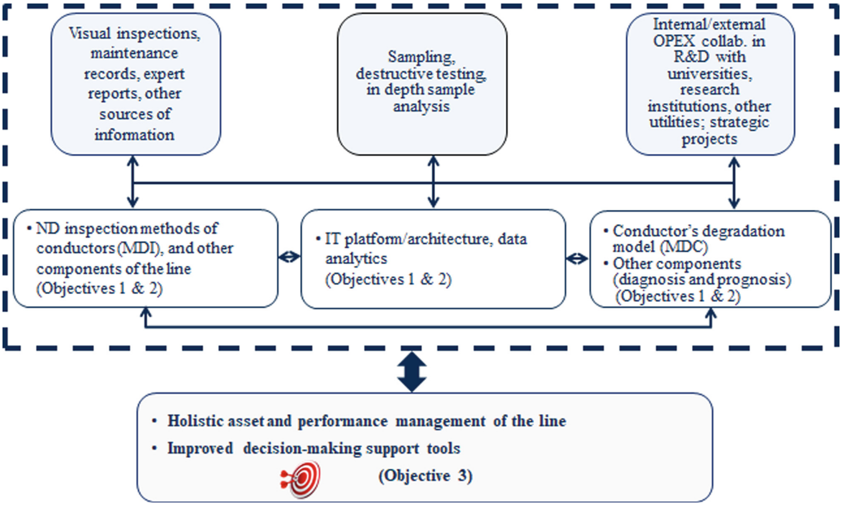


Fig. 1. Relationship between strategy activities and objectives.

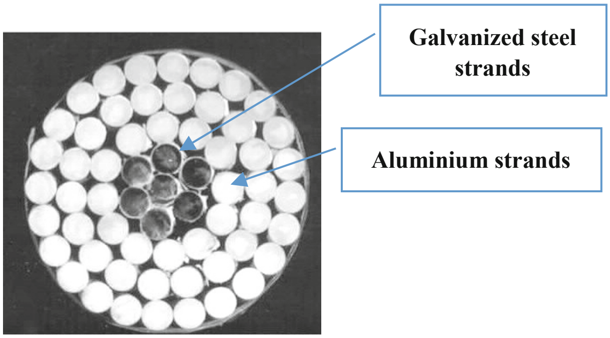


Fig. 2. Typical ACSR conductor cross-section.

A reliability and degradation model has to take into account all the relevant factors to fully assess reliability and RUL. In the case of a powerline, conductors are only one of the many components that comprise the power system but are of great importance in terms of cost and function. The first step was therefore to identify the primary conductor degradation mechanisms, as listed in Table 1. These mechanisms had to be quantified with proper sensors and lab measurements, in order to feed the aging model.

Conductor aging can vary greatly depending on the environment, type and size of conductor, maintenance history, and electric load. Conductors are subjected to a variety of degradation mechanisms, the most damaging being mechanical fatigue, corrosion, and ice/snow accretion. Some mechanisms can have an impact on both performance and residual life whereas others will affect only performance, for example by reducing surface emissivity. While both types are important to assess, the first bears more

Table 1. Degradation mechanisms and their effects on overhead conductors

Degradation mechanism	Failure mode	Failure cause	Effect on conductor
Mechanical fatigue	Broken strands	Wind-induced vibrations, ice/snow accretion	Loss of mechanical strength and transit capacity
Surface degradation	↑ Surface absorptivity/emissivity	Environmental/pollution	Loss of electric capacity/overheating
Corrosion	Broken strands/loss of cross-section	Environmental/pollution	Loss of mechanical strength and transmission capacity
Lightning	Broken strands/loss of cross-section	Environmental	Loss of mechanical strength and transmission capacity

importance in the current study since it has an impact on reliability. Also to be considered are the strategic ranking of the circuit and civilian liability in the area.

As a powerline passes through different environments, the conductor aging model under development has to take into account the geographical variability of degradation mechanisms and their cumulative effects. Figure 3 shows an example of how such variability can be treated.

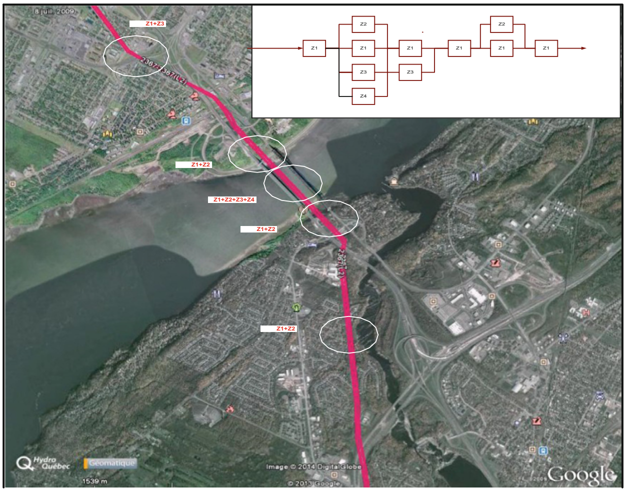


Fig. 3. Influence of different degradation mechanisms along powerline path.

Each section of the powerline is given a different probability of degradation by the various mechanisms depending on its location. In urban areas, pollution and salt spray corrosion will be an issue which may not necessarily be present in a rural area.

Degradation mechanism probabilities can be treated as either series or parallel reliability factors. In the model, the various mechanisms are being treated as possibly collaborating in failure, and this aspect is presently being investigated by laboratory tests. One example of such collaborative failure is the influence of having simultaneous corrosion and mechanical fatigue in the same area.

4 Data Collection Using Robotics and Sensors

Hydro-Québec's asset management strategy requires the gathering of field data to evaluate actual asset condition. In the case of powerlines, accessibility and harsh environments constitute a real challenge that calls for advanced methods, including robotics. The core of the strategy therefore consists of robotic platforms that were developed at IREQ to reach and travel along conductors while energized. Also to be mounted on these platforms is a portfolio of Non-Destructive Testing (NDT) technologies developed to match each of the identified conductor aging mechanisms.

4.1 Powerline Robotic Technologies

The introduction of remotely operated robotic systems provided new, complementary approaches for traveling along power lines while gathering visual cues that can be transmitted to the ground and archived for future viewing and comparison. That already provided a paradigm shift, allowing human operators to stay away from the most challenging areas like steep-slope spans or damaged conductors/overhead ground wires, while generally improving efficiency and avoiding outages often required for human work on power systems (Li 2016), (Menendez 2017).

Energy-wise, the most efficient way to travel along a powerline is through the use of wheels directly mounted on the conductors. The first system developed, called LineROVer, is a simple, remotely operated trolley. (See Fig. 4) It is limited to a single span, but is able to roll over splices and provides high traction force (>600 N), and can work on energized conductors up to 315 kV. It can carry an HD camera and a few optional maintenance sensors and tools for temporary repairs.

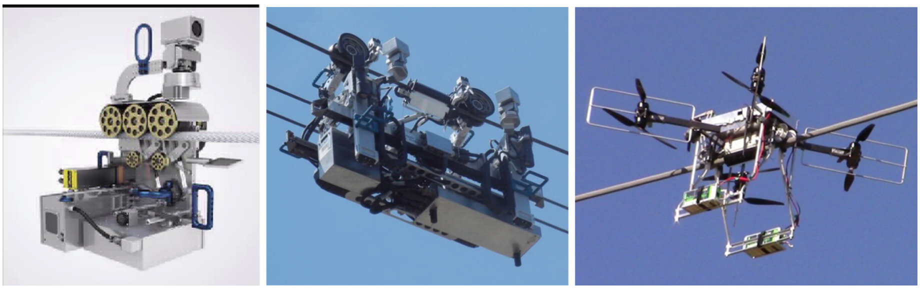


Fig. 4. Robots developed at IREQ for powerline inspection and maintenance: LineROVer (left), LineScout (center) and LineDrone (right).

The next system developed, LineScout, is a ruggedized and versatile moving platform able to support large payloads (up to 12 kg). It can cross obstacles up to 760 mm in length, such as suspension clamps, spacers, and aerial markers. This feature makes it one of the only systems capable of performing inspection or maintenance on several consecutive spans, or on bundled conductor spans where spacers have to be crossed every 40 m or so. Both LineROVer and LineScout have been used extensively on energized and deenergized powerlines in utility grids around the world. They have been the topic of several publications in both robotics and power utility forums (Montambault 2013; Montambault 2014).

In recent years, it has become clear that drones will play a key role in powerline maintenance tasks detailed inspection, general surveying, and LIDAR surveys, etc. An IREQ research team is now focusing on more specific and challenging applications, for example using drones to string conductors between towers, and even other applications that involve actual contact with live line components. Custom drones have been developed that are capable of landing on a live transmission line and bring NDT sensors on them to gather strategic data. This maneuver is simplified by the use of an on-board cable detection system that actually steers the drone towards the conductor so that it lands on it. Custom designed motor-wheels roll along the line and inspect it over long distances while consuming very little energy.

4.2 Powerline Sensors

The robotic devices featured in the previous section are equipped with HD cameras that are needed for visually spotting areas of interest such as broken strands due to lightning strikes, birdcages, areas where abrasion has occurred, or even sections darkened by contaminants or surface reactions to the environment. When it comes to assessing the actual *internal condition,* one can rely on conductor sampling but this tends to be costly and time-consuming, and it only provides information about the sampled location - often, whatever section is most accessible. To overcome these limitations, Non-destructive Testing (NDT) technologies were developed for each of the main aging mechanisms. The various technologies are described below and seen on Fig. 5.

Fig. 5. Sensors developed at IREQ for powerline inspection: LineCore (left), LineScan (center) and LineOhm (right).

4.2.1 Steel Galvanic Protection Assessment by LineCore

The need for a lightweight sensor that can probe inside the conductor to evaluate the condition of the zinc coating was addressed by developing the LineCore, based on eddy current technology. The LineCore sensor is an encircling solenoid that can be split open to facilitate installation on a conductor or - since the opening of the doors is motorized– to cross a hardware component. The eddy current electronics are mounted below the sensor and include a WiFi transmitter that sends the data to the ground. The base station runs software that controls sensor calibration and data acquisition. What is peculiar to the LineCore is that the software can analyze a wide range of ACSR conductors; there is no need to possess a sample of the ACSR before the inspection. An overlay is automatically computed to grade the LineCore signal according to five categories: brand new, possible zinc loss (13–20 µm), partial zinc loss (5–13 µm), and severe zinc loss (steel strands have bare spots), and of course aluminum or steel strand corrosion. LineCore is thus able to evaluate the condition of the galvanic coating along a span and to identify regions with complete loss of galvanic protection–or worse, with incipient corrosion of the aluminum and or steel strands (Lavoie 2018).

4.2.2 Strand Internal Break Detection by LineScan

Strand breaks mostly occur at the exit of a suspension clamp between the U-bolt and the last point of contact. The breaks are usually located between the outside layer and the clamp, or even more often -between subsequent inside layers. The difficulty here is that broken strands are not visible even if the clamp is removed. To solve this problem, a portable, low radiation X-ray delivery system, dubbed LineScan, was developed. Given that the system would have to be handled at the top of a tower by line workers, it needed certain features: low radiation leakage, exposure monitored by a ground crew, fast and easy to install on single conductor or on a double or quad conductor bundle. Review and demonstration of the commercially available systems narrowed the choice to a compact and lightweight pulsed spark-gap X-ray generator with a small and ruggedized digital radiography (DR) X-ray detection panel. The total weight of the X-ray system is 7.5 kg. Following successful deployments of LineScan, it was deemed technically desirable to improve the system by developing a robotic arm and installing it on the LineScout robotic platform. Careful review of the design constraints made it possible to derive a practical solution for flexible positioning of the DR panel and generator in most situations. The implementation is in its final testing phase and should become available in 2018.

4.2.3 Defective Joint Assessment by LineOhm

Last but not least among the sensors developed at IREQ is LineOhm, a lightweight and compact electrical resistance probe for assessing ACSR joint quality. It measures electrical resistance (in micro-Ohms) on a length of conductor, and then compares this with the resistance measured on the same length but over the joint. The ratio of these two measurements is highly correlated to the electrical quality of the joint, since the resistivity of the joint should be no greater than that of the conductor. To evaluate resistance, LineOhm measures the exact voltage drop between the two electrodes in contact, as well as the AC current flowing in the line. The electrodes are designed to ensure full and even electrical contact with the ACSR strands. The sensor is compact and weighs about 500 g, making it easy to mount on robotic platforms, including drones.

4.3 Field Data Gathering Surveys

While Hydro-Québec has been using robotic tools on powerlines since 2000, the goal was initially to conduct visual inspections and small maintenance tasks. Since 2014, however, field deployments have been evolving towards mostly data gathering campaigns. Initially, they were focused on a specific purpose, like assessing corrosion inside the conductors of long river crossings, or looking for inner-layer broken strands on certain lines subjected to heavy vibration-induced fatigue. These surveys will likely become more and more frequent, as longer-term decisions need to be taken about whether or not to refurbish some of the older lines (Fig. 6).

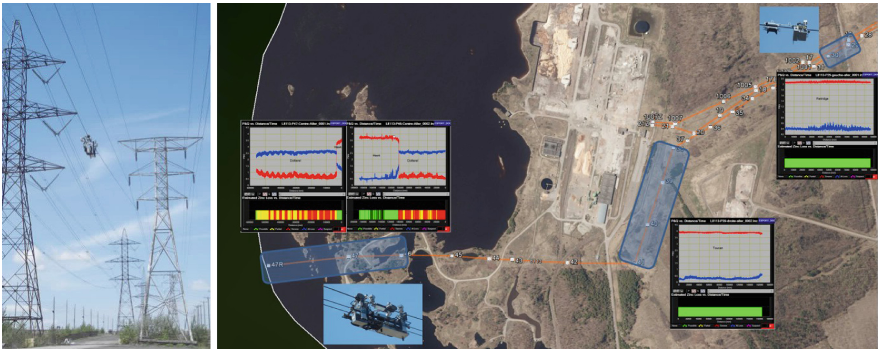


Fig. 6. Field measurement of galvanic protection layer by using LineScout and LineCore, enabling assessment of several spans in a single day.

5 Integrated Database

As previously discussed, knowledge of the current and future condition of transmission line components is of critical importance, especially in the case of conductors. Diagnosis and prognosis are greatly facilitated by field data gathering and the development of aging models. This requires the archiving and processing of several types of data: historical data, field surveys, non-destructive probe data, corroboration with destructive test analyses of samples, etc.

At the moment, various databases within the company contain raw data gathered during different phases of the life cycle: construction, maintenance, operation, weather events, various analyses, maintenance records, etc. An integrated information system will need to be developed and deployed to provide seamless integration between all these data, including the results of the conductor aging model and actual reports filed by field technicians. The creation of such a knowledge base will enable the simultaneous use of all these data to develop and validate degradation models and to know the state of degradation of components and their deterioration over time. The data will be structured to allow multidimensional and geographic analyses, operational solicitations, and effects of maintenance and inspection results. In addition, the relationship with the climatic environment, contaminants and soil types will help refine the models. Furthermore, developing processing tools will allow experimenting and inferring the health of the different portions of lines according to their component degradation models, as well as the impact of maintenance actions, inspection and asset replacement (see Fig. 1).

6 Conclusion

Transmission asset maintenance is an ongoing challenge for electrical utilities. Hydro-Québec has undertaken a large research project aiming to develop a comprehensive asset management strategy for its transmission lines. Various robotic and non-destructive technologies have been developed for this purpose. The strategy will allow vertical integration in order to provide tools for supporting decision-making in asset management. Analytical tools to simulate a transmission line's state of degradation and performance as a whole will improve existing decision support tools and develop new ones to guide decisions on medium- and long-term investment, while complying with risk, performance and cost-balancing criteria.

References

Li, L., et al.: A state-of-the-art survey of the robotics applied for the power industry in China. In: Proceedings of 2016 4th International Conference on Applied Robotics for the Power Industry (CARPI), pp. 1–5 (2016)

Menendez, O., Cheein, F.A.A., Perez, M., Kouro, S.: Robotics in power systems: enabling a more reliable and safe grid. IEEE Ind. Electron. Mag. **11**(2), 22–34 (2017)

Montambault, S., Pouliot, N., Hannon, M.: Robotic Repair of Overhead Ground Wires Using Linescout Technology on National Grid's Network. In: B2-040, Symposium Cigré 2013. Auckland, New-Zealand (2013)

Montambault, S., Pouliot, N.: Hydro-Québec's power line robotics program: 15 years of development, implementation and partnerships. In: Proceedings of 3rd International Conference on Applied Robotics for the Power Industry (CARPI 2014), 14–16 October, Foz do Iguassu, Brazil (2014)

Lavoie, E., et al.: LineCore Technology: Probing Below the Surface, Transmission and Distribution World Magazine, pp. 38–41 (2018)

Developing a System Vision of Assets at Elia, the Belgian Transmission System Operator

Suzanne Guérard(✉), Stéphanie Hammer, and Wim Michiels

Elia Asset, Quai Léon Monnoyer 3, 1000 Brussels, Belgium
suzanne.guerard@elia.be

Abstract. This paper is about the latest developments to support risk management of assets at Elia, the Belgian transmission system operator operating the HV grid from 30 to 380 kV. The story of risks management within the company started more than 5 years ago, when a methodology was developed to assess the criticality for continuity of supply of individual assets. The output of this methodology is still used nowadays. As an example, maintenance policies do take into account the criticality of the asset on the network to define its maintenance frequency. So do replacement decisions. However, as many other TSO's, Elia is also facing new challenges e.g. the energy revolution is pushing the ageing assets more and more to the limits. Scarcity in resources (both financial as in workforce) forces Elia to make sound and good decisions and prioritize actions despite uncertainties and scarceness of data. Risk management should help tackle these problems. In more concrete terms, there is a need for better addressing many diverse questions such as dealing with system redundancy, with the mutual impact of assets, with the impact of low voltage assets (protections) on high voltage assets, better assessing load at risk, asset end of life, the impact of a failure on several risk dimensions, the efficiency of risk mitigation measures, etc. In the frame of this paper, it was decided to focus on the solutions found to better assess the impact of protections on high voltage assets and the interaction between assets. To tackle these topics, one could think of implement- ing sophisticated methods such as probabilistic assessment, but this in turn requires asset reliability data, resources, skilled persons and lots of time for development, which are all sparse quantities. Therefore progress was made with pragmatic 80/20 methodologies, the 80/20 methodology is a free interpretation of the Pareto principle, stating that 80% of the results and valuable information come from 20% of the efforts which though not perfect, allows delivering valuable information. The corresponding developments are discussed in the frame of this paper.

1 Introduction

Elia is the Belgian transmission system operator (TSO). It operates over 8000 km of lines and underground cables throughout Belgium. The voltage levels concerned range from 30 to 380 kV, what makes from Elia both a TSO and a distribution sys- tem operator (DSO). As many other key players in the interconnected electricity system, Elia is faced to ageing asset fleets, in particular at medium voltage levels (DSO role). TSO's often pay an attention to their ageing assets which is in some ways correlated

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 39–45, 2020.
https://doi.org/10.1007/978-3-030-48021-9_5

with their net book value. It is often assumed that redundancy is used in the protection schemes of our primary assets and secondary assets may not be further investigated. Many TSOs are satisfied to know this redundancy is available.

Elia wants to investigate further this topic because amongst the energy not supplied (ENS) events which cause is not related to an intrinsic risk or a third party, it has been observed that system response plays a significant role. The particular case of automatisms like reclosers or automatic transfer switches, whose correct functioning helps in limiting the duration of interruption, is worth emphasising. There- fore, it follows logically that the analysis of combinations of primary and secondary ageing assets should be made, besides the analysis of ageing of individual primary assets. As a concrete illustration of that, let us take the case of a two extremity cable circuit protected by redundant protections. Let us assume some repairs of the cable have already taken place following local jacket degradation which in turn led to insulation degradation. During the previous failures, the cable protections were able to selectively eliminate the fault by sending an appropriate trip signal to the circuit breakers. However, these protections are now approaching their retirement age and the technology type of these protections, which is electronic, does not allow their condition to be monitored. All this, combined to the high ionic content of air caused by the location in a coastal area and slow degradation of the barrier between electronic and their environment has led to a common failure mode of these protections which is not detectable. Once the next cable failure occurs, these protections will be unable to answer adequately and send a tripping signal. In other words, selectivity is lost and an N-X contingency arises, as conceptually illustrated in Fig. 1.

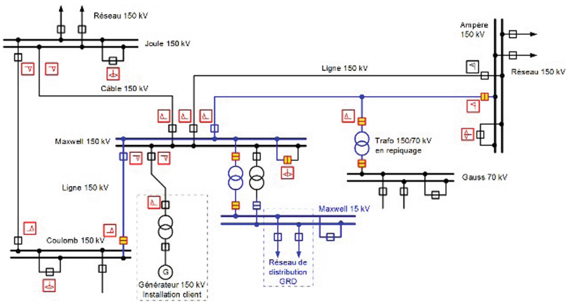


Fig. 1. Example on a conceptual network of an N-X contingency generated by the faulty elimination of a short-circuit on a 150 kV line: without backup protection, all elements connected to the busbar are switched off

In the particular case of a single busbar substation supplying an industrial customer this is particularly bad news, because it means the whole substation will be unavailable. These examples clearly highlight that once the statement of ageing combination of assets is made, the next step is to check how important the problem is. The originality of this paper is obtained through combining statistical method for assessing asset life information together with reliability block diagram analyses, so that both primary (high voltage) and secondary (low voltage) assets are taken into account. This permits a more effective assessment of the impact of a failure and its potential cascading effect. This will in turn enable a sound prioritization of investments and a greater consistency between the decisions taken for both primary and secondary assets. The main issue is selectivity and the most impacted business risk dimension is continuity of supply.

2 Methods

The following references [1, 2, 6] describe how, based on asset survival/reliability data, it is possible to deduce failure rates for end of life failures. In the case of Elia, removal information is available for some asset fleets like transformers (Fig. 2), though the reason lying behind removal is not systematically recorded. As a first assumption, it has been assumed that all removal is due to end of life failures. A more advanced correction method is described in [6].

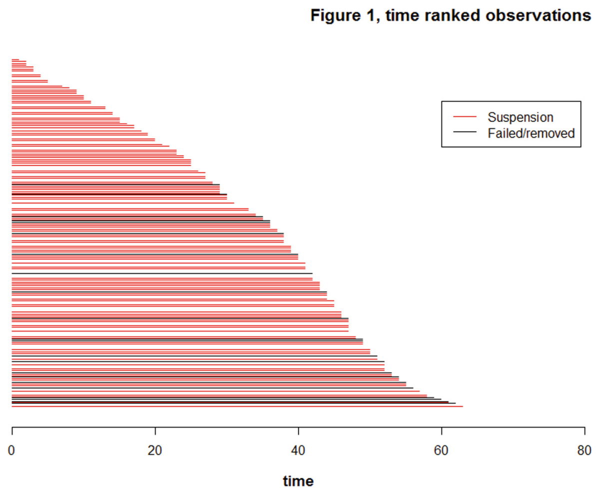


Fig. 2. Removal/right censored data for power transformers

Using this data and a Weibull distribution, the following reliability curve as a function of age can be deduced (Fig. 3). It correlates perfectly well with other transformer reliability curves published in [3].

For all asset fleets where there was no sufficient survival/removal available, a reliability curve was established based on expert judgement. Then, the best information available was used to position each asset on its fleet reliability curve: when- ever available, the asset condition was taken into account through its equivalent age. Reliability was then used as an indicator for probability of failure due to internal ageing, and classified in 5 categories, to be in line with the corporate risk policy. Having completed these steps, data is ready to check the interaction between assets located within the same bay. Two contributions are analyzed. On the one side, the

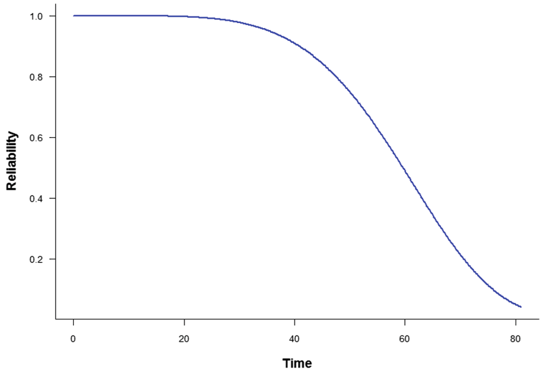


Fig. 3. Reliability curves for power transformers modelled with a Weibull distribution (mean time to removal = 58.9 years, Lower 0.95cl_lo = 57.4 years, Upper 0.95cl_hi = 61.2 years)

possibility to have a short-circuit (SC) caused by ageing of primary asset. A typical example would be an insulation fault. This contribution is assessed by the maximum probability indicator for HV equipment located in/connected to the bay (instrument transformers, transformers, cables, etc.).

$$P''_{\#} = \max(P_{*+,-}) \quad (1)$$

On the second side, assuming a short-circuit has happened, the probability of faulty elimination (P_{FE}) within the bay, might result in a loss of selectivity. Let S be the success probability of an event (S = 1 – P). Then P_{FE} is obtained with a reliability block diagram analysis (see Fig. 4) through combining the probability indicator of circuit breaker faulty elimination $P_{CB} = 1 - S_{CB}$ with the probability indicator of protection system faulty elimination $P_{9+} = 1 - S_{9+}$ as follows:

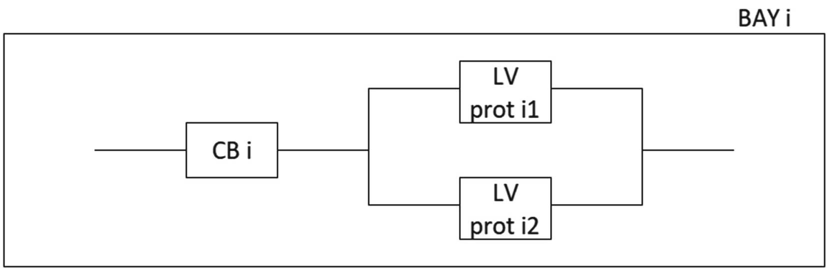


Fig. 4. Reliability block diagram analysis of the probability of faulty elimination of short-circuit by bay i

$$P_{<=} = 1 - S_{\#,} * S_{9+} \quad (2)$$

The redundancy of low voltage (LV) protection systems, reflected by the presence of protection 1 and 2, is taken into account as follows:

$$P_{LV} = P_{LV1} * P_{LV2} \quad (3)$$

Where P_{9+B} is assumed to be equal to one for the rare cases where information related to protection i is not available.

As a next step, an indicator to have a short-circuit in a bay caused by ageing of primary asset, followed by a faulty elimination can be computed as follows:

$$P_{BAY} = P_{SC} * P_{FE} \quad (4)$$

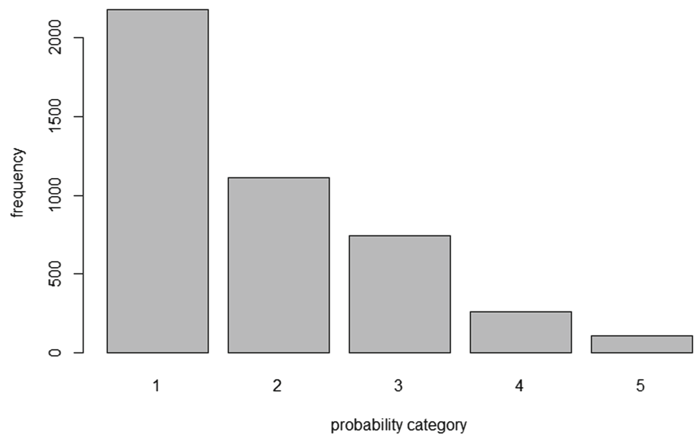


Fig. 5. Distribution of bays according to their probability indicator for short-circuit caused by ageing of primary asset followed by a faulty elimination

For the purpose of alignment with the corporate risk policy, results are then classified again in 5 categories (see Fig. 5).

These results, reflecting the chances to see the combination of unwanted events occur in a bay can be used as an input for further analyses, e.g.: evaluate the quality of asset interaction in a substation, assess the plausibility of scenarios and their im- pact on continuity of supply, challenge the consistency of asset fleet policies, check the alignment of previous decisions with corporate risk appetite, etc.

The reliability block diagram analysis can even be further extended to the analysis of a whole substation, as illustrated in Fig. 6 for substation j (SUBSj), or even to a project involving several substations.

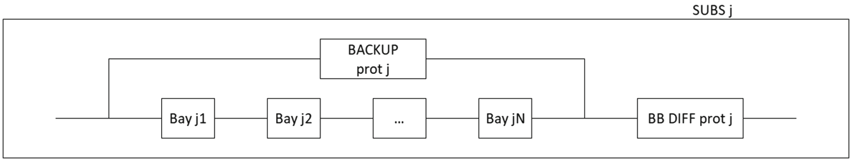


Fig. 6. Reliability block diagram analysis of the probability of faulty elimination in a substation SUBS j

3 Discussion

Though the model is a simplified one, it also offers several advantages. In the first place, its implementation takes into account the condition of individual assets. Then, once all input data has been loaded, running this model on the whole asset database only takes a few seconds on a standard laptop. To perform the analyses, R which is a free software programming language and software environment, widely used by data engineers has been used. A customized version of this methodology has been successfully implemented within Elia in 2018 in order to help prioritizing capital expenditure (CAPEX) projects.

The assumptions lying behind this simplified methodology can be clarified and discussed. In the first place, failures due to external events are outside the scope of this paper. Since the methodology assesses the chances to have a sequence of events starting with a short circuit of primary assets due to ageing, all asset failures not directly linked to asset ageing are outside the scope of this methodology. In the third place, the analysis made is a snapshot of the quality of interactions between primary and secondary assets. This means this model must first be combined to an ageing model of assets in order to predict future interactions between assets. To continue, the output does provide an indicator of probability of poor asset interaction, which is suitable for the purpose of comparing/ranking different asset systems and high- lighting needs, but which is not appropriate for quantified analyses like calculating the impact on performance indicators like system average interruption frequency index (SAIFI) or system average interruption duration index (SAIDI) [5].

Moving to a model providing correct absolute values will require a significant time, because a better view of the network topology and redundancies must be avail- able, the significant impact of automatisms must be taken into account, more failure/survival data is required to calibrate the ageing and maintenance models of all asset fleets, etc. The latter issue could be improved with benchmarking, through shortening the time needed collect enough data to develop asset ageing models and maintenance impact models.

4 Conclusion

This pragmatic 80/20 methodology[2] provides a probability indicator for a combination of unwanted events to occur in a bay, namely a short-circuit caused by ageing of primary assets followed by a faulty elimination at branch level. Such indicator for selectivity issues is a support for decision making process which does take into account the expert knowledge of people. Unlike other methodologies where the significant impact of low voltage equipment on energy not supplied is not considered, the present approach allows to take into account both the reliability and the redundancy of bay protection systems. It enables asset managers to develop an helicopter view of the interactions between assets, so that consistent decisions are made between the different asset fleets, and finally sound prioritization choices are made.

Acknowledgement. The authors wish to thank all their colleagues from the asset management department for their valuable suggestions and feedback.

References

1. Boudreau, J., Doche, R.: Impact of Mechanical Operations on Minimum-Oil Circuit Breaker Reliability. Cigre, Paris (2014)
2. Boudreau, J., Poirier, S.: End-of-life assessment of electric power equipment allowing for non constant hazard rate, application to circuit breakers. Electr. Power Energy Syst. **62**, 556–561 (2014)
3. Cigré Working Group D1.39, 2017. Cigré TB 706 Guidelines for the use of statistics and statistical tools on life data, s.l.: Cigré
4. IEC: Strategic asset management of power networks. IEC, s.l. (2015)
5. Marriott, R.J.: Reliability Analysis in R- A demonstration of some tools that can be used to conduct a reliability analysis in R (2016). uwa.engineering/project/example-of-reliability-analysis-in-r/
6. Picher, P., Boudreau, J.-F.: Use of Health Index and Reliability Data for Transformer Condition Assessment and Fleet Ranking. Cigre, Paris (2014)

Exploitation of Compressed Natural Gas Carrier Ships in the High North

Marina Solesvik(✉)

Nord University Business School, Nord University, Bodø, Norway
mzs@hvl.no

Abstract. There are huge reserves of hydrocarbons in the Arctic region. However, owing to the harsh climate, darkness, problems with communication, long distances, lack of infrastructure and other challenges, offshore operations related to hydrocarbon extraction are more difficult and more expensive. Nowadays, gas extracted from the Arctic shelf, is transported mainly by pipelines and liquid natural gas (LNG) carriers. Both methods are quite expensive. In this paper we consider the application of innovative compressed natural gas (CNG) technology to transport gas extracted from the hydrocarbon fields in the Arctic. The technology is not widely commercially used yet. The calculations show that the gas transportation costs using CNG carriers are lower than those using LNG carriers and pipelines for smaller gas fields and distances from the source to consumer up to 4000 km. The paper would be interesting for practitioners from the oil and gas and shipbuilding industry, policymakers and scholars.

Keywords: Compressed Natural Gas (CNG) · CNG carriers · Arctic · High North · Offshore operations

1 Introduction

Governments in many countries consider the development of renewable energy sources as one of the priority political tasks. Though the alternative energy sources are actively being developed and implemented worldwide, the role of hydrocarbons (oil and gas resources) is still very important, and will continue to be so for many years in the future. The role of natural gas as the cleanest and most hydrogen rich of all hydrocarbon resources is steadily increasing (Economides and Wood 2009). Burning gas generates less CO_2 emission per unit of generated electricity compared to traditional fuels like diesel, fuel oil or coal (Thomas and Dawe 2003). The results of geological studies suggest that about 12% of the world's oil and 25% of gas resources are situated in the Arctic. Five countries have a border with the Arctic (Canada, Denmark and Greenland, Norway, Russia, and the USA), and intend to use this opportunity and participate in oil and gas extraction in the Arctic. Taking into account the importance of the Arctic hydrocarbon resources, the governments of interested countries have established Arctic strategies; i.e. the Norwegian Government has developed long-term programmes and strategies for the Arctic, such as the Barents 2020 project; the Russian Government has worked out the Development Strategy of the Arctic Zone of the

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 46–54, 2020.
https://doi.org/10.1007/978-3-030-48021-9_6

Russian Federation and National Security for the period up to 2020 (RIAC 2013); and Canada has approved its Northern Strategy (Government of Canada 2014). The oil and gas sector is very important for the economies of these countries. In Norway, the oil and gas sector and related industries (suppliers, shipping companies, etc.) are the major driving force of prosperity and employment in the country. The same can be said about Russia. The depletion of traditional fields of hydrocarbons, such as in the North and Norwegian Seas in Norway, and West Siberia in Russia, are driving oil and gas companies and interested governments to look forward to exploiting rich hydrocarbon resources in the Arctic.

The extraction of hydrocarbons from the Arctic shelf is still more expensive than operations in non-Arctic areas, owing to the bad weather conditions (ice, cold weather, polar lows), remoteness from the shore and warehouses, lack of conventional infrastructure, and the lack of expertise. The cost of gas transportation (per unit of energy) to remote markets is about 10 times higher than the cost of oil transportation (Thomas and Dawe 2003). The purpose of our research is to consider the possibilities for making the extraction of the gas and its delivery to consumers cheaper, in order to make the final cost of gas extracted in the Arctic competitive in the world energy market. The technology of gas extraction from the shelf is complicated and consists of many areas where radical and incremental innovations can be applied in order to increase the efficiency of gas extraction and delivery and, thus, it will decrease the costs of Arctic gas. Several technical and technological challenges face the practitioners during the exploitation of deepwater gas fields in the Arctic. In this paper, we focus on the problem of extracted hydrocarbon transportation to shore terminals, and further, to the consumers. Currently, the problem of transportation is mainly solved with the help of underwater pipelines and liquefied natural gas carriers (Liquefied Natural Gas; LNG). For example, delivery of natural gas from the Snøhvit gas field in northern Norway is supported by LNG technology. However, the construction of LNG terminals and LNG vessels, as well as alternative or supplementing underwater pipelines, is both time-consuming and expensive (Staalesen 2014). Thus, scientists and practitioners are looking for cheaper and faster gas delivery modes from the continental shelf. The technology of compressed natural gas (CNG) sea transportation is a viable new alternative to the widely accepted LNG and pipeline systems for gas fields having resources less than 5 trillion cubic feet and the distances to markets up to 4000 km (Asibor et al. 2013).

This innovative method of natural gas delivery from the Arctic area is supported by compressed natural gas ship technology. In this paper, we aim to explore technical and economic components of the CNG technology application in the Arctic. Furthermore, we aim to analyse the prerequisites for the gas delivery with the help of CNG technology, and to develop the concept of the transportation system to bring gas from the Arctic shelf to shore terminals in the form of CNG.

The CNG technology has several important advantages: i.e. an opportunity to fill natural gas into the tanks on the board of a CNG ship straight from the shelf field, and discharge into the gas pipeline or gas network onshore. The use of the CNG ship technology considerably reduces capital investments and transportation costs, in comparison to developing the system of underwater gas pipelines, notably, in deep water, as well as LNG terminals and vessels. Therefore, further development of the

CNG technology adopted for the Arctic is an important issue for: policymakers, who are concerned with the diversification of gas supply channels; practitioners from the oil and gas industry, who are interested in further cost reduction and better efficiency of product delivery; and scholars concerned with the environmental, technical and economic issues of hydrocarbon extraction and transportation in the Arctic. Previous research focused mainly on LNG technology for the Arctic (Fadeyev 2014), or explored the use of CNG technology, without taking into account the Arctic challenges (Thomas and Dawe 2003). A novel conceptual contribution of this study is the exploration of innovative CNG technology application in the Arctic.

The paper is organised as follows. In Sect. 2, we highlight the context of the study and draw attention to the specifics of the Arctic region, and to challenges that face stakeholders for offshore operations there. In Sect. 3, conceptual insights from prior studies that have focused on natural gas transportation are discussed. Section 4 outlines our suggestions related to the use of CNG technology and ships in the Arctic. Finally, the implications of the study for further research, policymakers and practitioners are discussed, and concluding comments are presented in Sect. 5.

2 Arctic Context and Challenges

Arctic areas can be defined as areas to the north of the longitudinal tree line. This area has a mean July temperature of 10 °C. The Arctic is very rich in oil and gas and other mineral resources (Economist 2014). However, there are a number of obvious and less obvious reasons that challenge extraction works in the Arctic. In this chapter, we will consider closely the challenges related to hydrocarbon extraction in the High North. A number of political, environmental, ecological, technological and economic issues will be taken into account regarding oil and gas firms' intent to explore for hydrocarbons in the Arctic. First, the Arctic is still a disputed territory from the international law point of view. Several adjacent states claim that they have rights to the North Pole and the surrounding areas (Sale and Potapov 2010). The border issues play an important role and will have an influence in the future. The border agreements between countries can also hinder hydrocarbon and other mineral resource extraction in the Arctic. For example, there is still a disputed offshore international boundary between the Yukon and Alaska. The maritime border agreement between Norway and Russia was only signed recently (in 2010), which opened up oil and gas exploration in the Barents Sea.

Second, the weather conditions in the Arctic are considerably more difficult than in the Northern Atlantic where major oil and gas companies have accumulated significant expertise. Severe weather conditions make offshore operations and subsea works exceptionally difficult (Borch and Batalden 2015; Masterson 2013; Staalesen 2014). Previous research has suggested that low temperatures (up to −40–−50 °C), huge temperature differences, Arctic ice cover, drifting ice and floating icebergs, permafrost, waves, sea spray, wind, polar lows (unexpected changes of weather when a severe storm can start very quickly in 0.5–2 h), and darkness represent major climate challenges in the Arctic. This leads to additional challenges. For example, the existing technology allows offshore operations to only be carried out in some areas during the periods when the water surface is free from ice. The offshore vessels then are used only several months per year. It means that they are hired at spot rates that are significantly higher than rates under long-term contracts for periods of up to 15 years.

Third, longer distances between the shore and the hydrocarbon fields are a drawback. The distances to the hydrocarbon reservoirs in the North and Norwegian Seas are relatively short. In some cases, it takes only three to four hours to reach the platform from the shore. In the Arctic, the distances between the mainland and the discovered oil and gas fields may be 500–600 km, and this is not the limit. This is a logistics challenge for the oil and gas companies and their subcontractors who are supposed to deliver and evacuate people, and bring cargo to platforms. The emergency response time is also increased, when taking into account the longer distance to the shore. In addition, the infrastructure is underdeveloped in the areas adjacent to the Arctic seas. The above are major concerns of experts and practitioners since the development of the appropriate infrastructure (e.g. railways, roads, bridges, telecommunication facilities, and so on) leads to significant capital costs and increases the return-on-investment period (Harrison 2006). Longer distances to the main supply bases increase the cost of supplies. Some offshore shipping companies that already have operations in the Norwegian part of Barents Sea (Goliat and Snøhvit fields) established local warehouses in Hammerfest, while others deliver necessary parts and materials almost daily from their home bases (e.g. Esvagt shipping company delivers what is necessary from their base in Denmark by plane). Longer distances to the main consumers are also reflected in the higher costs for oil and gas extracted in the Arctic.

Fourth, the GPS, mobile connection, and telecommunications do not always work properly in the Arctic. This leads to difficulties communicating with the personnel on the ships and platforms. The dynamic positioning is complicated when the GPS does not work properly.

Fifth, the environment in the Arctic is very fragile. That is why the Technology Strategy of the Arctic of Norway stipulates the strictest safety requirements for offshore operations there. The governments of other Arctic states are also concerned with the environmental issues. Oil and gas companies pay a lot of attention to the preservation of the unique nature of the Arctic. Possible oil spills there can have unpredictable consequences for nature. Subsequently, oil spills would lead to a negative media image and large costs related to oil recovery. Oil and gas companies will pay higher prices to subcontractors who need to offer more expensive technical solutions that would work under low temperatures, ice and harsh conditions. For example, ships will additionally need winterization of the vessels, deicing systems, metallization of the surfaces, icebreaking capacities, thicker hulls, and so on. All these additional options are reflected in the price of the vessels. The fuel consumption will increase as well.

Arctic offshore operations will need new technology, routines, and equipment (Research Council of Norway 2006). The Arctic can be divided roughly into three areas: (1) an ice-free zone; (2) a zone covered with the ice several months of the year; and (3) an area covered with the ice the whole year. For the areas without ice (e.g., the Barents Sea), the existing technology of oil extraction and offshore support is available with minor changes. For the areas covered with ice for some months of the year (e.g., the Kara Sea), incremental innovations are necessary to be developed, and for the areas with surfaces covered with ice during the whole year, radical innovations need to be developed and implemented. For example, from ice-covered Arctic hydrocarbon fields, oil and gas can be delivered by submarine tankers and LNG ships respectively (Jacobsen and Murphy 1983).

3 Theoretical Background

Natural gas is extracted both from gas reservoirs and from oil reservoirs as a by-product. Some gas is always extracted when crude oil is produced (Cranmore and Stanton 2000). Transportation of natural gas is expensive. It is suggested that the transportation of natural gas costs about ten times more than the transportation of crude oil (Thomas and Dawe 2003). A CNG ship is equipped with a cargo-piping system with tanks for pressurized gas transportation. The idea of using CNG technology for the transportation of the natural gas by the sea is not new. The first attempts to use CNG technology in the sea where made in the US in the 1960s. However, at that time, technologies related to the heavy weight of high-pressure gas cylinders (about 2 t/m3 of available storage) and their material characteristics were not sophisticated enough. In combination with the low prices of hydrocarbons fifty years ago, the use of CNG ships was not economically advantageous. Recent development in composite technology and increased demand for hydrocarbons, combined with higher prices, have made the application of CNG ships a viable, and often cheaper alternative to the competing LNG vessels and pipelines. Recently, we have witnessed the advancement of major components of CNG technology in different parts of the world (Gaiotti et al. 2016). For example, in Indonesia it was decided to use CNG ships to deliver natural gas to the islands. The construction of pipelines for natural gas transportation would be more expensive there. Since 2000, leading firms produced conceptual designs for CNG ships that have an advanced structural arrangement of the cargo system. More specifically, the experts at the Trans Canada Pipeline Corporation presented a variant of a gas carrier with horizontal cylinders, manufactured as a gas pipe reinforced with fibreglass composite material (The Street 2006). Their competitors, Ener Sea Transport and Trans Ocean, presented another solution, where the cargo cylinders are placed in special airproof cooled modules. Furthermore, Coselle Corporation designed the system of the cargo tank in a form of a long pipe wound in a spiral called Sea NG (Coselle 2014). However, such a solution diminishes the safety coefficient for the cylinders from 2.5 to 1.7. In addition, CNG technology has been enthusiastically embraced in Russia (Blinkov et al. 2006). However, these projects have not been implemented yet.

4 Analysis of CNG Technology Application in Arctic

The Arctic area is vast and occupies about 6% of the earth’s surface. The Arctic area consists of one-third of the land and two-thirds of the water or ice-covered surface. As has been pointed out earlier, from the point of view of hydrocarbon exploration, the Arctic area can be divided into three zones: (1) an ice-free zone; (2) a zone covered with the ice several months of the year; and (3) an area covered with the ice the whole year. Apparently, the CNG vessels that are required for each of three zones would be different.

Ships aimed to be exploited in the Arctic have different ranges of ice class, which is reflected in the shape of the hull, degree of winterization, thickness of steel plates, ice-breaking capacity, etc. All these construction features are reflected both in the price of construction and fuel consumption (Borch and Solesvik 2013). For example, it is much more expensive to use a CNG carrier with an ice-breaking capacity in ice-free areas compared to a conventional CNG ship since the latter consumes much more fuel caused by a different hull shape. One of the deterrents for the rapid growth of gas production in the Arctic is a partially vague and expensive system of natural gas transportation to the shore. Therefore, it is timely to develop a balanced system of sea transportation for the exploitation of the shelf fields in the Arctic. For example, the Shtokman project in the Barents Sea has been postponed owing to the high costs of gas transportation. The initial plans were to build an LNG terminal in Teriberka, which is located 100 km from Murmansk. However, the calculations showed that the final cost of gas would be too high, when taking into account building expenses. Using CNG could be a viable alternative for the LNG in small gas reservoirs in the Barents Sea. The CNG ships are appropriate for use for middle-range distances in the smaller gas fields, as well as fields where traditional methods of gas transportation are not profitable (Blinkov et al. 2006). Using sea transport on a CNG-technology basis is reasonable, since it allows money to be saved on the construction of industrial pipelines and the LNG terminal. In particular, a means of gas transportation by sea is stipulated by the special traits of the systems that provide the implementation of the CNG technology. The arrangement and parameters of the system essentials would depend on the parameters of the gas supply source and the gas transportation procedure where it will be offloaded further. A CNG ship that would receive gas from shelf fields should be equipped with a system for complex gas-handling, in a composition similar to the gas-handling systems for uploading into the gas pipeline on the shore fields. The gas compression system will be installed on a ship to make it possible to pipe gas into the on-board tanks.

Each natural gas offshore field needs an individual transportation plan. Capital expenditures on a CNG can be reduced if one selects an optimal number of ships, capacities, and routes. In order to use CNG ships as profitably as possible, it is necessary to carry out an optimization investigation, in order to outfit ships with the equipment that would maximize the ship's performance. At the same time, it is necessary to support the unification of technological and transport ships. This will make it possible to use them indifferent gas fields. For the Arctic environment, special aspects must be taken into account, i.e. materials and coatings; hull construction, arrangement and equipment; ship systems and machinery; safety equipment and procedures; crew considerations and training (Legland et al. 2006).

Taking into account that the construction of the specialized CNG vessel is still expensive (although cheaper than building an LNG vessel), scholars have extended the technology of compressed natural gas shipping by a standard container vessel. They offered a special cargo system in the shape of separate freight sections so that each of them corresponds to a standard 20-ft or 40-ft sea container. This allows both to save time in the construction of the CNG ship (about 2 years), as well as save money, since a standard container ship is significantly cheaper than a specialized CNG ship. Moreover, if the vessel is ordered at the bottom of a shipbuilding cycle, the price of the vessel can be significantly lower than the price of construction during the cycle's peak (Solesvik 2011).

5 Conclusions and Implications

The barriers to the exploration, transportation, and development of hydrocarbon resources in the Arctic are significant (Harrison 2006). The governments in countries adjacent to the Arctic are looking for ways to reduce these barriers. In this paper, we considered the option of the application of novel CNG technology for the transportation of gas from the Arctic offshore fields to consumers, in order to contribute to the solving of the transportation barrier problem. Logistics and the cold climate are major challenges in the offshore exploitation of hydrocarbons in the Arctic. We compared the pros and cons of CNG vessel utilization with widely used LNG ships and underwater pipelines. So far, CNG ships have only been used in conventional seas, where the climate and the environment are not specific challenges for oil and gas and shipping companies. However, CNG technology offers significant advantages for offshore operations in the Arctic as well. First and foremost, it is cheaper to use CNG ships to deliver gas from smaller fields from Arctic islands and offshore than LNG ships, in combination with shore LNG terminals that liquefy gas from the oil fields to be carried by the LNG ships. Second, it is time-saving technology; i.e. it takes several years to build pipelines and LNG terminals. As a result, it costs 1.5–2 times less to deliver natural gas by CNG ships than LNG ships to middle-range distances of 1,500–1,700 miles. Finally, the operation of CNG ships is a flexible means of gas transportation. When the gas field is depleted, the vessel can be easily moved to a new field; something that is impossible for pipeline systems or onshore LNG terminals. We have elaborated on the idea of CNG vessel utilization for the Arctic. The utilization of CNG ships in the Arctic needs design modification and/or elaboration of the special design of these vessels for some areas in the Arctic. The design would be different, depending on the area of exploitation, since the Arctic is not homogeneous area. Weather and ice conditions differ in different areas. For the relatively calm Barents Sea, CNG vessels do not need special ice-class modification, and the only necessity is the instalment of a winterization package (i.e., de-icing equipment and metallization of some surfaces). Moreover, CNG ships can deliver associated gas from the oil fields in the Arctic and the North Sea. CNG technology can be utilized in Far East gas fields as well. There are about thirty small and medium-sized gas fields in the Sakhalin area. Transportation of the extracted gas using pipelines and LNG technology is not beneficial for smaller fields. Thus, CNG ships can be used there.

In this paper, we considered existing CNG technology, and compared its pros and cons compared to LNG technology, pipelines or their combination. We proposed ways to apply CNG technology in the novel research context of the Arctic. The study can be interesting for scholars in various disciplines concerned with effective technology and innovation development in the Arctic. The proposed results can be a building block for future research.

CNG technology is rather innovative and has several advantages. First, wider use of CNG technology will allow better use of natural gas by consumers in Norway, Russia and other countries. It is well known that although Norway is one of the biggest gas producers in the world, domestic consumption is quite low, both by households and industrial customers. At the same time, the Norwegian Government is concerned with the promotion of ecological sources of energy. The use of smaller CNG ships that can deliver gas directly to local pipelines that lead to end customers is one of the applications of CNG technology in Norway. Moreover, the ideas related to the wider and more effective use of CNG vessels in the Arctic can be useful for policymakers and practitioners in all countries of the circumpolar area, who are concerned with the effective and clean hydrocarbon extraction of the Arctic reservoirs (Fadeyev 2014; Harrison 2006; Solesvik 2018).

The depletion of oil and gas resources in the traditional fields forces oil companies, operating in countries with access to the Arctic shelf, to consider moving their operations further north above the Arctic Circle. Practitioners in Western countries are concerned with the lowering of production costs and preserving the high quality of goods and services at the same time. The concept of lean manufacturing is more and more popular among small and large enterprises. The application of CNG technology is, in many ways, cost-beneficial for producers and carriers of hydrocarbons, who are pressed with the fluctuations in price for oil and gas. The lower price for both oil and gas observed in 2014 has forced oil and gas companies to look for cheaper alternatives, and save on capital investments and operating costs. The novel CNG technology can be the remedy, which will allow savings on capital costs to build expensive LNG terminals and develop infrastructure around the LNG terminals. The cost of CNG ships is also half price as that of comparable LNG ships. This makes CNG technology a viable alternative or supplement to the LNG technology used in the Arctic so far (e.g., Snøhvit oil field in the North of Norway). The proposed CNG technology for the Arctic can be used in future projects and oil fields in the Arctic, Sakhalin and neighbouring smaller fields. Our study is limited by considering general aspects of CNG technology use in the Arctic environment. We have developed our ideas based on the relatively ice-free conditions of the Barents Sea. Future studies can consider the application of CNG technology in the East Arctic which is more ice covered. CNG ships for the East Arctic might need an ice-breaking capacity. Future studies and research in other disciplines might consider technical issues related to CNG technology in the Arctic. They might be interested in the further development of CNG technology for the Arctic, and might propose innovative technological solutions to meet the challenges of the harsh weather conditions observed in the Arctic. Novel ship designs for Arctic navigation would be a subject of cooperation between representatives of different industries; i.e. ship designers, shipbuilders, ship owners, and oil and gas companies (Solesvik and Westhead 2010).

References

Asibor, E., Marongiu-Porcu, M., Economides, M.J.: Oil production optimization without natural gas constraints: the harvesting of upstream natural gas. J. Nat. Gas Sci. Eng. **15**(1), 59–68 (2013)

Blinkov, A.N., Vlasov, A.A., Litcis, A.V.: Seaborne transportation of compressed natural gas from shelf. Maritime Exchange **3**(17), 82–87 (2006)

Borch, O.J., Batalden, B.M.: Business-process management in high-turbulence environments: the case of the offshore service vessel industry. Maritime Pol. Manag. **42**(5), 481–498 (2015)

Borch, O.J., Solesvik, M.: Collaborative design of advanced vessel technology for offshore operations in Arctic waters. In: Luo, Y. (ed.) Cooperative Design, Visualization, and Engineering, pp. 157–160. Springer, Berlin (2013)
Coselle: Coselle System Overview (2014). http://www.coselle.com/coselle-system/overview
Cranmore, R.G., Stanton, E.: Natural gas. In: Dawe, R.A., (ed.) Modern Petroleum Technology, Upstream Volume, pp. 337–382. Institute of Petroleum, John Wiley, Chichester (2000)
Economides, M.J., Wood, D.A.: The state of natural gas. J. Nat. Gas Sci. Eng. **1**(1), 1–13 (2009)
Fadeyev, A.: Arctic Offshore Exploration: An International Perspective. Moscow, Russian International Affairs Council (2014). http://russiancouncil.ru/en/inner/?id_4=3396#top
Gaiotti, M., Rizzo, C.M., Rizzuto, E., Vernengo, G.: Material selection for the gas containment system of a compressed natural gas carrier fleet. Appl. Ocean Res. **55**(1), 37–47 (2016)
Government of Canada: Canada's Northern Strategy (2014). http://www.northernstrategy.gc.ca/index-eng.asp
Harrison, C.: Industry perspectives on barriers, hurdles, and irritants preventing development of frontier energy in Canada's Arctic Islands. Arctic **7**(2), 238–242 (2006)
Jacobsen, L.R., Murphy, J.J.: Submarine transportation of hydro-carbons from the Arctic. Cold Regions Sci. Technol. **7**(1), 273–283 (1983)
Legland, E., Conachey, R., Wang, G., Baker, C.: Winterization guidelines for LNG/CNG carriers in Arctic environments, ABS Technical Papers, pp. 305–316 (2006)
Masterson, D.M.: The Arctic Islands Adventure and Panarctic Oils Ltd. Cold Reg. Sci. Technol. **85**(1), 1–14 (2013)
RIAC: International Cooperation in the Arctic, Report, Russian International Affairs Council, Moscow (2013)
Sale, R., Potapov, E.: The Scramble for the Arctic. Ownership, Exploitation and Conflict in the Far North. Frances Lincoln, London (2010)
Solesvik, M.Z.: Interfirm collaboration in the shipbuilding industry: the shipbuilding cycle perspective. Int. J. Bus. Syst. Res. **5**(4), 388–405 (2011)
Solesvik, M.Z.: Partner selection in green innovation projects. In: Complex Systems: Solutions and Challenges in Economics, Management and Engineering, pp. 471–480. Springer, Cham (2018)
Solesvik, M.Z., Westhead, P.: Partner selection for strategic alliances: case study insights from the maritime industry. Ind. Manag. Data Syst. **110**(6), 841–860 (2010)
Staalesen, A.: Dramatic return from Dolginskoe, Barents Observer, 19 November 2014. http://barentsobserver.com/en/energy/2014/11/dramatic-return-dolginskoye-19-11
Thomas, S., Dawe, R.A.: Review of ways to transport natural gas energy from countries which do not need the gas for domestic use. Energy **28**(14), 1461–1477 (2003)

Smart Life Cycle Management of Social Housing Assets

Quynh T. H. To[1,2](✉), Isam Shahrour[3], and Trang T. T. To[4,5]

[1] Lille University, Lille, France
to.t.h.quynh@gmail.com
[2] National University of Civil Engineering, Hanoi, Vietnam
[3] Civil and Geo-Environmental Engineering Laboratory, Lille University, Lille, France
isam.shahrour@univ-lille.fr
[4] National Economics University, Hanoi, Vietnam
trangulsa@yahoo.com, trang0312@gmail.com
[5] University of Labour and Social Affairs, Hanoi, Vietnam

Abstract. The social housing sector is submitted to increasing challenges, related to the aging of this sector, from poor quality, degraded life environment, high running expenses, tenants' low income; as well as reduction of the public funding for this sector and increasing sustainability requirement to reduce both energy consumption and greenhouse gas emission.

To meet these challenges this sector has to innovate both management and operation methods in order to optimization effectiveness in this sector together with the improvement of its performances. This strategy should be based on smart management. This paper proposes a smart life cycle management of social housing assets base on the result of the technology innovation in this sector. It contains both the original and the real-time information of assets, in order to build a smart-knowledge strategy for whole life cycle of asset management to improve the social housing efficiency and quality.

The paper presents a discussion of this smart life-cycle management system. It proposes the methodology for each process of this system's strategy. Its goal is tracking the optimal management of the social housing asset.

Keywords: Smart management system · Sensor and multi-sensor · 3D technologies · Social housing assets · Innovation · Tenant behaviour · Life cycle cost

1 Introduction

French social housing or public housing (moderate rent habitations, *habitations à loyers modérérs, HLM* in French) was initiated privately and early from the year of 19th century, in the same period with the English public housing. But till to the last decade, the first social housing agencies in France were formed (Le Blanc and Laferrère 2001). Nevertheless, the start of French social housing stock was significant after the Second World War (1948) as known as modernization and reconstruction period in France. The creation of HLM 'low-cost housing' response the demand for helping moderate-income

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 55–66, 2020.
https://doi.org/10.1007/978-3-030-48021-9_7

families domicile (Habitat 2003). During nearly 50 years, French government had adopted many policies to improve both quantity and quality of social housing stock. For example: the FNAH (National Fund for the Improvement of Housing) concentrated on building maintenance's problems resolving; the Orientation Act for Cities (LOV - La loi d'orientation pour la ville) rounded up the conditions of living and housing; specially the Urban Solidarity and Renewal Act (La loi SRU (Solidarité et Renouvelleme Urbain) which was valid from the 13th December 2000 in order that the large municipalities reserve at least 20% of French housing stock (Habitat 2003). As the result of the Government's efforts, in parallel with the increasing of housing stock volume, French social housing stock had grown quickly in 40-year-period (Schaefer 2008). With the number of social rented dwellings reached 5,4 million units in 2015, the social housing stock accounts for 18,4% of the 29,3 million units total housing stock (OECD Affordable Housing Database 2016; Rochard et al. 2015).

However, since the major part of this stock is old, it suffers from poor quality as well as high running expenses and degraded life environment. With the increase in concern for optimizing the operation and maintenance cost, improving the quality of life for tenants and reducing both of the energy consumption and greenhouse gas emission, public authority and social housing managers are interested by the development of innovation such as smart technologies to meet the increasing challenges in the social housing sector which HLM organization in North of France is highlight example (Aljer et al. 2017a).

2 Building Asset Management in Social Housing Sector

2.1 Essentialness of Asset Management for Superintending Social Housing Asset

Global social housing management is having a transformation to react with the innovation in institution and economy in many countries. Public housing stock also has some revision as becoming private - social part or public services are decentralized. These conversions have brought to social housing stock over the world more market-oriented management. The most mentioned trends are: British public housing associations are more and more detached from government to be more "real market parities"; or fading away the gap between social and commercial landlords after adoption of the Public-Use Housing Law in 1984 in Germany; or Australian State housing strategies as long as be appropriate for their State are all allowed even when they did not follow national consistency previously arranged (Gruis and Nieboer 2004).

Housing systems reformation go with reducing of government finance and subsidisation as in Sweden, England, The Netherlands and Australia makes social landlords more rely on the private capital market. Hence as accompanied result, their operation turns into more business-like way to have satisfactory of their financial suppliers. Cannot be outside the inevitable trend, since 2010, "l'Union sociale pour l'Habitat – the umbrella organisation bringing together all HLM federations - has launched a process of "thinking on a new business model taken into account the needs for affordable housing in France and the pressure on public funds for the State and the local authorities." (Gruis and Nieboer 2004).

One of the conspicuous changes in social housing's activities is management. Not only stopping at categories like technical management (maintenance, renovation, etc.) or social management (housing allocation, etc.) like in the past, at the present time social housing managers have been extending their concern to financial management (treasury, rent policy) and even tenure management (letting, buying, selling) as private housing landlords do (Gruis and Nieboer 2004). According to researchers' point of view, housing management used to distinguish various types: property management, facility management and asset management. In which, Asset management is term often associated with financial investments (CHOA's 2013). The efficiency gained from judicious asset management could help social housing landlords solve totally their limitations (Gruis and Nieboer 2004).

Be interested from the late of 20th century when financial performance played more important role in management decisions of social landlords, till recent years, asset management in social housing sector is not a new concept anymore. By analysing asset's performance, it has a major influence on the decision-making about holding, selling and repositioning, simultaneously satisfying requirements of optimising financial effectiveness.

2.2 French Social Housing's Asset Management

Stemming from the private sector where managers care only about optimizing the financial performance of an organization's assets, going into the social rented sector, the financial performance of asset has a lower position than social housing objectives efficiently. In spite of this, social housing asset management still keeps all the characteristics of methodologies, tools or resources from the private's one (Gruis and Nieboer 2004). Indeed, asset management systems for Quebec school building (Canada), Mosman council (Australia), or European social asset management practices are shreds of evidence.

French social housing stock has much in common as well as differences in its context and characteristics to compare with other countries. In asset management field, building asset types are like general, but management strategies and stake-holders are not the same. Indeed, buildings in general or building assets in particular are individual. Even they built base on other's design, their location, holder and user made them be dissimilar. Therefore, asset managers could not use the one strategy for all buildings they have – "there is no one-size-fits-all model". In addition, with several types of subsidies, finance and providers join to assess managing social housing asset, the requirement of planning strategic guidance for social landlords is the most necessary. In France, strategic asset management in the social housing sector used to be not as widespread as in the private sector. Although the social housing companies are 'private' institutions, they are subject to regulation by the French government concerning for example rents, allocations and investments. Social housing management companies were non-profit organization, which have not been market-oriented for a long period. However, in early the 21st century, they were initiatives to develop a more sophisticated approach towards strategic asset management (Gruis and Nieboer 2004). Little by little, managing assets of social rented buildings is having similar characters like private one does.

Since 2002, the French government has strongly encouraged social housing companies to build strategic managements approach to optimize their scarce resources and to put an end to the misallocation of grants. Social housing companies without a strategic approach could not touch to state subsidies to build or refurbish their housing stock. Added to, the competition from the private sector leads a result that several companies have to improve their approach to horizontal integration and change the performances registered by past refurbishment policies which have been very poor (Gruis and Nieboer 2004). The requirements for asset management in the social rented sector in France also changed a lot. Before, "The building and housing code state the minimum maintenance requirements. Maintenance is required to ensure safety standards for elevators, heating systems, gas appliances, collective gas-controlled mechanical ventilation systems, garage doors, detection systems, fire protection, etc." (Gruis and Nieboer 2004). After tenants have to pay for on-going maintenance depending on characteristics of their dwelling through service charges. They ask for a better quality of housing. Some tenants are attracted by parts of the private sector that may provide dwellings with better services and similar rents (Bat'im Club 2001; Gruis and Nieboer 2004). Day by day, this requirement is more and more difficult to satisfy. According to estimation, about 25% of French social housing stock had built from 30–40 years ago, more than 20% of the stock is over 50 years – old and only 35% stock are recent or renovation buildings (Schaefer 2008). In the building's - life – cycle point of view, they are all in "Adulthood" and "Old Age" stage. With poor management, it was not by accident that "Many of these dwellings suffer have a mediocre technical quality and a poor market position" (Gruis and Nieboer 2004).

So the social landlords must think most about asset renewals and upgrade their building assets or even rebuild while planning asset management strategies. However, social housing sector has complex stakeholders with the different viewpoint about lasting asset's longevity. Cause the age of assets to be used to calculate depreciation is different with technical-age and far from use-age. Managing asset systems of complex – building, under governing of special policies, complicated stake-holders joined into decision making process and keeping the balance between economic and social goals obviously is not straightforward. Besides, building asset management is long-term and complicated. It requires effort contributions and consultants from all involved parties from forming to using of property in order to succeed. In social rented sector, the managers that related to construction period are contained of: Architectures or engineers of design consultant and building contractors corresponding to each asset types. They have responsibilities and duties, in order are author supervision and warranty in the using period. After all, social land-lords and maintenance - renewal contractors play the most important roles in managing assets. Among them, accountants and property managers in social housing manage organizations work toughest. The reason is they must balance between financial performances and social effects in limited budget. Without technology support, asset managers in social rented sector may have a huge challenge.

2.3 Optimizing Asset Management Performance by Applying Smart Technologies in the French Social Housing Sector: Experiments and Challenges

Nowadays, applying new technologies for sustainable development has become the Global trend in all the fields of life. Smart technology like BIM (Building Information Modelling) for designing and supporting digitalized information or PMS (Project Management System) for governing construction process had used regularly in constructing from a long time ago (Jupp 2014). As a leading result, building manager used BIM design for supervising the property constructed with BIM. There are many result of researches related to BIM adoption in construction and manage building prove that BIM not only reducing workload for managers, but also optimizing efficiency in using (Jupp 2014; Love et al. 2014; Volk et al. 2014; Barry 2013). Among them, Peter E.D. Love's research "a benefits realization management building information modeling framework for asset owners" (Love et al. 2014) and Rebekka Volk's paper "Building Information Modelling – BIM - for existing building – Literature review and future needs" (Volk et al. 2014) all published in 2014 had provided social landlords theoretical basis as well as best practices to change their management behaviours.

Besides, other studies focused on modernizing the asset management decision process. Four researchers could be mentioned are: Dino Gerbsic has studied asset management system AMS in Canada (Gerbasi 2005), Ashish Shah analysed Australian practices on Building Asset Management (BAM) system types (Shah et al. 2004) or Paul Dewich and Marcela Miozzo's research mentioned both the latest technologies to collect data and the barriers approach in social housing sector (Dewick and Miozzo 2004) and Jan Kitshoff, Robin Gleaves and Gordon Ronald tried to figure out how to increase productivity and performance of Social Housing Asset Management by innovating approach (Kitshoff et al. 2012).

Viewed from different sides, favourably always goes with challenges in innovation adoption. However, that does not affect the owners on the way following the trend of modernization management. Up to now, there are many countries successes in modernization in building management. Typical patterns could mention are: Sustainable technologies in Scottish social housing (Dewick and Miozzo 2004); Building Information Modelling - BIM & Facility Management - FM in Hong Kong's public rental housing (BIM forum 2013); Energy monitoring programme for social housing in Ireland (Sinnott and Dyer 2012); Return on investment of BIM to management cost to HLM organisations in France (Caisse de deposts 2015). Synthesis lessons learned from 6 models show that even though social landlords in each project focused on different works in the asset management process, they all used smart technology to improve management performance. BIM system built the digital framework contained initial information of building asset (technical indicators; installation location, linkage with other devices…). While the sensor monitoring system updated real-time data and asset management system works on data mining (big data). It is not difficult to realize that smart technology can hold up during the building life-cycle whole asset management process. Accordingly, by means of making the best use of smart technologies that have been studied and applied in Smart city project (Lille University) for long period, optimizing social housing asset management performance is prospective and no longer be a big challenge.

3 Smart Building Asset Management System Design for Social Housing Asset in HLM Organization – One Case Study

The benefits of BAM are proven in many researches, improving the performance and the contribution to safety, health along the building asset lifecycle and protection of the environment by reducing buildings' expenditure and trash. Together with the organizational commitment to quality, performance or safety, it helps to mitigate the legal, social and environmental risks associated to building assets. Building Asset Management, as a discipline, allows organizations to optimize the whole life value of managing building assets portfolios. For a single building management organization, the list of assets or portfolio may contain diverse assets in types, distributors, and subjected to differing demand/utilization requirements.

In the 21st Century, Smart Technologies are introduced to various stakeholders in building asset management, namely owners and users, to improve performance, quality, and to increase satisfaction. Smart Technology has changed stakeholders understanding and perspective towards performing business, collaboration cooperation, communication and connection. Stakeholders are using smart technologies for accessing, exploiting and managing assets. Adopting these smart technologies in such sectors poses a huge challenge in training and implementation. Nonetheless, using such technologies will improve several areas performance and productivity; reduce costs, improve sustainability, increase users' satisfaction, retention, and loyalty in the long run. However, new threats will be introduced for example, security, privacy, legal conflicts and risky reputation, if it is used inappropriately. From the state of using smart technologies for building asset management, it could not be denied that up to now there is no smart management system that combines all the technologies been able to full support for BAM. This following part of research aims to build a smart building asset management system, which formed by smart technologies and strategies after investigate and assess some of the effects and influences of smart technologies.

3.1 Applying Smart Technologies into Building Asset Management

Parlikad et al. (2017) presented their report about state of art of using intelligent technologies into infrastructure assets management, in which, there are a lot of common aims to building asset management. From this research and others related to smart technologies needed, it is not difficult to find the fittest smart asset management system for building assets. Indeed, a smart building asset management system also needs to be smart into seven steps (Fig. 1).

The research of Parlikad et al. (2017) mentioned that there is a wide variety of product and smart technologies are now available on the market, offering different

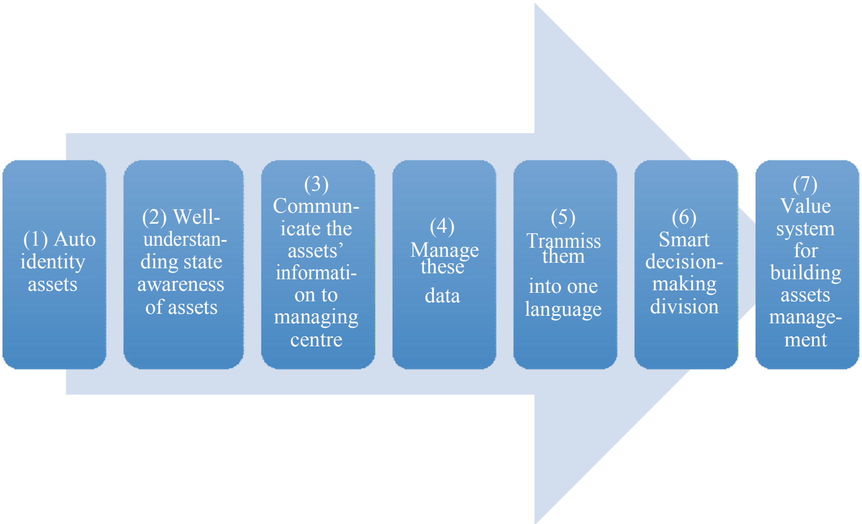


Fig. 1. Smart asset management system (Parlikad et al. 2017)

capabilities in terms of users' requirements. For asset identification ((1) in Fig. 1), two-dimensional barcodes and Radio Frequency Identification (RFID) tags are the most popular due to their advantages in cost and easy to apply. To help understand the state or condition of assets ((2) in Fig. 1), a variety of sensing technologies have been developed. These sensors track the critical parameter that can provide an indication of the rate of progression or the likelihood of development of different failure modes of an element, an asset or the system. Then, for communication database purpose ((3) in Fig. 1), the data captured by these sensors needed to be communicated to data management systems and analysis systems for making decisions. This process can be performed either by connecting the sensors physically using wires for data transmission or through wireless sensor networks. After that, for managing data ((4) in Fig. 1), the big challenge for managers is how to identify the right data to be collected for effective whole life management of single or multiple assets. The standards such as the PAS-1192 series and the development of BIM compliant solutions offer industry the capability to manage asset database more efficiently. The intention of the intelligent asset model is that a mechanism be available to enable all asset data to be accessed via a single query, without requiring searches across multiple databases and storage locations in different organisations, which can draw on the developments in the field of the Internet of Things (IoT). Nevertheless, how to transmission of asset information ((5) in Fig. 1) in a standardised format is a challenge in a world where different organisations (or different departments within the same organisation in reality) use the information systems and databases in different types. One of the major developments towards this is the move towards standard formats for exchanging asset data that the Construction Operations Building Information Exchange (COBie) is the most popular in the civil engineering sector. COBie is a formal schema that can help establish the organise formation of new and existing assets. COBie supports capturing and recording the important data at the point of origin, including equipment lists, asset data sheets, warranties, spare parts lists and maintenance schedules.

After having the "right" asset databases, how the making decision support ((6) in Fig. 1) from this information is the main question for asset managers. Parlikad et al. (2017) reckons because new sensors may often produce the new engineering datasets (that were not available in the previously systems), so novel techniques for the interpretation of this new engineering data are required. In addition, advances in the area of predictive, analytics help asset managers not only to understand the state of their assets, but also to predict impending failure and create the opportunity to tale optimised preventative action. New developments in the area of software agents enable the concept of smart asset management to become a reality; for example, where every asset is represented by a software program that continuously seeks data from various sources (including sensors), analyses it and takes decisions without the need for human intervention. The ISO 55000 family of standards (and PAS -55 before it) has helped shift the emphasis of asset management from minimising cost to realising value. Parlikad et al. (2017) believes that the value of an asset essentially ((7) in Fig. 1) depends on three factors:

1. The benefits arising from the asset to the stakeholders through effective performance of the system;
2. The risks posed by the asset (its operation and its condition) to the system and its stakeholders; and
3. The expenditures incurred on the asset. Value-driven Asset management is, therefore, the aggregate of activities carried out to realise benefits (and opportunities for further benefits) while minimising both the costs and risks over the lifecycle of the asset. CSIC has developed a structured methodology to determine how an asset contributes to the system's value, how its condition can affect that value, and how value can be managed by making the right decisions. A key element in future smart asset management system is the use of sensory data (condition, environmental, and operational) in combination with other relevant data sources to develop a better understanding of value, including improving the predictability of the effect of asset condition on value. This will enable the right asset management decision to e made with better confidence to maximise asset value.

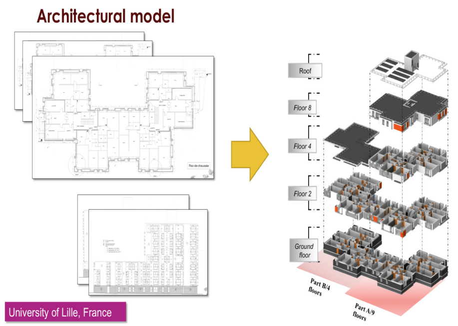
3.2 Residential Building Triolo – Case Study

The use of smart technologies like BIM, Sensor system and data mining into the asset management of an existing social housing residence - Triolo has been investigated in one research of Sunrise Smart City project of Civil and geo-Environmental Engineering Laboratory, Lille University, France. The result of this study was reported in Shahrour et al. (2017). In which, the usefulness of BIM for the social housing managers in order to overcome the increasing challenges related to the asset ageing as well as the necessity to reduce the current expenses was proved with Alileche and Shahrour (2018) research. The study reveals that the construction of the BIM model for existing buildings requires important effort for collecting data and information from existing sources (plans, sheets etc.) as well field survey. After the collection of this information, the BIM model is established according to conventional methods. The BIM model could also integrate dynamic data recorded by sensors and could be used to track and to supervise some exploitation operations (Shahrour and Jnat 2017). In addition, the construction of the BIM model for Triolo residence offers to the manager a powerful tool with ease of access to information concerning the building, the equipment, the exploitation and the comfort conditions. It could be enhanced in the future with additional information concerning property and security as well as predictive maintenance (Alileche and Shahrour 2018) and building condition evaluation (Aljer et al. 2017b).

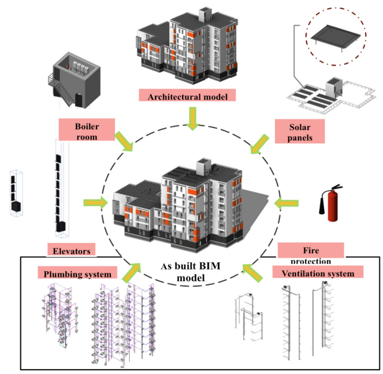
Triolo social housing residence was built in 1973 and rehabilitated in 2012. It is composed of two parts. The north part includes 9 levels and 35 dwellings, while the south part includes 4 levels and 15 dwellings. The BIM model created for this building was created from both the following available documents (1. The architectural plans in paper format (Fig. 2a), 2. Technical documents and maintenance manuals for some equipment such as the ventilation system, extractors and the boiler (Fig. 2b and 3). Maintenance documents, which concern corrective and preventive maintenance of some building equipment and the recent information database which smart sensor system and data mining system provided (Figs. 2c and 3a).

Aljer et al. (2017b) paper presented the design, fabrication and use of an innovative system for monitoring fluids consumption and comfort parameters in social housing. The specifications of the system were determined through concentration with social housing tenants as well as technical and administrative staffs. The system is based on the use of a central unit (Raspberry Pi), which tracks and controls the indoor environment (temperature, humidity, air quality, lighting, noise…), the water and energy consumption as well as the state of doors and windows (open/closed). The system uses a friendly interface, which allows users to follow real-time data and to access to historical data enhanced by information concerning the quality of the indoor environment and the expenses. The use of the monitoring system in a demonstration social housing apartment confirmed its good performances and robustness. This system is used to monitor Triolo social housing apartments as well (Fig. 2d and 3b) (Shahrour and Jnat 2017).

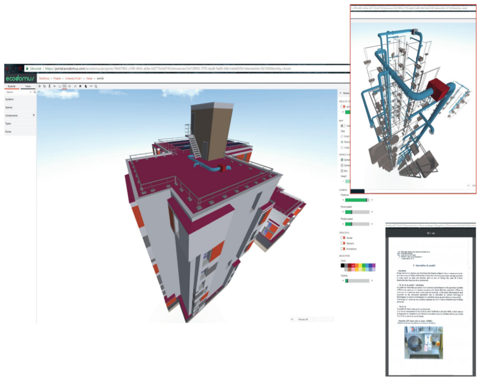
From then, the smart technologies-based management process was implemented in this residence uses a global information system put into BIM, which includes two types of information for the residence management: first is the static data obtained from the handover of Revit project to the FM Ecodomus platform it contains all the information about the building, manuals and the technical documents and second is the dynamic data received from the sensors installed in the real building it concerns the water and electricity consumption and indoor comfort parameter and alerts in case of an abnormal event or for maintenance operations. The maintenance operations were integrated into the BIM model. The preventive maintenance program for each equipment was also introduced. This system offers managers a friendly graphics platform with comprehensive information on the physical components of the residence and the maintenance operations. This cloud-based solution for the building management is based on BIM allowing the building manager to have all the necessary information for an effective management of the residence, each equipment contains its attributes (Fig. 3a). From the large database, which collected from sensor system, one friendly and simple interface is



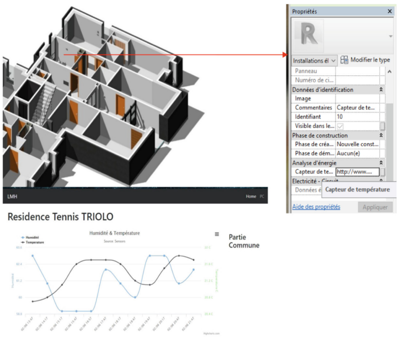
a) The architectural model of Triolo building



b) Full concept in BIM model of Triolo building



c) BIM model of Triolo building with maintenance doccuments added

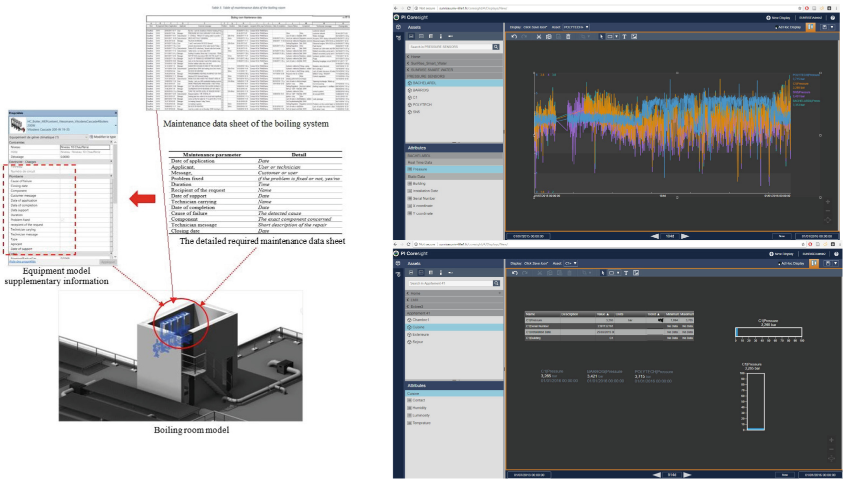


d) Sensor system's results was presented in Triolo building's BIM model

Fig. 2. BIM model for Triolo building

created on PI server (Fig. 3b) to provide for asset managers. With this general information about building and asset performance, managers could be able to have an overall view about performance of Triolo building asset management.

At the same time with technical information management of Triolo building assets, the other analysis system for managing value of social building asset. The calculations related to the annual cost of Triolo building is compared with the expenditure of others HLM buildings (Fig. 4). Form that, HLM managers could realise immediately if there some expenditure are not in the normal zone of value.



a) Building asset maintenance with BIM model

b) Real-time data mining on PI server

Fig. 3. Smart technologies for Triolo building asset management

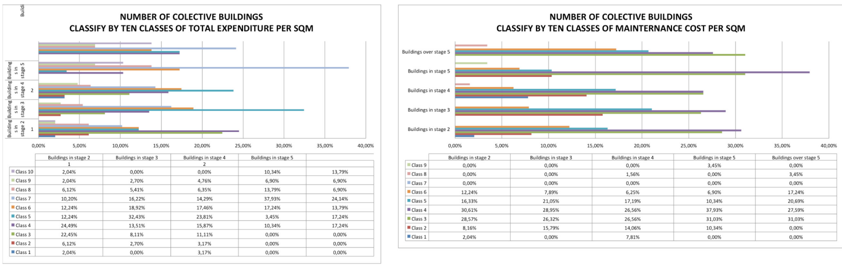


Fig. 4. Calculation of HLM buildings' expenditure

4 Conclusion

Smart asset management for social housing assets is applied to HLM asset management that Triolo building assets is named. Nevertheless, with the results which analysis above, it is not difficult to show that smart technologies be used for managing Triolo building asset are not completed like the model in Fig. 1 presented, specially for identifying assets and managing value of building asset management. The reasons of these consists are natures of social building assets such as belong to non-profit organization or age which make the auto identifying process for each asset difficulty. However, undeniable that smart technologies brought a new generation for social housing asset management. They already help all stakeholders manage their assets more automatically and effectively. From all results that showed above, full model of smart social housing asset management in near future is entirely foreseeable.

References

Alileche, L., Shahrour, I.: Use of BIM for social housing management use of BIM for social housing management. In: 17th International Conference on Computing in Civil and Building Engineering at Tampere, Finland, p. 8 (2018)

Aljer, A. et al.: Smart System for Social Housing Monitoring, (iii), pp. 17–20 (2017a)

Aljer, A., Loriot, M., Shahrour, I.: Smart system for social housing monitoring. In: 2017 Sensors Networks Smart and Emerging Technologies (SENSET), pp. 1–4 (2017b). https://doi.org/10.1109/senset.2017.8125057

Le Blanc, D., Laferrère, A.: The effect of public social housing on households' consumption in France. J. Hous. Econ. **10**(4), 429–455 (2001). https://doi.org/10.1006/jhec.2001.0300

CHOA's: What is Building Asset Management ..? (2013)

Gruis, V., Nieboer, N.: Asset Management in the Social Rented Sector: Policy and Practice in Europe and Australia (2004). http://books.google.co.uk/books?id=DvVfynlX8_0C

INTA CC Habitat: INTA CC Habitat – small introduction of social housing in France (2003). https://inta-aivn.org/images/cc/Habitat/background%20documents/CCHabitat_Housing%20social%20France.pdf

Shahrour, I., et al.: Approche inclusive de l'innovation dans le logement social - 3ème colloque LMH – Lille 1 - _LMH 260617 Residence tennis (2017)

OECD Affordable Housing Database: OECD Affordable Housing Database (2016-2017) - Public policies towards affordable housing (PH): 'Social rental housing' and 'Affordable housing programmes', Public policies towards affordable housing (2017)

Parlikad, A.K., et al.: Intelligent assets for tomorrow's infrastructure: guiding principles (2017). https://doi.org/10.13140/RG.2.2.16125.26085

Rochard, U., Shanthirablan, S., Brejon, C.: Bâtiments résidentiels: Typologie du parc existant et solutions exemplaires pour la rénovation énergétique en France (2015)

Schaefer, J.-P.: Social rental housing and housing markets in France. In: Social Housing in Europe II: A Review of Policies and Outcomes, pp. 95–104 (2008). https://doi.org/10.1080/14616710903357268

Shahrour, I., Jnat, K.: Smart Technology for a comprehensive approach of comfort in social housing. In: 2017 Sensors Networks Smart and Emerging Technologies (SENSET), p. 4 (2017). https://doi.org/10.1109/senset.2017.8125059

Barry, D.: Asset Management Point of View Management Capacity in Rural Communities (2013)

Dewick, P., Miozzo, M.: Networks and innovation?: sustainable technologies in Scottish social housing, pp. 323–333 (2004)

Gerbasi, D.: An Asset Management System for School Buildings in Quebec (2005)

Jupp, J.: Technology adoption and management innovation in construction. In: International Conference on Computing in Civil and Building Engineering ASCE, pp. 753–760 (2014)

Kitshoff, J., Gleaves, R., Ronald, G.: Social housing asset management: an innovative approach to increase productivity and performance. In: Case Studies in Service Innovation, pp. 107–109 (2012)

Love, P.E.D., et al.: A benefits realization management building information modeling framework for asset owners. Autom. Constr. **37**, 1–10 (2014). https://doi.org/10.1016/j.autcon.2013.09.007

Shah, A., Tan, T., Kumar, A.: Building infrastructure asset management?: australian practices. In: CIB World Building Congress 2004, 2–7 May 2004, Toronto, Canada, pp. 1–8 (2004)

Volk, R., Stengel, J., Schultmann, F.: Building Information Modeling (BIM) for existing buildings - literature review and future needs. Autom. Constr. **38**, 109–127 (2014). https://doi.org/10.1016/j.autcon.2013.10.023

Benchmarking for Asset Service Options and for Analysing Their Value Capture Potential

Helena Kortelainen(✉), Minna Räikkönen, Pasi Valkokari, and Susanna Kunttu

VTT Technical Research Centre of Finland Ltd., P.O. Box 1300, 33101 Tampere, Finland
{helena.kortelainen, minna.raikkonen, pasi.valkokari, susanna.kunttu}@vtt.fi

Abstract. The new era of technology and business is forcing companies to seek new business by rendering services associated with their products and by outsourcing a number of tasks or services. In addition, companies are in the middle of a transformation process driven by rapid information technology development. In this context, the use of benchmarking could provide clear direction and methods for comparing own practices and processes with peers and for learning of the practices in relevant sectors. Moreover, service providers could exploit benchmarking approaches together with their customers when looking for development needs in the asset management practices and in the asset service initiation. In this paper, we present a systematic benchmarking framework for identifying novel asset service opportunities, for optimizing the offered service solutions and, for analysing their value creation potential. The special emphasis is in the minerals and metals processing industry. The developed benchmarking approach and tool helps to compare different sites according to their operational and maintenance characteristics and to identify potential sources of value.

1 Introduction

Growing market complexity and increasing competitive intensity are forcing companies to seek new business by developing services associated with their products. Numerous companies are striving with mixed results to become "solution providers" or "value partners" by adding services to their portfolio of tangible products (Oliva and Kallenberg 2003). For example, technology providers have been actively developing novel asset based services such as condition monitoring and remote control. In addition, global product manufacturing companies have shifted their focus from product delivery to value adding asset life-cycle services. The literature, however, is surprisingly sparse in describing to what are the services to be delivered, to what extent services should be integrated and how the transition from products towards services should be managed (Neu and Brown 2005; Oliva and Kallenberg 2003).

Competition in international markets has led to a situation where industrial services, for example maintenance services, are outsourced to greater extent. Many regional

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 67–74, 2020.
https://doi.org/10.1007/978-3-030-48021-9_8

clusters have fragmented into global value chains, in which geographically scattered companies specialize in specific activities. As such, the competition has transferred from taking place between individual companies into competition between business networks or, as the increasing interest in sustainability and industrial ecology suggests, between industrial ecosystems, (Ashton 2009; Baldwin 2016). Companies are also in the middle of a transformation process driven by rapid information technology development. Data, information, knowledge and analytics, and the use of data in decision-making and processes, are in the core of this transformation process. The different industry sectors are currently learning to exploit more and more data and information in their business. The new IoT-based service concepts can introduce smart, automated products to the manufacturing business (Pletikosa Cvijikj and Michahelles 2011). These kinds of developments can be seen to have a vast business potential for equipment manufacturing companies or information service providers both in consumer and business-to-business markets (see e.g. Pletikosa Cvijikj and Michahelles 2011).

In this new era of technology and business, the benchmarking could be utilized more effectively than now. Benchmarking can be described as a tool to compare product and service attributes, quality attributes, operations and processes. Benchmarking is the most preferred tool for business performance improvement (Clarke and Manton 1997; Jain et al. 2008; Asrofah et al. 2010). Benchmarking can help understand how the best-in-class companies carry out their business activities and eventually leads to learning how to deal successfully with competition from these companies (Meybodi, 2009). Customized service delivery, digitalization, automation and integration of customers' processes offer new possibilities to create and capture the value revealed by benchmarking. Thus, the use of benchmarking could provide clear direction and methods for learning from customers by initiating value-added services that exceed their expectation and - moreover - help to sustain a company's performance and competitiveness in the long-term. However, benchmarking is still an underestimated concept in the asset service development and companies do not recognize the whole potential of benchmarking in asset service identification and valuation.

2 Approach

This paper describes a systematic benchmarking framework for identifying novel asset service opportunities, for optimizing the offered service solutions and, for analysing their value capture potential. The framework was realised and tested in the Fleet information network and decision-making research project that aimed at finding new ways to gather, analyze and understand and utilize information in fleet level operations and create extended service solutions around information (Kortelainen *et al.* 2017). The project was conducted as a part of the DIMECC Service Solutions for Fleet Management (S4Fleet) program during the timeframe 2015–2017. The project was carried out in close co-operation between participants from Finnish universities and ICT and manufacturing companies.

This paper aims at answering the following research question: *How the benchmarking can be used to identify asset services and to analyze their value capturing* *potential?* The research question is tackled by using qualitative research methods (Yin 1994). Qualitative research approaches emphasizing empirical field data are beneficial when the topic is in its early stage and extant theories are limited (Creswell 2013; Eisenhardt and Graebner 2007). Creswell (2013, p. 160, 183) presents the data collection methods and analysis procedures in qualitative research, including data collection methods such as interviews and observations, and procedures from data managing as representing and visualizing e.g. with matrices, trees, and propositions. The state-of-the-practice has been observed through numerous discussions and interviews with the individual companies, and through observations and results from other recent research projects and literature that provided insight into the state-of-the-practice.

In the case study part of the research, companies have laid down the current development challenge, and the co-creation with the company representative(s) and researchers has led to the result. Case studies have been the dominant methodology used by qualitative researchers. It is considered a particularly useful approach for increasing understanding of topics that are previously under-investigated (Gummesson 2000) and in situations, where there are complex and multiple variables and processes (Yin 1994). In this research, the collision of research and practice has enabled the combination of different views and generated scientifically interesting and practically relevant research results. The benchmarking framework presented in this paper is the result of the fruitful co-operation between VTT Technical Research Centre of Finland Ltd and Outotec. Outotec provides technologies and services related to minerals and metals processing, energy and water.

3 Benchmarking for Asset Service Options and for Analysing Their Value Capture Potential

Benchmarking is the process of identifying "best practices" in relation to both products and the services by which the products and services are created and delivered. The search for "best practices" can take place both inside a particular industry, and in other industries, for example, the lessons learnt from other industries. (Asrofah et al. 2010) Benchmarking encourages a company to become open to new methods, ideas, processes, and practices to improve effectiveness, efficiency, and performance (Deros et al. 2006; Asrofah et al. 2010). Though a number of definitions of benchmarking within the literature, they all essentially share the same theme as shown in Table 1.

Accordingly, the objective of benchmarking is to understand and evaluate the current position of a business or organization in relation to the "best practice" and to identify areas and means of performance improvement. Benchmarking has been defined as a continuous, systematic process for evaluating the products, services, and work process of organizations that are recognised as representing the best practice, for the purpose of organizational improvement (Sarkis 2001).

Since benchmarking focuses on continuous improvement of specific product characteristics or processes, which are critical to the success of a firm's business strategy, it is recognised as a cost-and time-effective method in meeting competition (Watson 1993). Furthermore, benchmarking can also be described as a structured

Table 1. Benchmarking definition in the literature (adapted from Asrofah et al. 2010).

Literature	Definition
McNair and Kathleen 1992	Benchmarking is an external focus on internal activities, functions, or operations in order to achieve continuous improvement
Watson 1993	Benchmarking is a continuous search for, and application of significantly better practices that lead to superior competitive performance
Partovi 1994	Benchmarking is the search for the best industry practices, which will lead to exceptional performance through the implementation of these best practices
Bhutta and Huq 1999	Benchmarking is first and foremost a tool for improvement, achieved through comparison with other organizations recognized as the best within the area
Vermeulen 2003	Benchmarking is the process of identifying, understand- ing, and adapting outstanding practices from within the organization or from other businesses to help improve performance
Yusuff 2004	Benchmarking is a means for assessing industrial competitiveness
Hurreeram 2007	Benchmarking is a systematic and continuous process of searching, learning, adapting, and implementing the best practices from within own organization or from other organizations towards attaining superior performance

approach where the structure of the benchmarking process is often developed by the development of a step-by-step process model, which provides a common language within organizations (Spendolini 1993).

4 Benchmarking Tool

Within the project, VTT and Outotec co-operated in developing a benchmarking approach and a tool for finding maintenance service opportunities and ways of visualizing potential sources of customer value (Kortelainen *et al.* 2017). Both the service provider's and site's viewpoints are considered. As implied in the various definitions offered (Table 1), benchmarking is a continuous process. Thus, the process of the developed approach can be presented in the form of benchmarking wheel where major process steps are linked together like the spokes of the wheel. The approach follows the basic principle of PDCA (plan, do, check, act) and it is applicable to different industry sectors (Fig. 1).

Step 1. Plan the benchmarking study. Regarding the benchmarking study, there needs to be a balance between time, resources and the level of information and understanding. The decision of what to benchmark must be aligned with the company's strategic direction. Moreover, all the assumptions and boundaries of the analysis need to be defined as well. Both the service provider's and site's representatives for carrying

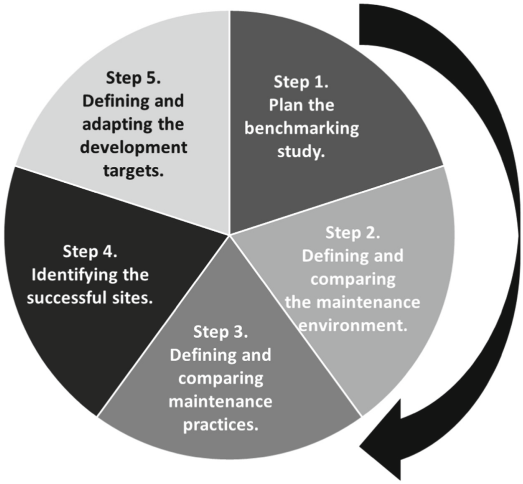


Fig. 1. The benchmarking wheel.

out the data collection and analysis should be nominated to make benchmarking more transparent. The aim is to conduct the data gathering during normal business negotiations.

Step 2. Defining and comparing the maintenance environment from both site and production line point of view. The second step is to recognize companies that are similar enough to be able to learn best practices from each other. As the focus is on the maintenance service development and value capture, the benchmarking method categorizes sites and production lines according to their maintenance environment. The tool includes questions describing the features related to the maintenance function, namely maintenance policy and maintenance activities. Results from five current sites define the first five reference environments of the tool. New sites are compared to these reference sites based on the similarity index. Maintenance environment of the new site will be the one having highest similarity index. Maintenance environments should be redefined after a certain data collection period when more data is available to make conclusions about relevant reference environments. From a service provider's point of view, the maintenance environment related aspects are external and they cannot be controlled by a service provider.

Step 3. Defining and comparing maintenance practices from both site and production line point of view. The second step is to define the level different maintenance practices of the site in question. The tool includes questions regarding the maintenance practices and compares the sites belonging to same maintenance environment (Fig. 2).

Step 4. Identifying the successful sites. This step aims to identify the best sites in each maintenance environment to describe the best practice performance. Availability, maintenance costs and replacement value are key performance indicators (KPIs) indicating the success of applied maintenance practices. In the tool, KPIs are collected and they can be used as enough data is gathered and available.

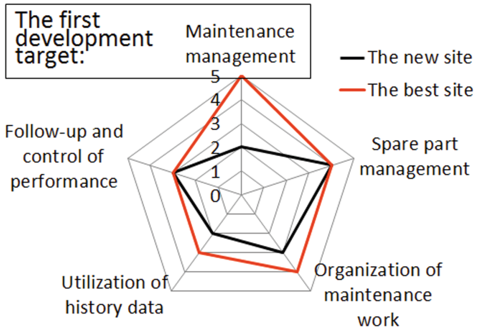


Fig. 2. Comparison of maintenance practices between the best site and the site under study.

Step 5. Defining and adapting the development targets for the site in question. The final benchmarking step involves adapting the best practices and implementing specific improvements. Adapting best practices is not to be confused with copying best practices. Best practices learned from others must be adapted to a company's strategy, value proposition, culture, technology and human resources. Development targets are appropriate for this phase. Some improvements will be immediate or short term, others will be long-term. The KPIs can also be used to track performance improvements.

5 Value Capture Potential for Network Actors

Traditionally asset owners collect data, create information and knowledge, and gain wisdom that is applicable and useful for their business and the service providers do not have the access to the relevant data (Kortelainen et al. 2018…). The value potential of service relationships and the competences available in the network often hidden (Ahonen et al. 2010; Ali-Marttila et al. 2016). The benchmarking tool helps in identifying and visualizing the potential sources of value. With the approach based on categorizing sites to comparable units and benchmarking them against each other, the service provider is able to improve its capability in (Kortelainen *et al.* 2017):

- Showing improvement potential in asset management and making recommendations of applicable asset management policies,
- Facilitating sales by optimizing the customer specific product and service offering, and
- Concretizing the customer value of the service provision.

The developed benchmarking tool allows Outotec to utilize installed base knowledge as well as insights into potential value sources on a very concrete level. The customer's actual needs are defined in close cooperation with the service providers and possible other network actors during the site assessment. The assessment gives insight into what types of services the customers demand the most and into the value capture potential.

6 Conclusions and Future Research

This paper proposes a systematic benchmarking framework for identifying novel asset service opportunities, for optimizing the offered service solutions and, for analysing their value creation potential. The developed benchmarking approach promotes service providers' ability to recognize improvement potential in customer's asset management practices and the ability to find improvement actions for current situation. The developed tool allows a value-based sales process and, more specifically, the matching of company's service offering against the customer's actual needs. It also support the development of a transparent sales process defined in close cooperation with the customer. The tool can also be used in company's internal product development being faced by the key challenges of customers. Addressing the service product portfolio accordingly will give insight into what types of services the customers demand the most.

In the near future, the research continues by defining the service offering based on benchmarking together with the associated value elements and by generating appropriate business models. To fulfil this, the research focuses on linking the substance areas in the benchmarking scheme to specific measureable quantities, or to qualitative factors that help to evaluate the impact and value of the service delivery. In this case, benchmarking contributes especially to a value based sales process and produces a set of verified business models that drive a feasible service offering.

References

Ahonen, T., Reunanen, M., Pajari, O., Ojanen, V., Lanne, M.: Maintenance communities - a new model for the networked delivery of maintenance services. Int. J. Bus. Innov. Res. **4**(6), 560–583 (2010)

Ali-Marttila, M., Marttonen-Arola, S., Ylä-Kujala, A., Ukko, J., Rantala, T., Sinkkonen, T., Pekkola, S., Saunila, M., Pekkarinen, O., Kärri, T.: Stagewise process towards collaborative and value-driven decisions in maintenance networks. In: Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015). Lecture Notes in Mechanical Engineering, pp. 41–50. Springer, Heidelberg (2016)

Ashton, W.S.: The structure, function, and evolution of a regional industrial ecosystem. J. Ind. Ecol. **13**(2), 228–246 (2009)

Asrofah, T., Zailani, S., Fernando, Y.: Best practices for the effectiveness of benchmarking in the Indonesian manufacturing companies. Benchmarking: Int. J. **17**(1), 115–143 (2010)

Baldwin, R.: The Great Convergence: Information Technology and the New Globalization. Belknap Press of Harvard University Press (2016)

Bhutta, K.S., Huq, F.: Benchmarking – best practices: an integrated approach. Benchmarking: Int. J. **6**(3), 254–268 (1999)

Clarke, A., Manton, S.: A benchmarking tool for change management. Bus. Process Manag. J. **3** (3), 248–255 (1997)

Creswell, J.W.: Qualitative Inquiry & Research Design: Choosing Among Five Approaches, 3rd edn. SAGE Publications, Inc. (2013)

Deros, B.M., Yusof, S.M., Salleh, A.M.: A benchmarking implementation framework for automotive manufacturing SMEs. Benchmarking: Int. J. (2006)

Eisenhardt, K.M., Graebner, M.E.: Theory building from cases: opportunities and challenges. Acad. Manag. J. **50**(1), 25–32 (2007)

Gummesson, E.: Qualitative Methods in Management Research. Sage, Thousand Oaks (2000)

Hurreeram, D.K.: Manufacturing strategy auditing for garment making companies. Benchmarking: Int. J. **14**(3), 272–288 (2007)

Jain, R., Yadav, O.P., Rathore, A.P.S.: The propagation of benchmarking concepts in Indian manufacturing industry. Benchmarking: Int. J. **15**(1), 101–117 (2008)

Kortelainen, H., et al.: Fleet service creation in business ecosystems – from data to decisions. Fleet information network and decision-making. VTT Technology, p. 309 (2017)

Kortelainen, H., Happonen, A., Hanski, J.: From asset provider to knowledge company - transformation in the digital era. In: Proceedings of the 11th World Congress on Engineering Asset Management (WCEAM 2015). Lecture Notes in Mechanical Engineering. Springer, Heidelberg (2018, in press)

McNair, C.J., Kathleen, H.J.: Benchmarking: A Tool for Continuous Improvement. Harper Collin, New York (1992)

Meybodi, M.Z.: Benchmarking performance measures in traditional and just- in-time companies. Benchmarking: Int. J. **16**(1), 88–102 (2009)

Neu, W.A., Brown, S.W.: Forming successful business-to-business services in goods-dominant firms. J. Serv. Res. **8**(1), 3–17 (2005)

Oliva, R., Kallenberg, R.: Managing the transition from products to services. Int. J. Serv. Ind. Manag. **14**(2), 160–172 (2003)

Partovi, F.Y.: Determining what to benchmark: an analytical hierarchy process approach. Int. J. Oper. Prod. Manag. **14**(6), 25–39 (1994)

Pletikosa Cvijikj, I., Michahelles, F.: The toolkit approach for end-user participation in the Internet of Things. In: Uckelman, D., Harrison, M., Michahelles, F. (eds.) Architecting the Internet of Things. Springer, Heidelberg (2011)

Sarkis, J.: Benchmarking for agility. Benchmarking: Int. J. **8**(2), 88–107 (2001)

Spendolini, M.: How to build a benchmarking team. J. Bus. Stragegy **14**(2), 53–57 (1993)

Vermeulen, W.: Benchmarking as an enabler of business excellence in the South African financial sector. Benchmarking: Int. J. **10**(1), 65–72 (2003)

Watson, G.H.: How process benchmarking supports corporate strategy. Plan. Rev. **21**(1), 12–15 (1993)

Yin, R.: Case Study Research: Design and Methods, 2nd edn. Sage publishing, Beverly Hills (1994)

Yusuff, R.M.: Manufacturing best practices of the electric and electronic firms in Malaysia. Benchmarking: Int. J. **11**(4), 361–369 (2004)

Offshore Assets in the Transition from 'Traditional Business Practice' to 'Connected Business Eco-systems'

Jan Helge Viste[1(✉)] and Jayantha P. Liyanage[2]

[1] University of Agder, Kristiansand, Norway
jan.h.viste@uia.no
[2] University of Stavanger, Stavanger, Norway
j.p.liyanage@uis.no

Abstract. In many industrial sectors we have already begun to see how new business ecosystems gradually emerge turning the traditional business practices upside down. Particularly within relatively conventional sectors, such as upstream oil and gas (O&G) and land-based process industry, we expect to see substantial changes in the years to come. In this paper, we elaborate on principal challenges related to major structural changes and core technology platforms for the oil industry identifying critical features and elaborating on emerging operational principles. This is based on a critical review of latest industry conditions and practices that are targeted towards the creation of new business models and processes due to strong industrial forces. Based on previous experiences from various operations and maintenance improvement projects as well as several topside new-build projects (O&G platforms) on the Norwegian Continental Shelf, two examples related to Integrated Operations (IO) are given to demonstrate gradual transformation of the business eco-system, and thus to reflect how the industry sector could be transformed into a new industrial paradigm.

1 Introduction

In relatively conservative industrial sectors, such as oil & gas production, the transition from a traditional business practice to a connected business eco-system is a demanding process, even though the change will be unavoidable (Liyanage 2012). This transition will involve adapting new business models that follows the principles of 'interconnected eco-systems', and strategic integration of 'information and operational technologies' (i.e. IT + OT). With reference to the emerging business ecosystems within the oil & gas production sector particularly due to economic and socio-political forces, it is interesting to explore for instance how such new eco-system would look like and what kind of players and roles can exist to cope with new demands and ongoing changes. In order to succeed with smarter, safer, and efficient management of modern asset portfolios, salient features and capabilities of the core technology platform also need to be considered, both within the IT domain (Information Technology, i.e. the traditional office network) and the OT domain (Operation Technology, i.e. the domain of Process Control Network). This is also important since new core technology

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 75–86, 2020.
https://doi.org/10.1007/978-3-030-48021-9_9

platforms can enable the creation of next generation core business processes enabled by for instance, digitalization trends and Industry 4.0. This may subsequently lead to changing of roles of different players in new business eco-systems within engineering and operations of oil & gas assets. Thus it is important to review challenging issues related to the two core technology platforms, firstly exploring the main aspects and afterwards highlighting issues that need to be addressed in order to prepare for a safe and secure asset management practice. Based on experience and observations within the offshore oil & gas sector, two examples from Integrated Operations (IO) on the Norwegian Continental Shelf are discussed at the end to provide a brief understanding on ongoing changes. The first example covers gradual changes towards a more effective and efficient maintenance regime that can be carried out based on new capabilities of operating assets and eco-system, and the second one highlights different options regarding the interplay in an oil business ecosystem between a manufacturer/vendor and the upstream oil E&P company when engineering new off-shore assets.

2 Developing Connected Business Ecosystems Around Smart Assets in the Oil and Gas Sector

In many industries, and around many assets, products and services, we have already seen how new business ecosystems have turned traditional business practice upside down. Similarly, also within the upstream oil and gas exploration and production (E&P) sector major production companies such as Exxon Mobil, BP, Total, Shell, Equinor, etc. expect to initiate substantial changes along the same development track in the years to come. The *benefit* of working in a network of organizations is that strategic cooperation can deliver greater value for the firm than the value it might have created alone (Makinen and Dedehayir 2012; Liyanage 2013). However, the transition process from conservative and traditional to a new and 'connected' is largely regulated by a number of conditions that vary from pure technological to structural issues (Fig. 1).

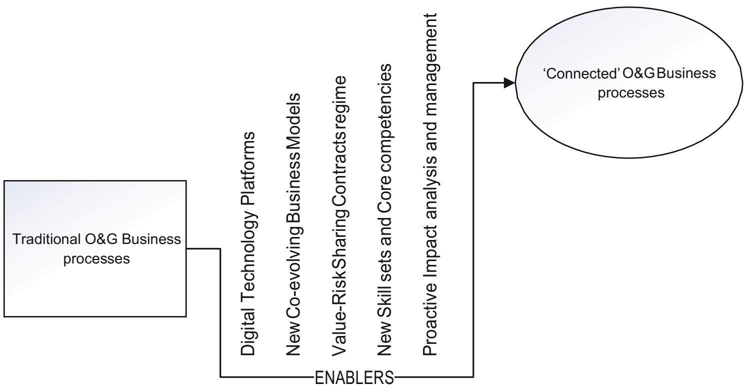


Fig. 1. The O&G business is subjected to a demanding transitional process.

The EU funded project *Sustainable value creation in Manufacturing networks* (‘*SustainValue*’) defines a business ecosystem as “*the network of organizations – including suppliers, distributors, customers, competitors, government agencies and so on – involved in the delivery of a specific product or service through both competition and cooperation. The idea is that each business in the ‘ecosystem’ affects and is affected by the others, creating a constantly evolving relationship in which each business must be flexible and adaptable in order to survive, as in a biological ecosystem*” (SustainValue Deliverable D1.1 2011, p. 7). Mean while, Makinen and Dedehayir (2012, p. 1) elaborates somewhat brief that “*business ecosystems describe the network of firms, which collectively produce a holistic, integrated technological system that creates value for customers*”. At the same time, Bharadwaj et al. (2013, p. 474) discuss as to how this also can influence the business strategy: “*In a digitally intensive world, firms operate in business ecosystems that are intricately intertwined such that digital business strategy cannot be conceived independently of the business ecosystem, alliances, partnerships, and competitors. Furthermore, the use of digital platforms enables firms to break traditional industry boundaries and to operate in new spaces and niches that were earlier only defined through those digital resources*”.

In 1994, Venkatraman (1994) published an article in the *MIT Sloan Management Review* which developed a framework for IT-enabled business transformation. This article elaborated on a business transformation framework and underlined that realizing benefits from IT applications require corresponding changes in how firms are organized and how they are interconnect with others in extended business networks (Venkatraman 1994). Already more than 30 years ago, the ideas of ‘Business Network Redesign’ and ‘Business Scope Redefinition” to reach higher degrees of business transformation were described, even if the term ‘business ecosystem’ was not used at that time, as far as we know. In his new book, Venkatraman states that the digital future relies much on the; big data, social web, mobile apps, internet of things (IoT), robotics, and cognitive computing (Venkatraman 2017). When we think about the upstream oil and gas E&P companies, all of those six criteria seems to match quite well, perhaps with some minor exceptions. If we then presume that those companies are going to be more smarter by resorting to digitalization and digital inter-connections than before, a major issue that arises then is that *how* this can or should be done in a safer and secure manner. The change is naturally considered to be important and inevitable in order to further develop and succeed. However, one of the key findings in the EU project SustainValue (2012) was that business model innovation is a key factor to the successful development of new industrial systems and implementation of new or modern business processes. At the same time, it was observed that there are strongly emerging new demands towards development of concrete business cases with due consideration on sustainability, i.e. a premise to design business models and processes that will integrate and foster linkages between economic, social and environmental value from new industrial activities. With the current high focus on carbon foot print related to engineering and operations of Industrial assets, environmentally friendly energy production, as well as reductions in CO_2 emissions, the *environmental value* will be even more critical for the oil industry in the future.

2.1 Challenging Times Across the Sector

The upstream O&G sector has been subjected to various transformations over the last decades. Shuen et al. (2014) has identified at least five factors that have triggered a climacteric for upstream entities:

- Increasing demand for oil and gas requires significantly increased production
- Unconventionals: new technologies, new geographies, new processes, new opportunity and the need for ubiquitous learning
- The rise of national oil and gas companies, large independents and service companies alongside the supermajors provides both competitive challenges and collaborative opportunities, increasing the complexity of strategic decisions
- Managing the human resource strategy in its parts and interactions
- Managing health, safety, security and environmental risks throughout the business ecosystem

When developing new business models to operate assets and creating adaptive strategies to stay competitive, the mastery of knowledge, sharing of competencies, and other unique technological and operational capabilities are very critical to ensure steady developments in various parts of the business ecosystem (Teece 2007; Liyanage 2012). We think this is also true for the oil and gas sector that has been facing challenging times especially during the last 2 decades.

2.2 Potential Structural Changes

It has sometimes been claimed that the upstream oil and gas companies actually are in the *transportation business* (!): just transporting oil and/or gas from a reservoir (deep down in the ground) to the processing plant. Of course, this is a formidable simplification. Some others consider the upstream oil companies to primarily be *knowledge companies* because that is the main issue they manage in order to be able to do the complex “oil transportation”. From an evolutionary perspective, these companies can be seen as *energy delivering companies*, which primarily have been based on hydrocarbons until now. However, in the light of major political and social forces and subsequent structural changes within energy production portfolios, that may in the near future be more based on a strategic energy-mix from offshore wind, solar energy, and/or hydrogen fuel cells.

According to Venkatraman (2017), success in the digital period demands a focus on three players, namely: *Industry incumbents* (historical, traditional competitors in your industry), *Tech entrepreneurs* (digitally born companies disrupting and reorganizing the business world), and *Digital giants* (huge companies that progressively have extended their influence beyond their traditional industry). Until now, it seems more appropriate to categorize the upstream oil and gas companies as *industry incumbents*, mainly in phase of *digital experimentation* but also seeing experiencing *collision at the core*. As observable, a *tech entrepreneur* can acquire a position challenging the traditional role of an upstream oil company, and can involve with national oil companies in developing and co-owning the license. On the other hand, a large drilling equipment production company can transform by starting to deliver drilling services, co-owning strategic rigs, and capturing the market of conventional drilling contractors. One big question that remains in this uncertain context is that *when* and *how* upstream oil companies will tackle the third structural transformation phase towards a digital age. This can well be a dramatic transformation that perhaps only some of them survive. A clear factor that potentially drive them forcefully into a critical transformation phase earlier than expected, is the growing demand for more environmental-friendly and economical energy supplies than pure hydrocarbons.

As we have seen in some other industrial sectors, it can also be speculated that for those companies in the oil and gas sector who deliver various products and services, emphasis may also turn from delivering traditional *products* and *services* (i.e. with a single-firm focus) into rather delivering *production platforms/energy assets* and *fully integrated solutions* (i.e. integrative ecosystems) as Venkatraman (2017) has termed, through non-traditional collaborations with strategic partners introducing a dominant structural impact across the sector. It is not easy to predict as to how such a radical transformation may subsequently look like and thus influence the entire sector.

2.3 Concerns Related to IT and OT

There are a number of ongoing initiatives that underline a distinctive difference between the IT domain (Office Network, typically based on Ethernet, office applications like Word and Excel, email and other standard Internet/Intranet technology) and the OT domain (Process Control Network). For instance, the Norwegian Oil and Gas Association ("Norsk olje&gass") is the professional body and employer's association for oil and supplier companies engaged in the field of exploration and production of oil and gas on the Norwegian Continental Shelf (NCS). It commits to solve common challenges for the members and to strengthen the competitiveness of the NCS. The organization has written a guideline for information security for process control, safety, and support ICT systems (Norwegian Oil and Gas Association 2016). The guideline is not mandatory but over the years it has become an informal industrial standard within the oil and gas sector (at least in Norway) for basic network topology.

The Information Security Gateway is the main responsible for the security function, separating the two IT and OT domains. The network topology could be quite similar to network topologies for other companies with a need for an "inner" and more secure network than the office network. However, for the upstream oil companies, the Process Control Network typically will be the one responsible for control and safety of a large offshore oil installation, and a hostile hacker attack can generate disastrous effects. Hence, in the Guideline 104 (Norwegian Oil and Gas Association 2016), there is a security baseline requirement that has crucial impact related to this Information Security Gateway ("Firewall"), namely "*ISBR 10 Change management and work permit procedures*" which states that "*Change management and work permit procedures shall be followed for all maintenance and changes in the PCSS ICT systems and networks.*" (Norwegian Oil and Gas Association 2016, p. 22). In practice, this implies that for all access to the Process Control Network (ex. related to a maintenance intervention, this for instance can be related to having access to a device or server for some diagnostic tasks), will need a work permit to be activated (often by a manual procedure), so that the control room operator is aware of what is going on. For pure

"Read" access to a device or unit, this could be deemed unnecessary. But in practice the functionality for such an access, for instance to a valve, is not restricted to be a pure "Read", so the maintenance monitoring role could in fact also by mistake, or as a hostile action, open or close the valve at the wrong time that can have fatal consequences.

The process control networks are normally not based on standard IT, but different communication protocols and technology that has mainly emerged from the automation domain. Even for new-build of platforms, the old HART protocol, i.e. digitally overlaid communication on an analog 4–20 mA two-wire thread (RS-232) to each device/sensor/transmitter, can still be chosen as the basic technology, often supported by RIO (Remote I/O) cabinets placed at strategic places on the platform to reduce cabling. Some oil installations have chosen data bus technologies instead of this. In process control systems, this market is dominated by FOUNDATION Fieldbus and PROFIBUS PA (note that PROFIBUS should not be confused with the PROFINET standard for Industrial Ethernet). However, within the oil and gas sector, the companies typically have external or internal recommendations/guidelines/requirement with restrictions on number of devices on each bus (e.g. maximum four devices). Hence, the gain is often small (compared to HART/RIO), and some oil companies have in fact returned back to using HART and RIO instead. Industrial Ethernet would be a good alternative, i.e. using more Internet technology also in the Process Control Network, but it is difficult to find projects among the most major players on the Norwegian Continental Shelf until 2017, where Industrial Ethernet has been chosen (as principal technology).

Moreover, when new offshore installations are going to be built, the EPC (Engineering, Procurement, Construction) contractor often develop installation specific network topologies adapted from Norwegian Oil and Gas Association (2016). However, there exist some main challenges regarding ICT for modern asset management purposes within such network topologies. Some selected concerns are highlighted in Table 1.

In general, a great number of companies face practical difficulties in information management and communication processes. Liyanage (2012) argues that this can lead to a loss of valuable information and knowledge in thousands of databases that can otherwise be used for an organization's advantage. We believe that also in the oil and gas sector (like in many industrial sectors), the asset management practices inherits a number of conventional elements that can introduce major barriers towards smarter assets (see for instance (Liyanage 2012; Liyanage 2013)).

It can be argued that in order to achieve business effects, technology is not all that matters. As Venkatraman (2017) emphasizes; "*One emerging and particularly dynamic area of change involves the way humans and machines are interacting now and how that relationship may evolve in the future. This new frontier may provide you with new ways to orchestrate across ecosystems, as well as new avenues to co-create with others within ecosystems*" (Venkatraman (2017), p. 166). There are many human and organizational aspects that can be even more important in both IT and OT context. In the oil industry where there is a dominant focus on technological lead, there is also a significant potential to undermine critical soft issues and soft interfaces. This is

Table 1. Notable core ICT concerns in the asset transition process.

Issue	Concerns
Core OT technology	• Is HART and FieldBus outdated? • Is Industrial Ethernet the solution? • How to integrate wireless sensor networks (ex. WirelessHART or ISA.100) - *Note that wireless sensor network is different from WiFi*
Separation between IT and OT	• How to create more dynamic solutions? • Possible to mitigate the problems using products as SHIELD (from the company IPnett), which digitalizes the work permit process and provides the oil company with full control over all accesses at all times, leaving a complete audit trail
Retrofitting asset management solutions	• What applications need to be installed in the Process Control Network? • Can it be accessed from the Office Network (or from a vendor via Internet)? • Will it generate various alternate routes through (or outside) the security gateways?
Work permit based access to equipment, servers, and transmitters on OT	• No separation between pure monitoring and control functions in the applications (software)? • Remote access from internal (or external) expert centers restricted? • Remote access from remote vendors and experts restricted? • Hacking fear? • Ideally two ports, one for control, and one for harmless monitoring and diagnostics (e.g. pure READ-functionality)?
Security	• New security regime with Industrial Ethernet (and also Industrie 4.0) compared with current configuration? • Will introduction of more sensors and measurement instruments on the equipment lead to more leakage points? • Will more maintenance on the sensors and instruments themselves (e.g. calibration) have a negative impact on HSE and/or maintenance load?

particularly true when it comes to the development and implementation of new business ecosystems around Smart assets.

2.4 New Competitive Roles and Positions in the New Business Ecosystem

As Makinen and Dedehayir (2012) highlights, the typical players in a typical business ecosystem includes; suppliers, complementors, system integrators, distributors, advertisers, finance providers (e.g. venture capitalists, corporate investors, investment bankers, and angel investors), universities and research institutions, regulatory authorities and standard-setting bodies, the judiciary, and customers. For the Norwegian oil and gas sector, one could more specifically add employee unions, safety authorities, service contractors, and EPC contractors (Engineering, Procurement, Construction) to the list of important players. According to Libert et al. (2017), in such a business model there are four different types of roles:

- *Asset builders* deliver value through the use of physical goods. These companies make, market, distribute, sell, and lease physical things.
- *Service providers* deliver value through skilled people. These companies hire and develop workers who provide services to customers for which they charge.
- *Technology creators* deliver value through ideas. These companies develop and sell intellectual property such as software, analytics, pharmaceuticals, and biotechnology.
- *Network orchestrators* deliver value through connectivity. These companies create a platform that participants use to interact or transact with the many other members of the network. They may sell products, build relationships, share advice, give reviews, collaborate, and more.

A major challenge for upstream oil companies under the current transformative context is that they need to early decide which role they have to undertake within the new business ecosystem emerging in the energy production segment. As we see, they can undoubtedly take the key role within the oil and gas sector related ecosystem particularly due to current dominating positions, i.e. the *keystone firm* role (Makinen and Dedehayir 2012), which others also referred to as *platform leader* or *ecosystem leader,* who plays the role of regulating the overall function of the eco-system and as a consequence its actions influence the success of all other members, including its own. Then they also should be able to act as *network orchestrators* (Libert et al. 2017), or *industry incumbents reinventing at the root*, (Venkatraman 2017) to cope with new strong demands. The role in the ecosystem can largely be influenced by the traditions, power, funds, position, as well as core capabilities that are needed to negotiate and undertake a vital and a competitive role. The study done by Makinen and Dedehayir (2012) underlines the vital role of keystone firms in designing the business ecosystem and guiding its evolution. Co-evolution among ecosystem members, a product architecture design that pursues modularity, and the degree of control assumed by the keystone firm emerge as important internal factors, while changes in the ecosystem's social, economic, technological, and competitive environment serve as external factors governing the evolution of the ecosystem. These dimensions will truly re-define the distribution of power and position within the new business eco-system related to energy production.

3 Some Examples of Changing Ecosystems in the Oil and Gas Sector

In light of the review done in aforementioned sections, two separate examples are given below to demonstrate as to how current trends in oil & gas business influence the transition from "traditional business" to "connected business eco-systems". In these examples, just a few of the most important players were chosen looking into how the underlying changes can be seen from different perspectives.

3.1 Example #1: Condition Monitoring and Maintenance Planning

For an upstream oil and gas E&P company who owns a portfolio of producing platforms, traditional business model has been such that the oil company itself recruit and develop the crew for each specific asset (offshore or onshore). The crew undertakes necessary roles and responsibilities for designated equipment of a given asset. Often experts can also be found with specific expertise on specific technical domains within the operational organization who can assist on demand. Some oil companies also have had centralized condition monitoring support centers for selected equipment, typically for heavy rotating machinery, but also for other types such as valves, automation systems, fire & gas detectors, etc. More centralized condition monitoring approach has been growing over the last years, primarily due to increasing attention towards more proactive maintenance practices. In fact, more proactive maintenance implies new competencies, often creating a strategic need to actively involve manufacturers or vendors and other external domain experts in diagnostic and prognostic tasks. During the current transformation period it is not uncommon to find situations that oil companies resort to combined solutions where in-house diagnostic and prognostic competencies and processes are coupled with external expertise. Such situations allow external vendors to utilize their own core resources meaningfully, for instance connecting their own remote monitoring and expert centers, and providing expert advice remotely. Naturally, vendors have the ability to perform distinctive diagnostics and prognostics tasks for several other companies simultaneously, within the capacity limits. However, this can often be met with various practical challenges due to sensitivity and security reasons. Moreover, various technical and operational issues still remain unresolved, for instance related to data integration within the ERP systems, maintenance work process optimization, data-driven decisions, etc.

Arguably, the whole oil industry ecosystem around diagnostic and prognostic services still appear to be quite immature, and much work remains to be done in order to establish more robust business models involving different stakeholders who can share functional responsibilities and underlying risks, and thus have dedicated roles in value creation process. One main reason for the current unsatisfactory situation is that the classical wisdom related to IT (information technology) and the OT (operational technology) platforms that the upstream oil companies have established over the years have created a major bottleneck. These domains still remain largely separated partly due to independent technological dominance, partly due to absolute conditions demanded for safe and secure operations, and partly due to lack of right level of competencies.

3.2 Example #2: Purchasing Models for New Builds

During the development of new oil fields and procurements related to new topsides, it is not uncommon to experience that the central discussions nowadays more and more trending to focus on, for example: what do we want to purchase from our main manufacturers/vendors? Physical equipment and units, or a "functionality"? For example, do we want to purchase "2 × 100% LM6000 gas turbines", or would it be better to purchase "stable 50 MW power supply for the coming 20 years"? The supportive reasoning for such conventional vs. novel solutions goes back and forth with lengthy discussions. The decision in such contexts has not always been straight forward since it is always affected by many internal factors, for instance; *what is normally done by the purchasing department?, who is better to assure high availability; us or them?, should more responsibility for maintenance be allocated to the general maintenance contractor (instead of individual manufacturers/vendor) or will that be affected?, who should be able to deliver added services or solutions over a period of time?, how can the roles between the project department and the operation department be affected?, what are the implications on the contracting process and contractual conditions?,* etc. Notably, there can also be other business conditions that affect both the underlying processes and decisions, for instance;

- The oil company wants to keep full control of the major operational costs and thus overall cash flow process, after making major investments to build and commission (a new offshore oil installation typically cost between NOK 10 billion and NOK 100 billion).
- There may be different tax rules for CAPEX (investment costs) and OPEX (operational cost) in different countries. Hence, introducing additional operational elements in the purchasing contracts may influence on the reimbursement of tax money, w.r.t. enforced laws and regulations.
- Even if an upstream oil company prefers to delegate more maintenance responsibility to main contractors/vendors or other third parties, there are strict regimes related to safety responsibility. As license owner and operator, the upstream oil company is always the ultimate accountable for the safety of the installation.

In spite of clear division of core responsibilities between an oil company and contractors following formal industrial practice, some new platforms have been developed and delivered requiring major upgrades prior to full scale production authorization. Some projects have lead to multimillion-dollar modification projects after delivery in order to make the platform "useable" within a normative operational and regulatory frame. As mentioned, the optimal role split will probably be dependent on what kind of physical products are in question, the conditions for functional assurance and maintenance programs, competency availability, as well as terms of warranties and other regulatory demands. In an effort to adapt to the changing demands from oil companies as well as to position in changing markets, various equipment and service providers have begun to develop flexible business options creating more space for a strategic interplay between a main manufacturer/vendor and the upstream oil company. Observably, this has gradually begun to influence the traditional mind-sets related to conventional oil business ecosystem, and thus creating positive signs towards new developments. The current efforts by supplier and vendor market can take many diverse forms, for instance; from *delivery of a product*, to *delivery of a product together with condition monitoring package*, *delivery of a product with an extended remote surveillance service together with a dedicated data platform*, and *delivery of a complete solution inclusive of functional assurance,* etc. Interestingly, some other form of alternative solutions also appear to emerge gradually as 'hybrid solutions' challenging the conventional business eco-system. For instance, an oil company may retain x% of the purchase price of a major equipment or a technical system, for instance for two years from purchase and commissioning date, until all performance targets are met as expected and agreed. Such solutions create opportunities for other third parties that for instance, can be contracted and thus become responsible through a mutual agreement for maintenance, testing, and inspection services. One of the major bottlenecks in this regard is the contract regime that has matured over the years towards a more or less a non-negotiable practice.

4 Conclusions

In this paper, we have elaborated on the transition process in the oil & gas sector from traditional to connected business ecosystems. This to a large degree is also affected by digitalization interests and subsequent challenges related to IT and OT aspects, as much as economic demands. Based on the observations of the E&P industry, particularly changes in the last decade and emerging challenges, it is clear that the best alternative (s) are not yet ready and that there is an array of complex issues with no easy answers. However, there are some clear signs and trends, which indicate that the oil and gas industry will be exposed to substantial changes and subsequently the traditional business model will gradually be transformed towards a different business ecosystem in the years to come. During the most challenging period that the industry has been going through, the E&P conditions clearly demanded sharper focus on cost-effectiveness and simplicity. This has been shaking the whole industry for some time gradually leading the path towards structural changes and new eco-systems. It can be anticipated that the changes will be accelerated during forthcoming years for better, breaking conventional practices and creating a more robust business eco-system to survive and prosper under tough industrial conditions.

References

Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N.V.: Digital business strategy: toward a next generation of insights. MIS Q. **37**(2), 471–482 (2013)

Libert, B., Beck, M., Wind, J.: In a networked world, it's time for leaders to follow. Leader Leader **83**, 41–46 (2017)

Liyanage, J.P.: Smart engineering assets through strategic integration: seeing beyond the convention. In: Van der Lei, T., et al. (eds.) Asset Management – The State of the Art in Europe from a Life-Cycle Perspective. Springer (2012)

Liyanage, J.P.: Coping with dynamic change: collaborative business interfacing for SMEs under integrated eOperations. In: Varajao, J.E. (ed.) Small and Medium Enterprises: Concepts, Methodologies, Tools, and Applications. IGI Global (2013)

Makinen, S.J., Dedehayir, O.: Business ecosystem evolution and strategic considerations: a literature review. In: 2012 18th International ICE Conference on Engineering, Technology and Innovation (2012)

Norwegian Oil and Gas Association: 104 - Norwegian Oil and Gas recommended guidelines on information security baseline requirements for process control, safety and support ICT systems. New revision 05122016 (2016). www.norskoljeoggass.no

Shuen, A., Feiler, P.F., Teece, D.J.: Dynamic capabilities in the upstream oil and gas sector: managing next generation competition. Energy Strategy Rev. **3**(C), 5–13 (2014)

Sustainvalue Deliverable D1.1: Sustainability gaps and stakeholder requirements (2011). http://www.sustainvalue.eu/publications.htm

Sustainvalue Deliverable D2.1: State-of-practice in business modelling and value-networks, emphasising potential future models that could deliver sustainable value (2012). http://www.sustainvalue.eu/publications.htm

Teece, D.J.: Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. Strateg. Manag. J. **28**(13), 1319–1350 (2007)

Venkatraman, N.: IT-enabled business transformation: from automation to business scope redefinition. Sloan Manag. Rev. **35**(2), 73 (1994)

Venkatraman, V.: The digital matrix: new rules for business transformation through technology. LifeTree Media Ltd. (2017)

A Tool to Assess Learning Processes Based on the Cooperation Principle

Gabriela Bjørnsen[1], Ove Njå[2(✉)], and Geir Sverre Braut[3]

[1] Rogaland Fire and Rescue IKS, Sandnes, Norway
gabriela.bjornsen@rogbr.no
[2] University of Stavanger, Stavanger, Norway
ove.njaa@uis.no
[3] Stavanger University Hospital and University of Stavanger, Stavanger, Norway
geir.sverre.braut@sus.no

Abstract. A strategy for cooperation in emergency management has been developed and politically agreed upon by the Rogaland County Council, Norway. The region comprised by the strategy consists of many different actors within societal safety and emergency management. The strategy aims at strengthening the existing cooperation, establishing professional centres and further developing competencies in their emergency response efforts within the region. The region has more than twenty road tunnels either in the planning phase, under construction or in operation. The emergency services have established a new organisation of their cooperation to ensure coordination, learning and supervision. This relates both to exercises and real event operations. An important tool in this respect is a recently developed handbook for cooperative exercises. The book is used in planning, execution and follow-up of all cooperation exercises. In this paper we present our newly developed evaluation model for following up the cooperation exercise guidelines, with special attention to events in road tunnels. We employ a learning model that extends the notion of learning from observed changes to also include confirmation and comprehension of cooperation activities.

1 Introduction

Four principles constitute the basis of the Norwegian emergency preparedness structure - *responsibility, similarity, proximity* and *cooperation.* The cooperation principle, (re) introduced formally as an overarching, cross-sectoral planning criterion in 2012, implies that authorities, voluntary, private and official actors are individual responsible for establishing appropriate interaction and coordination with relevant parties in all activities regarding prevention, emergency preparedness and crises management. In this paper, *cooperation* is used as a translation of the Norwegian term "samvirke". Our definition of cooperation in emergency management includes both coordination and interaction between different independent actors, vertically and horizontally at all organisational levels.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 87–95, 2020.
https://doi.org/10.1007/978-3-030-48021-9_10

Interagency coordination is necessary for successful implementation of critical decisions and prevention of overlap, conflicts and miscommunication (Boin et al. 2005). However, such coordination is challenging to obtain. Technical and cultural communication problems can undermine the horizontal cooperation during emergency response (Boin et al. 2005), different terminology and interagency conflicts can add further pressure on the emergency management (Paton and Flin 1999). In the review of the incidents in Oslo and at Utøya in 2011, the “July 22 commission” points at the lack of ability to coordinate and interact, and to learn from exercises, as two important factors explaining the unfortunate performance of the emergency response (Gjørv 2012).

A recently developed handbook in Rogaland for cooperative exercises constitutes the basis for the planning, execution and follow-up. It is a tool aimed for improving practices in and across emergency services. Previous experiences showed limitations in learning from full-scale cooperation exercises due to the participants’ lack of prerequisites (Vik et al. 2014). The handbook was published in August 2014, but the principles in the handbook have been applied in planning and execution of cooperation exercises since the autumn 2013. The correlations between the handbook’s exercise concepts and how learning is achieved needs to be further studied. To succeed with the intentions of establishing better cooperation in emergency prevention and management the actors must constantly challenge the learning principles, learning as a phenomenon and the established practice for training activities.

It is necessary to assess and evaluate the regime for follow-up initiated learning processes. This article presents an evaluation model of cooperation exercises based on the handbook’s guidelines with special attention to tunnel fires. The tool is a first edition of parameters that contribute to the learning processes from stimuli, such as exercises, training activities or real events, has been subjected to learners (emergency services and tunnel management actors) until effects are observed. Our evaluation tool is based on a combination of learning theories and empirical data.

2 Theoretical Framework for the Evaluation Tool

Exercises are common means in building experiences and competencies in interagency cooperation (Lonka and Wybo 2005). In addition to uncover limitations in emergency plans, cooperation exercises contribute to establishing networks and personal relationships between emergency responders from different organisations (Kettl 2003). Training and exercises are also important tools for emergency responders to acquire and learn how to use necessary knowledge and skills (Lonka and Wybo 2005; Sinclair et al. 2012; Sommer et al. 2013).

Both the emergency response systems and the context in which they operate can be described as complex, an example is tunnel fire responses (Njå and Svela 2018). There is a need for a holistic view on systems and accident factors, because events, actions and behaviour of the different system components can only be understood by considering their *“role and interaction within the system as a whole”* (Leveson 2011). The control of safety within a system involves many levels of actors, ranging from the Government at the top-level to the operators at the bottom. Each level applies constraints on the level beneath, through legislation, policies, rules, routines, work instructions etc., thus forming the boundaries for the system's practice and performance (Rasmussen 1997; Svedung and Rasmussen 2000; Leveson 2011). Such hierarchy of control must be based on adaptive feedback mechanisms and communication to ensure that *"the information needed for decision making is available to the right people at the right time"* (Leveson 2011). Abrahamsson et al. (2010) argues that a holistic system approach is suitable to deal with the complexity of the situations and systems involved in emergency response.

2.1 Learning and Experience in Emergency Response Work

The concept of learning is subjected to different definitions and perspectives (Braut and Njå 2009). The individual cognitive approach views learning as acquisition, using individual factors to explain the development of competence. While the sociocultural approach focuses on learning as participation, and explains the development of competence through contextual factors (Sommer et al. 2013; Sommer 2015). It can be argued that the two approaches complement each other (Sommer 2015; Sfard 1998; Sommer et al. 2013), and in order to understand how emergency workers develop their skills and knowledge, these two approaches must be combined. Sommer et al. (2013) have developed a model for learning in emergency response work based on such a combination, cf. Fig. 1.

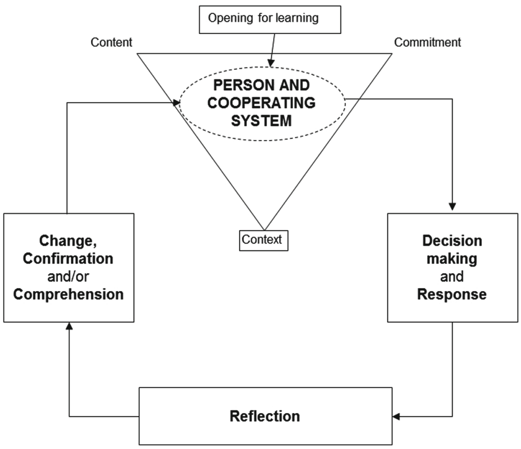


Fig. 1. Learning model adapted from Sommer et al. (2013)

This model sees learning as a continuous process. The starting point for understanding learning is the individual (the person) placed within the contextual elements of *content, context* and *commitment* (Sommer et al. 2013; Sommer 2015). The practical and theoretical content of the learning activities, such as skills, behaviour, how to interpret situations or the use of equipment, must be relevant in order to improve performances. Social climate, relationships, trust and openness are conditions affecting the learning environment, which constitutes the context. The individual's commitment to the learning activities also strongly influences what and how much learning that occur (Sommer et al. 2013; Sommer 2015). In this theory it is believed that individual learning in concert creates the organisational and cross-organisational learning, even though it is not an equation of simple aggregation of the individual learning.

During real emergency situations or exercises the emergency workers must be able to consider relevant situational cues in their decision-making. The result of these decisions are the individuals' behaviour and response, which in the end form the outcome of the emergency or exercise situation. Through subsequent reflection on their performance, the individuals can learn from their experiences, and this will influence the individual performance in following situations (Sommer et al. 2013; Sommer 2015).

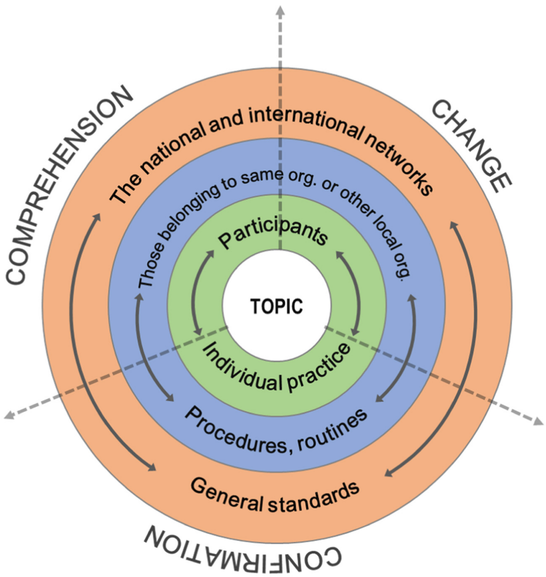
The outcome of learning can be categorized as *change, confirmation* and/or *comprehension*. Learning results have traditionally been expressed as changes in structure, behaviour, cognition, processes or organisations, but the model for learning in emergency response work also includes confirmation and comprehension as potential results from learning (Sommer et al. 2013; Sommer 2015; Braut and Njå 2009). Confirmation is some kind of positive reinforcement verifying that the emergency workers' normal practices, tools and existing skills are working very well. Learning as comprehension occurs when the emergency workers gains a deeper understanding of existing practices, tools and behaviour. A more comprehensive understanding of the mechanisms working in different emergency situations in and across organisations and how different practices and behaviour can provide possibilities or limitations, enable emergency workers better prepared when facing new situations (Sommer et al. 2013).

2.2 The Evaluation Model Directed at Cooperation Exercises

Emergency workers need to meet emergency situations in a functional way. They must be able to know which tasks to implement in which situations, be familiar with own tasks and how to perform them satisfactorily, and assess the results of their own behaviour (Njå and Sommer 2010). Because exercise situations never will be identical to real situations, the objective of exercises are to develop competencies and knowledge that can be generalised to similar real-life situations.

A method for evaluating the quality of a training and exercise program must consider the interaction between the situation characteristics, the individual's competencies and the probability of personnel showing functional behaviour in the situation. An evaluation must analyse the type of situation (scenario) used in the exercise and if it matches "real life", it must define the personnel's level of competence and the behavioural objectives of the exercise. The chosen evaluation method can start with either situation, person or behaviour when assessing if the exercise is a suitable way of establishing required connection between the three (Njå and Sommer 2010).

The three layers in the circular model in Fig. 2, represent the different levels in the emergency response system hierarchy, starting with the individuals in the inner circle. In order to evaluate if and how learning has taken place through exercises, we claim that expressions of change, confirmation and comprehension must be identified, qualitatively or quantitatively on different levels in the participating organisations, from the individual level up to the level where regulations and general operative standards are made, cf. Fig. 2.



	CONFIRMATION	COMREHENSION	CHANGE
Individual practice			
Procedures/routines			
General standards			

Fig. 2. Model for studying and evaluating cooperation exercises

2.3 The Evaluation Tool

The control loop that forms the basis of the handbook for cooperative exercises is consistent to the feedback mechanisms described above (Leveson 2011). Each step in the exercise process shall be subjected to evaluation and subsequent feedback in order to make necessary adjustments before initiating the next step in the process, in which cooperation activities are specifically focussed.

Content, Context and Commitment

Content, context and commitment are the elements that constitutes a person's learning of cooperation tasks. Thus the concerted development of skills and knowledge depend on a combined approach to learning and how the exercise is designed in order to make the participants receptive to learning. According to the handbook for cooperative exercises the evaluation shall clarify experiences and weaknesses for further action, and for use in developing and improving both the cooperation between the organisations and the individual agencies (Vik et al. 2014). This process is vital in order to obtain learning as confirmation, comprehension and change at all levels in the emergency response system. There need to be explicit cooperative goals for the exercises, which will be considered through the assessments;

Content:
- Identification of boundary objects
- Phenomena involved
- Flexible vs. standardised cooperative behaviour
- Motor vs. cognitive collaboration behaviour
- Responsibility and decision making
- Communicative challenges

Context:
- Physical requirements of joint forces
- Mental requirements of joint forces
- Emotional requirements of joint forces
- Contact with physical energy
- Training arenas

Commitment:
- Degree of individual involvement in cooperating tasks
- Instructors' competencies
- Motivational aspects
- Participant competencies and preparations
- Socio-cultural factors including boundary awareness

Decision Making and Response

Decision-making and response corresponds to individuals' performance in collaboration with others in the training situation. Individuals' behaviour and response is thus a result of the decisions they make in specific contexts, which consequently form the outcome of the emergency situation. Important features of the tool are as follows:

Decision making:
- Information processing (use of available information, search for additional information etc.)
- Situational awareness
- Ability to recognise relevant situational cues
- Mental simulation
- Communication and coordination
- Development of tactics and measures

Response:
- Choice of action
- Allocation and use of available resources
- Performance of tasks

Reflection

Reflection is a vital prerequisite for learning, thus exercises must facilitate such activities. Collective debriefings are necessary in order to discuss and exchange experiences across different organisations. Discussions directly after the exercises are perceived as valuable, and concluding evaluation and collective debriefing has been requested by exercise participants during international studies (Berlin and Carlström 2014; Berlin and Carlström 2015; Andersson et al. 2014).

The time factor is paramount. Debriefing immediately after the exercise is important to share observations about the event and discuss the reception narrative and which individual and collective features of the exercise content, context and commitment that were interesting for learning. Participants and organisations must be encouraged to carry out step two of the reflection, which is digging deeper into the experiences, much in line with root cause assessments. By providing process feedback the participants are given information to make them understand what led to a particular outcome. Critical cues and judgements about actions should be the focus of the emergency workers reflection (Sommer et al. 2013). Informal discussions and assessments are as important as a formal gathering some days after. Thus the study of cooperation exercises must include an evaluation of the debriefings and their contribution to reflection, which we present as questions:

- Is the narrative agreed upon or is there opposing views?
- Did the collective debriefing contain elements of Gibbs' reflection circle (1988); descriptions; feelings; evaluation; analysis; conclusion; action plan?
- Did the participants focus on process rather than outcome?
- How was uncertainties presented and discussed?
- How is trust between individuals and organisations reflected?
- Which boundary objects were important, and did the debriefing concentrate on being aware of boundaries of future cooperation?
- If there were alternative interpretations – did the debriefing include discussing other strategies?
- Did the debriefings clarify the cooperative abilities, confirming good practice?
- Did the debriefings identify areas in which the cooperation exercise provided new knowledge and need for changes?

Change, Confirmation and Comprehension

Identifying expressions of change, confirmation and comprehension from cooperation exercises, dependant on trust, understandings of responsibilities, phenomena involved, etc., have been scarcely studied in the research literature. Our model in Fig. 2 tries to grasp the dynamics of the individuals at the core and how general standards at the national and international networks are influenced. Our approach is explorative and it will be developed during the research activities of exercises and real event assessments started up in Rogaland.

In order to properly understand the concepts of change, confirmation and comprehension, our research is directed at describing inferences, activities developed, measures and expressions of individual and collective reflections being internalized in the individuals, the organisations, and across organisations and networks.

Change in:

- Response actions
- Plans, procedures

- Situation assessment
- Practical handling
- Participants experiences from the exercises

Confirmation of relevant:

- Cooperation and teamwork
- Self-evaluation
- Knowledge acquired
- Working across organizational boundaries
- Joint response work
- Joint situational awareness

Comprehension in the perspective of:

- Physics in the situation development
- Constraints established
- Interaction between services
- Variations in human (victim) behavior
- Responses to toxins from various smoke compositions

A preliminary summary of this work with the new tool is that there is an urgent need for a systematic credible approach to learning. The tool shall no undergo testing, first and foremost in exercise and training situation, and furthermore as a tool to assess learning from investigations of real events. The success of this approach require that involved actors understand the assessment and acknowledge the assessments and results produced as meaningful input to the services' works.

3 Conclusion

Cooperation in emergency services is necessary in order to achieve a successful emergency response work. It is essential that emergency services has the ability to interact and coordinate resources, as well as learn from exercises. Exercises are significant tools in strengthening skills and improving the cooperative activities between the different agencies involved in crisis management. It is however necessary to study how and if learning takes place as a result of cooperation exercises. The exercise handbook developed in Rogaland is a very good initiative in order to take the cooperative exercises one step further. The intentions and cooperation behind this handbook and its guidelines needs to be maintained and developed.

A main objective with the exercise handbook is to facilitate learning at both individual level and within and across the organisations, as well as identify knowledge gaps and areas where competence needs to be increased. It is therefore needed to evaluate how the handbook provides additional knowledge and competence in the emergency services prevention and management work. Based on the evaluation model presented in this article we wish to examine the correlation between the handbook's exercise concepts and if and how learning is achieved. Our learning evaluation model will be tested in various activities the next year.

References

Abrahamsson, M., Hassel, H., Tehler, H.: Towards a system-oriented framework for analysing and evaluating emergency response. J. Contingencies Crisis Manag. **18**, 14–25 (2010)

Andersson, A., Carlstrom, E.D., Ahgren, B., Berlin, J.M.: Managing boundaries at the accident scene – a qualitative study of collaboration exercises. Int. J. Emerg. Serv. **3**, 77–94 (2014)

Berlin, J.M., Carlström, E.D.: Collaboration exercises—the lack of collaborative benefits. Int. J. Disaster Risk Sci. **5**, 192–205 (2014)

Berlin, J.M., Carlström, E.D.: Collaboration exercises: what do they contribute? J. Contingencies Crisis Manag. **23**, 11–23 (2015)

Boin, A., Hart, P., Stern, E., Sundelius, B.: The Politics of Crisis Management: Public Leadership Under Pressure. Cambridge University Press, Cambridge (2005)

Braut, G., Njå, O.: Learning from accidents (incidents). Theoretical perspectives on investigation reports as educational tools. In: Reliability, Risk and Safety. Theory and Applications, pp. 9–16 (2009)

Gibbs, G.: Learning by doing: a guide to teaching and learning methods (1988)

Gjørv, A.B.: Rapport fra 22. Juli-kommisjonen [22 July commission report]. Norges Offentlige utredninger, Oslo (2012)

Kettl, D.F.: Contingent coordination practical and theoretical puzzles for homeland security. Am. Rev. Public Adm. **33**, 253–277 (2003)

Leveson, N.: Engineering a Safer World: Systems Thinking Applied to Safety. MIT Press, Cambridge (2011)

Lonka, H., Wybo, J.-L.: Sharing of experiences: a method to improve usefulness of emergency exercises. Int. J. Emerg. Manag. **2**, 189–202 (2005)

Njå, O., Sommer, M.: Evaluering av øvelsesopplegg i forbindelse med snøskredredning i Rogaland 2010: perspektiv på ledelse og skredredning [Evaluation of exercise arrangements in avalanche rescue in Rogaland 2010: perspectives on management and avalanche rescue]. Universitetet i Stavanger, Stavanger (2010)

Njå, O., Svela, M.: A review of the competencies in tunnel fire response seen from the first responders' perspective. Fire Saf. J. **97**, 137–145 (2018)

Paton, D., Flin, R.: Disaster stress: an emergency management perspective. Disaster Prev. Manag. Int. J. **8**, 261–267 (1999)

Rasmussen, J.: Risk management in a dynamic society: a modelling problem. Saf. Sci. **27**, 183–213 (1997)

Sfard, A.: On two metaphors for learning and the dangers of choosing just one. Educ. Res. **27**, 4–13 (1998)

Sinclair, H., Doyle, E.E., Johnston, D.M., Paton, D.: Assessing emergency management training and exercises. Disaster Prev. Manag. Int. J. **21**, 507–521 (2012)

Sommer, M.: Learning in emergency response work. University of Stavanger, Stavanger (2015)

Sommer, M., Braut, G.S., Njå, O.: A model for learning in emergency response work. Int. J. Emerg. Manag. **9**, 151–169 (2013)

Svedung, I., Rasmussen, J.: Proactive risk management in a dynamic society. Swedish Rescue Services Agency, Karlstad (2000)

Vik, H., Berg, H.O., Lilleeng, B.: Håndbok i øvelses planlegging [Handbook for Exercise Planning], Stavanger (2014)

Digital Twins, a New Step for Long Term Operation of Nuclear Power Plants

Christophe Varé(✉) and Patrick Morilhat

EDF R&D, 6 quai Watier, 78401 Chatou Cedex, France
{christophe.vare,patrick.morilhat}@edf.fr

Abstract. Digitalization is one of the key technology to improve safety and performance of Nuclear Power Plants (NPP). EDF has initiated R&D in this field for more than 10 years and some tools are now commonly used by EDF operators. They cover outage management (use of virtual reality for outage preparation, 3D visualization of plant maintenance and upgrade…), human performance in operation (plant field workers mobile technologies, augmented reality to improve situation awareness, advanced training…), advanced plant control automation and digital architecture (fab lab for design and fast checking of future concept of operation…), as well as on-line equipment and process monitoring. Now the new challenges are to develop "digital twins" which are digital models fed by on-line data from operation and to benefit from data analytics science which made great progress during the last years. EDF has engaged new programs in these fields that will support long term operation of most critical nuclear equipment. A first development, dedicated to a 1:3 scale mock-up of a NPP containment is now operational. It gives an idea of the technical challenges to handle before deploying such technology for EDF fleet.

1 Introduction

Digitalization is said to be one of the new frontier of the nuclear industry. In reality, digitalization was introduced in the nuclear field a long time ago through various developments. As a well-known example, scientific modelling has been used for many years by engineering division to improve the safety and the performance of Nuclear Power Plants (NPP).

Other developments have been completed and implemented in EDF NPP operation processes. In this paper, we will first introduce some of them. They cover various operational and engineering fields such as outage management, human performance in operation, advanced plant control automation and digital architecture, as well as on-line equipment and process monitoring. Despite these successes, some challenges are still to be overcome, giving new focus for R&D activities.

Digital twin may be the new age for the digitalization of the nuclear industry even if there is actually a clear connection between the concept of digital twin and developments that have been achieved during the past years in instrumentation, data analysis, numerical simulation… To evaluate the potential of this concept, EDF has undertaken the realization of the digital twin of a 1:3 scale mock-up of a 1300 MW nuclear reactor containment. This first work, which is described in this paper has shown the technical

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 96–103, 2020.
https://doi.org/10.1007/978-3-030-48021-9_11

issues that have to be solved. It also showed how much this new approach can be virtuous in a complex engineering process.

2 First Steps of EDF Nuclear Activities Digitalization

Digitalization of nuclear activities is not new at EDF. The first numerical models for modelling of mechanical phenomenon or process operation were developed by EDF R&D more than 25 years ago. It is worth mentioning that in the mid 90's, EDF with French industrial partners have digitally designed and built two 1450 MW nuclear plants with fully digitalized control rooms, which was a premiere at that time (Meng 2018).

Since then, other areas have been digitalized for numerical computations: thermal-hydraulics (computational fluid dynamics, CFD), core calculations, electromagnetism… In addition, considerable progress has been made in the field of numerical simulation with the development of ever more accurate models and the provision of ever greater computing power.

Current digital projects at EDF address the full life cycle of the plant: the design phase with the extensive use of simulation tools and rapid prototyping interfaces, operation and maintenance activities with the development of solutions based upon 3D virtual and augmented reality for field operators (Dupin 2016), artificial intelligence for enhanced predictive maintenance, and robotics for decommissioning works.

The relevance of these models for design activities and the understanding of complex physical phenomena has been demonstrated for long. However, many of these tools are too complex and theoretical to be used in nuclear power plant operation process. Last IT developments make it possible to overcome this obstacle. Following, we show with some examples how these recent developments have been used in operational nuclear activities.

2.1 Virtual Reality Applied to Nuclear Power Plant

EDF R&D has developed a solution for a 3D virtual visit of the reactor building (Hullo 2015). Based upon 360° photographs, laser scans and reconciled CAD data and drawings, the tool provides a complete as-built view of the reactor and allows maintenance preparation staff to prepare safely and correctly the outage prior to the opening of the reactor building. The tool has been used for 2 years in several plants and is currently been deployed across the entire EDF's fleet. Thanks to a gaming technology, the maintenance operator is able to find rapidly the location of the equipment for which the maintenance work is planned, identify specific requirements such as a need for scaffolding, with a precision of about 2 cm. A deep learning algorithm helps the operator to find quickly and automatically one specific equipment with only its identification tag among the thousands of 3D photos included in the solution. This enables a saving of significant work hours on the outage critical path (Fig. 1).

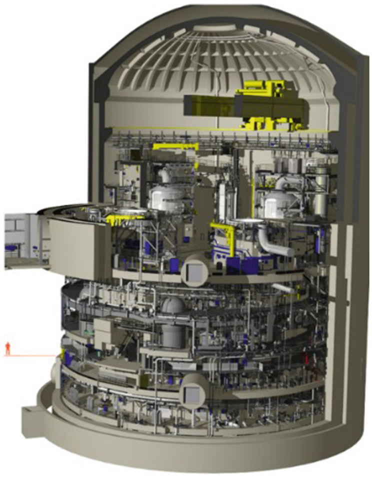


Fig. 1. 3D virtual reactor building

2.2 No Data Processing Without Data

The data analytics lifecycle follows a step by step process: data acquisition and extraction, data cleaning and validation, data analysis and visualisation, feedback to the process when needed (data driven automation). Garbage in, garbage out: in this data lifecycle, the quality of data is key for a further use with data science techniques. As many other industries, we estimate that about 70% of the efforts are spent on data extraction and cleaning, by far more than the data analysis phase itself. Data may be incomplete, corrupted, inconsistent, etc. The need for a methodology to speed-up this phase is prominent and rarely underlined by the most enthusiastic advocates of data analytics.

Compared to other industries, nuclear data is not "big data". The total volume of EDF plant digital process data amounts to some TBs, the equivalent of Twitter microposts per year. In terms of sensors, a typical plant has about 8000 digital sensors, to be compared to 400 000 for a last generation Airbus aircraft. In fact, the problem to face in nuclear plant is less the volume, even though not negligible, than the diversity of data. Developing a sound predictive maintenance solution leads to mix together process data (numbers), results of non-destructive examinations (images) and maintenance reports (text or even written documents). The information used for diagnosis and prognostic has to be extracted from various types of sources and formats and reconcilied to provide the relevant information (Baraldi 2015).

2.3 Progress of Data Science

When it comes to the modelling of systems or components, two different paths are equally possible: on the one hand, a model driven approach, on the other hand a data driven approach. We do think that these approaches are not in opposition, but they have to be associated to improve the value of analytics.

The model driven approach, also known as model-centric, consists in modelling a system or a component with numerical analysis tools (e.g. finite elements, finite differences or finite volume) when we know what are the phenomena and they can be described explicitly by physical conservation laws, generally under the form of differential equations. It is the world of "classical" engineering sciences. This approach requires a deep knowledge of the physical phenomena (thermal-hydraulics, mechanics, electromagnetism…), of the design of the components geometry and boundary conditions.

The data driven approach, also known as data-centric, is more recent and lifted by three factors: the increase in the amount of data, advances in data analytics algorithms and significant progress in high performance computing facilities. Data analytics techniques extract information from both structured (time series, Excel or SysmL files) and non-structured data (raw text, images, video…). This approach is well adapted to situation where the physical knowledge is sparse or missing, or the phenomena are too complex, but there is a large amount of data, possibly of various nature, related to the phenomena to be described.

Going for digital twins relies on a sound mix between this two approaches. An example of innovation brought by the alliance of model driven and data driven approach is the development of a new generation of predictive maintenance monitoring tools. The first generation of monitoring tools has been deployed since beginning 2000's across the EDF fleet, mainly based upon model-driven tools using process data. In addition to these tools, EDF is developing a "layer" of data analytics solutions, mainly aimed at crossing the results of numerical models with existing data non explored so far, such as maintenance reports.

3 Toward Digital Twins

As it has been presented above, physical modelling computation, virtual reality, instrumentation and data analysis have shown great progress in the past years. Thus, the integration of all these disciplines in one unique tool seems now achievable. This gives an idea of the digital twin concept: on one side the real equipment (a pump, a thermal-hydraulic loop, a complete reactor…), on the other side its model (model or data-centric) and between them some ways to exchange information (online monitoring, modelling results…) feeding both the model and the real equipment (operation or maintenance decisions). To evaluate the potential of this concept, and the technical challenges to overcome, EDF has undertaken the realization of the digital twin of a 1:3 scale mock-up of a 1300 MW nuclear reactor containment.

Following, this mock-up named "Vercors" is first introduced. Then we present data that are measured and the finite element model which has been done to simulate its behavior. The digital twin of Vercors mock-up is then described. The main conclusions of this work are finally discussed.

3.1 Vercors Mock-Up

"Vercors mock-up" (Fig. 2) is a 1:3 scale mock-up of a 1300 MW nuclear reactor containment, which has been built in EDF Lab Les Renardières (France) to observe and analyze ageing phenomena and their effects on such concrete structures. In particular, it aims at studying the evolution of tightness with time of nuclear double-walled concrete reactor containment buildings (Masson 2015).

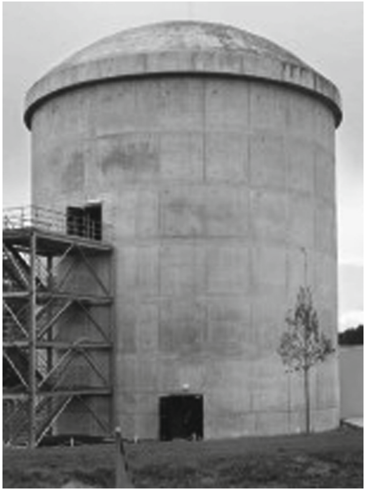


Fig. 2. EDF Vercors 1:3 scale mock-up for NPP containment ageing study

3.2 Data and Modelling

What guided the development of the digital twin, at the beginning of Vercors project, was the mass of data to be handled and the will to make them available to all concerned actors. We needed to capitalize them in a tool that is immediately operational for engineers such as it can be used on the whole duration of Vercors Program. The data we are considering are construction as well as measurement data. These latters come from tests and on-line measurements as illustrated below.

To characterize the behaviour of the concrete of Vercors mock-up, more than 1000 tests were carried out in EDF and university laboratories. The aim of these tests is to get data on the creep and the shrinkage of concrete to set the constitutive material laws that are then used to make prediction on the containment overall behaviour. When we built Vercors mock-up, we also installed lots of sensors, traditional ones and innovative ones. 2 km of optic fibers were thus implemented in Vercors mock-up. They can gather more than 2000 measurements per hour. Their processing gives on-line deformation and temperature maps of Vercors that can be visualized and that can be also used to feed models or to validate the modelling approach.

A finite element model of Vercos (see Fig. 3) was developed, using advanced EDF-developed digital tools such as Salome Méca. It enables to consider the data measured locally and to extrapolate it to the whole spatial domain, rebuilding fields, like temperature or mechanical fields and thus giving a precise image of the reality of the structure, like a snapshot. It can also be used to predict the evolution of the behaviour of Vercors mockup, and in particular to predict its tightness over time.

Fig. 3. Finite element model of Vercors mock-up

3.3 Digital Twin

All technological bricks are thus ready to build the digital twin of Vercors. To define the main features of Vercors digital twin, a hackathon bringing together all EDF engineer working on the Vercors Project was held. Through this event, a modular approach was chosen and it was decided to focus the first work on extraction and visualisation of data. As a result, the digital twin takes the form of a set of software designed to communicate with each other as interoperable components (see Fig. 4 below). They are used to capture data, to check their consistency and give visual representations that are meaningful to Vercors engineers. They also integrate the finite element model of Vercors enabling computations of interest (Mathieu 2018).

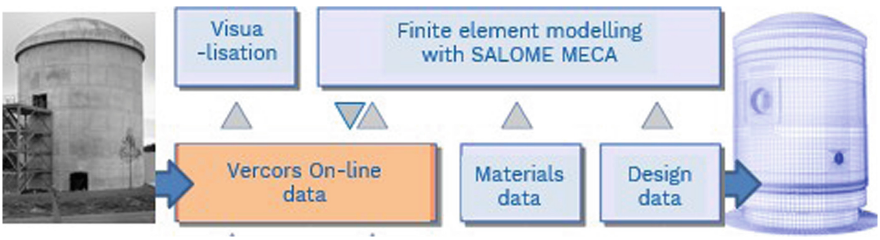


Fig. 4. Vercors modular digital twin

The Vercors digital twins is starting to prove its worth. Its visualisation tool has already been used by the Technical Direction of EDF to facilitate the choice of maintenance of certain nuclear power plants. In a longer run, it is planned to create a digital twin of each containment building for which maintenance is expected.

4 Conclusions

In maintenance and operation at EDF Nuclear, the digital transformation is well on tracks. More than 800 solutions have been developed and deployed: electronic work-packages for field workers, 3D visualisation tools for outage preparation, IOT devices, etc. For engineering staff, a Plant Life Management Infrastructure allows the sharing of reference data and documents among EDF and its main suppliers.

Digital twin is a new step of the digital transformation of EDF nuclear activities. It can rely on the know-how that has been acquired by EDF engineers for long. A first use-case has been developed by EDF R&D using a 1:3 scale mock-up of a 1300 MW nuclear reactor containment. This first experience shows that digital twins help different teams (R&D, measurement and engineering) speaking a common language and working efficiently all together. In addition, digital twins are seen to be useful to transmit engineering knowledge beyond generations, a major challenge of long term operation of nuclear power plants.

However, the digital transformation raises a set of questions that are not well answered yet. Data governance has to be made clear: who owns the data, how to access the data repository or data lake, what are the responsibility regarding the data quality and the conclusions that data analytics may lead to. The business model, and behind the value of data analytics, is not that obvious, as opposed to what is said by many. Data is on the cost side, information is on the value side. Where does the cost go to is quite clear (generally the plant operator who has to deploy and maintain a data acquisition and management system). Where does the value go is not that clear, because it could be shared by numerous players: the operator, the suppliers, consultancy companies selling data analytics services, etc… In any case, a strong use case approach has to be set up to assess, on a case by case basis, what and where is the value of any digital solution to be implemented.

References

Baraldi, P., Di Maio, F., Rigamonti, M., Zio, E., Seraoui, R.: Unsupervised clustering of vibration signals for identifying anomalous conditions in a nuclear turbine. J. Intell. Fuzzy Syst. **28**(4), 1723–1731 (2015). https://doi.org/10.3233/IFS-141459

Dupin, L.: Digital technology strengthens EDF's nuclear safety (in French). Usine Nouvelle (3479–3480), 8–10 (2016)

Hullo, J.-F., Thibault, G., Boucheny, C.: Advances in multi-sensor scanning and visualization of complex plants: the utmost case of a reactor building. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **XL-5/W4** (2015). https://doi.org/10.5194/isprsarchives-xl-5-w4-163-2015

Masson, B., Galenne, E., Oukhemanou, E., Aubry, C., Laou-Sio-Hoi, G.: Vercors: a 1/3 scaled mockup and an ambitious research program to better understand the different mechanisms of leakage and aging. In: Fontevraud 8: Conference on Contribution of Materials Investigations and Operating Experience to LWRs' Safety, Performance and Reliability, Avignon, France, 15–18 September 2014 (2015)

Mathieu, J.-P., Charpin, L., Sémété, P., Toulemonde, C., Boulant, G., Haelewyn, J., Hamon, F., Michel-Ponnelle, S., Hénault, J.-M., Taillade, F.: Temperature and humidity-driven ageing of the VeRCoRs mock-up. In: Computational Modelling of Concrete Structures, pp. 215–224 (2018)

Meng, Q.J., Liu, Z.B., Sun, Y.B.: Overview of development of NPP advanced main control room console. In: Xu, Y., Gao, F., Chen, W., Liu, Z., Gu, P. (eds.) Nuclear Power Plants: Innovative Technologies for Instrumentation and Control Systems, SICPNPP 2017. Lecture Notes in Electrical Engineering, vol. 455. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-7416-5_24

Defining Success in Air Force Infrastructure Asset Management Through Use of the Delphi Technique

Brendan Maestas(✉), John Elshaw, Andrew Hoisington, and Donald Ohlemacher

AF Institute of Technology, Wright-Patterson AFB, Dayton, OH, USA
brendan.maestas@afit.edu

Abstract. Asset Management has a history of policy mandates within the US Government dating back to 1990's. In order to accomplish these many directives, the Air Force Civil Engineer community has adopted a mindset and framework commonly referred to as Asset Management. Despite numerous references and guidance to establish Asset Management principles, the AF has not yet developed a clear and concise way to define or measure overarching success in Asset Management. This research effort focuses on closing the knowledge gap between issued policy and implementation. Through interviews from Subject Matter Experts at various levels of the AF Civil Engineering structure, this research identifies: key elements that constitute and promote success, internal success identifiers, and currently established measures for success. Using this information and recommendations from the AF SMEs, novel suggestions are presented for measuring and incentivizing Asset Management success within an organization.

1 Introduction

Asset management is a relatively new field of study originating out of maintenance management of the oil and gas industry in the 1980's (Woodhouse 2003). Asset management has been gaining popularity in industry as companies seek to minimize lifecycle costs and maximize the operational longevity of their built infrastructure and equipment. The United States Air Force (USAF) first entered into the realm of asset management through Executive Order EO 13327 (Bush 2004). EO 13327 stated that "It is the policy of the United States to promote the efficient and economical use of America's real property assets and to assure management accountability for implementing Federal real property management reforms." The executive order also directed that "executive branch departments and agencies shall recognize the importance of real property resources through increased management attention, the establishment of clear goals and objectives, improved policies and levels of accountability, and other appropriate action." Despite this mandate, Executive Branch agencies are still struggling to effectively implement asset management principles (Carper 2013).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 104–110, 2020.
https://doi.org/10.1007/978-3-030-48021-9_12

1.1 Purpose

The purpose of this research effort is to help overcome the difficulties in implementation and policy. By developing consensus on what AF asset management is, what asset management should look like, and what goals the AF should set, the researcher intends to align the AF Civil Engineer (CE) community in effort and spirit. This research is primarily focused on answering the following question: How should success in asset management be objectively defined and quantified? Through the process of attempting to answer this question we hope to provide a unity of purpose and direction to the AF CE enterprise.

2 Literature Review

The term asset management is a relatively vague and nebulous term. It can be used to refer to a variety of loosely connected fields. The primary cause for the confusion is that many different industries value and subsequently manage a wide variety of assets. This paper is not focused on investment or enterprise asset management, but rather on infrastructure asset management, defined as "the balancing of costs, opportunities and risks against the desired performance of assets, to achieve the organizational objectives" (ISO 2014a). Additionally, when used in this paper, civil engineers do not necessarily have to possess a Civil Engineer degree or license. Rather, civil engineers are members of the AF CE functional community, and are thus identified by their profession, not their personal expertise or education.

2.1 Asset Management Standardization

Although some consultancy firms and property-owning companies included many best practices, it was not until the publishing of Publicly Available Specification (PAS) 55 in the United Kingdom, that standards for asset management became widely known and available to the industry. Once PAS 55 was released as a British standard in 2004, the Institute of Asset Management began working the process of formalizing their work to be submitted to the International Organization for Standardization (ISO), as the worldwide federation of international standards bodies. This effort came to fruition with the release of ISO(s) 55000-55002. The ISO 55000 series notably emphasizes the importance of an internal 'Strategic Asset Management Plan' (SAMP) that aligns an organization's asset management policies with its individual corporate vision and strategy (ISO 2014a; b; c; Woodhouse 2017). Through the process of following creating a SAMP, an organization is able to tailor the generalized asset management systems described in the ISOs to the needs of the individual organization. Additionally the prescribed SAMP allows organizations to measure their success against asset management objectives that have been tailored to them. The ISO 55000 series provides a globally recognized asset management vernacular which allows companies to benchmark best practices and become certified through and internationally recognized process. The USAF does not currently adhere to the ISO 55000 series and is not obligated to do so. However, the AF might achieve great success by referencing the ISO and building on its principles.

3 Methodology

The primary goal of this research is to examine questions that relate to defining and measuring success within the context of AF asset management, as stated above. We used the Delphi technique for collecting data, with three rounds of responses required. In order to help create policy and drive a shared understanding, this study relies on the opinions and experiences of Subject Matter Experts (SMEs). The researchers originally identified twenty experts as SMEs in the areas of AF asset management application, implementation, or policy. All SMEs were employed by the USAF as Active Duty military members or Government Service (GS) civilians, to ensure that their expertise was relevant to the most current policies and state of the AF. Although no formal criteria were established to be considered as a SME, the researchers selected individuals that worked within the AF Civil Engineer community who were well versed in asset management policy and practice.

Twenty SMEs were originally emailed during the first round of the Delphi technique study and asked to participate. Of those SMEs, only nine individuals provided responses. The second round of the Delphi questionnaire was only sent to the nine SMEs who responded to the first round. Of those nine participants, seven SMEs provided responses for the second round of questions. The third round of the Delphi questionnaire was sent to the nine SMEs that responded to the first round. Of those nine, seven SMEs provided responses for the third round of questions.

3.1 Representativeness of Participants

This research achieves representativeness across the AF CE enterprise by including input from SMEs at a variety of ranks and functional levels. Since asset management is still maturing within the AF CE community, we did not feel that it was advisable to widely solicit opinions from non-experts or non-civil engineers. AF members with less experience may differ in their understanding, but they are unlikely to have a better understanding than the experts.

3.2 Instrumentation

The goal of round one of the study was to solicit individual suggestions and responses in support of our research question. The instrument contained 19 open ended questions. In round two, the participants were provided nine questions from round one, and asked to rank the importance of each of the themes that were identified in the previous round. The goal of round three was to assess agreement on the rankings from round two. To prepare for round three, the individual rankings from round two were aggregated. In round three, the nine participants from round one were asked to rate their level of agreement with the aggregate ranking of themes, where a score of 1 indicated strong disagreement, a score of 2 indicated disagreement, a score of 3 indicated neither agreement or disagreement, a score of 4 indicated agreement, and a score of 5 indicated strong agreement. Participants were also asked to provide feedback whenever they disagreed with the aggregated rankings.

4 Results

When asked to identify and rank the top ways to know if an organization is doing Asset Management well, the SMEs developed the following priority from their aggregated rankings: 1. Meeting defined performance standards, 2. Absence of premature facility failure, 3. Asset management plan progress, 4. Greater preventive maintenance to corrective maintenance ratio, 5. More funding for the unit for identified requirements. This prioritization had an average agreement of 4.43 on the 5-point Likert scale, indicating a moderate agreement.

When asked to identify and rank the best indicators that an organization is effectively applying asset management principles, the SMEs developed the following priority from their aggregated rankings: 1. Improvement on existing performance measures, 2. Reduced lifecycle cost, 3. Complete/accurate inventory, 4. Absence of critical infrastructure failure, 5. Following work priority lists, 6. Fewer inquiries from senior leadership, 7. Meeting minimum facility condition standards, 8. Percentage of personnel trained in asset management. This prioritization had an average agreement of 4.29 on the 5-point Likert scale, indicating moderate agreement.

When asked to identify and rank the ways that the AF should measure success in Asset Management, the SMEs developed the following priority from their aggregated rankings: 1. Whether or not collected facility data is used to inform decisions, 2. Progress towards broad enterprise goals, 3. Amount of work generated by asset management plans, 4. Difference between prioritized work lists and work actually performed, 5. Percent of complete inventory inspections, 6. Mission availability, 7. Percent of personnel that understand asset management, 8. Whether or not the major project priority list reflects outputs from the sustainment management system. This prioritization had an average agreement of 3.14 on the 5-point Likert scale, indicating that there is no consensus of agreement from the SMEs.

5 Analysis

The results from our research did not provide a definite positive answer to how success in asset management should be objectively defined and quantified. However, the results do not prohibit asset management from being objectively defined and quantified. Thus, this section is devoted to summarizing the answer to this research question and seeks to overcome the differences in understanding, as well as provide some of the researcher's suggestions for how success in Asset Management might be objectively defined and quantified.

A large part of the difficulty in this study stems from the AF not establishing a comprehensive definition for asset management. Furthermore, the results indicated that the SMEs had very different understandings of asset management. Without a clear understanding of what asset management is, it is difficult to define what successful asset management looks like. Based on the responses and the available research, we suggest the following comprehensive definition of asset management for use in the AF: "The deliberate and ongoing process of managing an asset throughout its lifecycle by understanding its physical attributes, condition, usage, performance, importance, and environment; while optimizing risk acceptance, financial constraints, future plans, and strategic organizational objectives." However, even with this definition, it would be difficult to define success with just one metric. The following paragraphs discuss some elements that could be used to measure different aspects of success in asset management.

One of the primary goals of asset management is achieving cost savings and increasing efficiency. To this end, the AF might develop metrics to indicate financial success. One potential financial measure could involve comparing operations and maintenance costs at a specific location with a pre-established baseline. While this comparison may be viable for a single location, it could prove problematic when trying to compare performance between bases of varied sizes, missions, and environmental conditions, which might make direct comparison between bases difficult. A secondary challenge could arise if bases were compared to their own baseline.

Although evaluating a base against its own baseline would ensure a fairer comparison, this method would not differentiate between bases that are already doing a good job with less room for improvement than bases that are doing a far poorer job. Another methodology for developing a financial success metric might be to evaluate each facility's annual cost against an AF-wide average for each facility type. This methodology would help to highlight facilities that deviate from the average (either positively or negatively) but would also likely face challenges in normalizing the cost data by location and facility size. Better methodologies for establishing financial success measures may exist within industry but may be overcome by the military spending practices, which would certainly skew cost reporting. All of these reasons make it difficult to develop a fair measure for financial success at each base.

Another potential measure for success would be the absence of facility failure. The AF could develop a metric that that tracks the number of times that a facility has failed over a period of time. This of course would necessitate the AF establishing criteria for what facility failure means. Because of the different aspects of a facility performance, a facility might be considered to have failed when: it is unsafe for occupants, a major system has ceased to work, or the facility otherwise does not enable its occupants to perform their assigned mission. Bases could then be evaluated on the number of 'failure' days that they experience over the course of a year. This might be considered fairer, because although larger bases might have more opportunities for failure (more facilities), they also likely have more personnel and larger budgets to fix or prevent failures than smaller bases. In any case, this metric is limited because it only defines success as the absence of failure but cannot differentiate between bases that are meeting the minimum standard of zero failures, and those that are truly excelling at asset management and have zero failures. If the AF were to take a long-term view of facility failure, a metric could be established that compares how long a facility remains operational and functional compared to its design lifecycle or AF created standard lifecycles. This metric could be fairly evaluated when a facility is demolished, since the ultimate lifecycle of a facility is unknown until that point. This concept might provide a useful benchmark at an AF enterprise level, where there are larger quantities of demolished facilities at a single time but may be too far reaching to evaluate at an individual base level.

Another metric for success suggested by the SMEs is how much work is being generated through the Activity Management Plan (AMP) process. This concept is based on the premise that units committed to asset management principles would be more proactive about identifying requirements before a facility needs repair or replacement. While proactive asset management minded units might generate more work, it is equally likely that creating a metric that measures the amount of work generated through the AMP process might incentivize units to create duplicate or unnecessary work requests, in order to score better on the metric. A similar, but potentially less problematic, metric for success would be to determine how well a unit's projected work list (created through sub-AMPs and BUILDER data) matches the actual work performed. This might encourage bases to strive for better work forecasting, while discouraging bases from addressing emerging requirements. Another potential difficulty with this type of measure is the length of time over which the projected and actual work is being compared.

As highlighted in the discussions above, creating comprehensive metrics for asset management success is fraught with difficulties and hidden challenges. Consulting the ISO 55000 confirms that evaluation of asset management is "indirect and complex" (ISO 2014a; b; c). The ISO does however suggest that any asset management performance measures should be directly tied to the organizational objectives (ISO 2014a; b; c). Since the AF does not have a SAMP or any document that clearly outlines asset management objectives, it will be difficult to create effective performance metrics that measure progress towards these goals. Another suggestion would be for the AF to work towards growing asset management maturity at all levels, as recommended by ISO 55002 and many industry sources (ISO 2014a; b; c; U.S. Department of Interior 2008; United States Federal Highway Administration; American Association of State Highway and Transportation Officials 2013; Woodhouse 2017). Enhancing asset management maturity should be considered an important step regardless of which criteria is selected. Without an underlying asset management maturity and shared understanding, it is likely that any performance measures could be 'gamed' so that units score well without actually doing well.

5.1 Proposed Way Forward for the AF

This research effort has produced many findings that the researcher believes would be of great benefit to the practice of asset management within the USAF. In order to implement the findings and realize the potential of benefits of this research, we recommend that the AF seek to accomplish the following five steps. First, CE leaders should attempt to organize an AF wide asset management summit that gathers the necessary policy makers and practitioners for the purpose of improving AF asset management. Second, the organizers of the summit should seek to inform the attendees that there is in fact a problem with how the AF is currently implementing asset management. The findings of this research and the demonstrated lack of agreement shown in this study may prove very useful for illustrating the AF's problems with asset management. Third, having agreed that there is a problem, AF leaders should designate a single office of responsibility for asset management policy and implementation. The current organizational hierarchy within the AF does not establish a single asset management authority, which can make policy updates complex and problematic. Fourth, this office of asset management responsibility should seek to create a Strategic Asset Management Plan, as recommended by the ISO 55000 series (ISO 2014a; b; c). This SAMP should align the specific mission priorities of the AF with goals for the infrastructure that supports those priorities. The AF might benefit from the vast wealth of knowledge within the asset management consulting industry. Finally, after the SAMP has been established, the office of asset management responsibility should focus on rewriting relevant policy to support the SAMP, and increasing asset management education within the AF, by leveraging the strengths of the Civil Engineer school and the AF Institute of Technology.

6 Conclusion

Asset management is a continuously developing field with great potential for the AF. Every organization is unique and as such, asset management programs must be tailored to specific organizational priorities, strengths, and weaknesses. This research effort solicited opinions from AF SMEs in order to develop an understanding of what success in AF asset management might look like. Some of the major findings of this study were the need to develop both a clear definition of what asset management is and an official SAMP for the AF. Other findings of this research effort included: the importance of leadership buy-in; complete and accurate facility inventory; and understanding of asset management principles at all levels of the organization.

References

Bush, G.W.: Executive order 13327, United States of America, pp. 1–6 (2004). https://www.gpo.gov/fdsys/pkg/FR-2004-02-06/pdf/04-2773.pdf

Carper, T.R.: Federal real property asset management reform act of 2013. U.S. Government Printing Office (2013)

ISO: ISO 55000: Asset management, pp. 1–26. International Organization for Standardization (2014a)

ISO: ISO 55001: Asset management system requirements. International Organization for Standardization (2014b)

ISO: ISO 55002: Guidelines for the application of asset management. International Organization for Standardization (2014c)

U.S. Department of Interior: Asset management plan. Version 3.0. (2008)

United States Federal Highway Administration: American Association of State Highway and Transportation Officials: AASHTO transportation asset management guide (2013)

Woodhouse, J.: Asset management: a science emerging. Int. J. COMADEM **6**(3), 4–10 (2003)

Woodhouse, J.: Developing a strategic asset management plan (SAMP). In: Institute of Asset Management Annual Conference (2017)

A Digital Model of Physical Assets for Long Term Network Resiliency

Thomas Lacroix[1](✉) and Pierre Stevenin[2]

[1] Cosmo Tech, Lyon, France
thomas.lacroix@cosmotech.com
[2] Réseau de Transport d'Électricité, Versailles, France
pierre.stevenin@rte-france.com

Abstract. This paper describes how asset management decisions in energy networks can leverage a new system-driven dynamic ageing model which considers system reliability and resilience. This model is an important part of the new strategic asset management tool called AIO, developed by Cosmo Tech in partnership with RTE.

1 Introduction

Transmission System Operators (TSOs) are responsible for the operation, maintenance, and development of energy transmission networks. One of their most important missions is to design and implement the right maintenance and asset investment strategies in order to ensure grid availability, service continuity, reliability, resiliency, and safety.

To help TSOs achieve these goals, Cosmo Tech is developing, in partnership with RTE, a new strategic asset management tool called Asset Investment Optimization, or AIO (Lacroix and Stevenin 2016). AIO allows TSO asset managers to optimize their maintenance and renewal policies using a complex system modeling simulator.

Part of this complexity relies on the different interactions that exist between entities like policies, assets, human resources, etc. As asset ageing plays an important role in the optimization of the whole system, there is a strong need of a reliable dynamic ageing model that combines different impacts on the degradation of an asset, like maintenance operations, service conditions, or the environment.

This paper describes this dynamic ageing model with a focus on different examples to demonstrate the importance of representing the complexity of the system.

2 The Need for a Dynamic Ageing Model

The ageing of a physical asset is a continuous and progressive phenomenon which is dependent on several actions. Those actions can either increase or decrease the degradation of the asset (Lannoy and Procaccia 2009). We can group those actions in two major classes:

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 111–117, 2020.
https://doi.org/10.1007/978-3-030-48021-9_13

- Workforce actions (operation, maintenance operations, etc.): actions made manually or automatically by a human or an automat/machine on the asset (cleaning, switching off, etc.);
- Environmental actions (corrosion, fatigue, etc.): actions made directly or indirectly by environment on the asset (wind, rain, snow, etc.).

To create an ageing model the principal difficulty is to match two opposing objectives:

- The ageing model must consider the most important actions that impact ageing;
- The ageing model must be easily understood and parameterized by non-expert users.

Moreover, some actions are discrete actions occurring over specific time steps (e.g. painting action on towers). Such actions cannot be predicted at the beginning of the simulation, as the completion of the action depends on several rules or constraints (e.g. painting action on towers can be made only if there are no workforce constraint and no network constraints).

All those considerations lead us to represent an assets ageing using one or several variables which evolve dynamically with actions in each simulation time step (Doyen and Douilhet 2016), (Catrinu-Renstrøm 2016).

Such a model aims to be generic (i.e. suitable for all kinds of assets) and efficient in terms of CPU consumption. It must simulate with a reasonable degree of accuracy a given behavior law, and not determine it.

3 Ageing Model Description

As mentioned previously, even if the model must be quite simple (i.e. two state variables or fewer) we still need to be able to represent at least three different cases. This paper highlights one out of three use cases:

- **the asset's ageing is modeled with a single variable.** This variable represents the apparent age of the asset, which evolves in time. This is the simplest way to model the behavior of an asset;
- **the asset's ageing is modeled with two independent variables.** This case is relevant for an asset made of two different main components ageing in parallel;
- **the asset's ageing is modeled with two dependent variables.** This case is an extension of case 2, where the two different main components of the asset are linked together.

As a result, Cosmo Tech and RTE decided to model the ageing of an asset using two random variables called "state variables" (the word "state" refers to expressions often used in dependability wording (ICE 2014).

The global ageing model is described as followed, for each time step δt:

$$A(t + \delta t) = A(t)(1 - \alpha + (t)) + \beta + \delta t - \gamma(B(t))\delta t$$
$$B(t + \delta t) = B(t)(1 - \alpha/(t)) + \beta/\delta t$$

With:

- A, B: state variables
- $0 \leq \alpha_A, \alpha_B \leq 1$: impact of workforce actions on the state variable;
- $0 \leq \beta_A, \beta_B \leq 2$: impact of environmental actions and/or operating conditions on the state variable;
- $1 \leq \gamma(B(t)) \leq 1$: function representing the coupling between state variables A and B. The formulation of this coupling is based on Hill equation (Hill 1910):

$$\gamma(B(t)) = \frac{p[B_2(t)^2]}{B_{345} + B(t)^2} + q$$

With:

- p: amplitude of the sigmoid;
- q: translation parameter;
- n: control the slope of the sigmoid;
- B_{max}: maximum value of B.

Notice that:

- For case 1, state variable B and γ function do not exist;
- For case 2, γ function is equal to zero. Equation of state variable B has then the same shape as equation of state variable A;

The Hill equation has been chosen for the flexibility and the facility it offers to parameterize the coupling between the two variables. The following curve gives an example of this flexibility (Fig. 1):

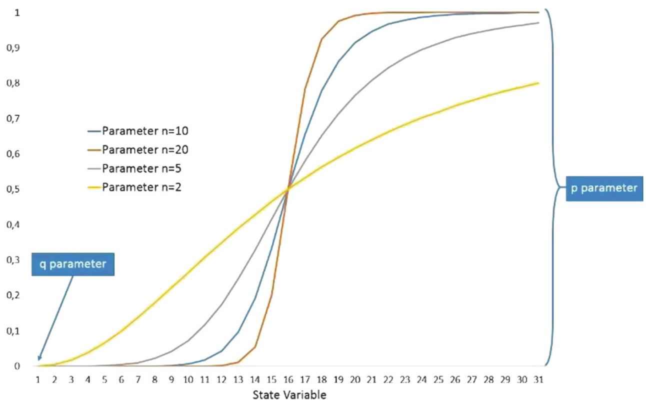


Fig. 1. Hill equation

With these equations, at the beginning of each simulation, we need to determine, for each state variable, two supplementary values:

- The initial value of the state variable at the beginning of the simulation. This value represents the history of the asset (or the component of the asset, if the asset is modeled with two state variables). To fix this value, several factors must be taken into account:
  - The installation date of the asset;
  - The environment where the asset takes place;
  - The maintenance actions made on the asset.
- The maximal value achievable by the state variable. This value represents the moment when the behavior of the asset is modified. This value is calculated using a random selection on a distribution function parameterized by the user. As we mostly work on ageing models for physical assets, we use Weibull distribution. However, it is possible to use other distribution functions like Normal distribution, Log-Normal distribution, etc.

Note: the distribution function must be determined under several hypotheses:

- As workforce actions can modify the value of the state variable, the distribution function must be independent from any workforce action. This means the distribution function is to be taken in case of no workforce action on the asset;
- As environmental actions can either increase or decrease the evolution of the state variable, the distribution function must be taken under normalized environmental and operating conditions, depending on values taken by parameter β.

4 Practical Example

The example provided in this section are based on fictional hypotheses, even if some of the maintenance and renewal operations described here are currently applied by RTE. Hypotheses have been chosen to focus on specific parts of the ageing model.

This case deals with an asset modeled with two independent state variables, A and B. To illustrate this case, we will take an underground oil cable as a practical example, with those fixed hypotheses:

- At the beginning of the simulation, the cable is new;
- The cable is modeled with two components ageing in parallel, in independent ways:
  - Paper: life duration of paper is 50 years old. No maintenance operation can be done on paper. Ageing is irremediable.
  - Oil: life duration of oil is 20 years old (if no maintenance or renewal operation is done). Maintenance operation is possible on oil called oil-degassing: quantity of clean oil push used oil inside the cable. At the end of the cable, used oil is recovered (Welsch 2000);
- A renewal policy is applied to oil cables: they are replaced when their age reaches 45 years old.

In this case, the ageing model is described with the following equations:

$$A(t + \delta t) = A(t) + \delta t$$
$$B(t + \delta t) = B(t)(1 - \alpha/(t)) + \delta t$$

With:

- A: effective age of the paper
- B: effective age of the oil
- $\alpha_B(t)$ = 0,5 if a maintenance operation is done at t
- $\alpha_B(t)$ = 0 if no maintenance operation is done at t

The following scheme explains the way the simulation works in this case (Fig. 2):

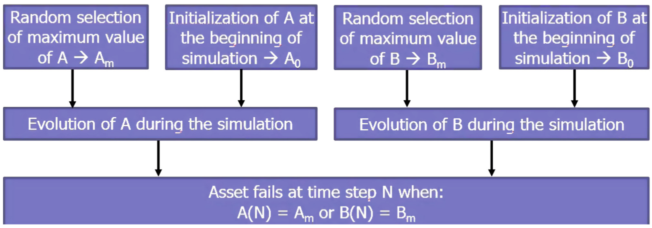


Fig. 2. Simulation steps (case 2)

For this case, we will analyze two different situations:

- Situation 1: we apply a predetermined maintenance policy to the cable: oil-degassing maintenance operation is done every 13 years;
- Situation 2: we apply a condition-based maintenance policy to the cable: an operation is made each year to analyze the oil-dissolved gases into the cable (Boche et al. 2014). If the quantity of gas in the cable is too large, oil-degassing maintenance operation is done the year after the analysis.

The following curves give the evolution of the two state variables of the cable in both situations (Figs. 3 and 4):

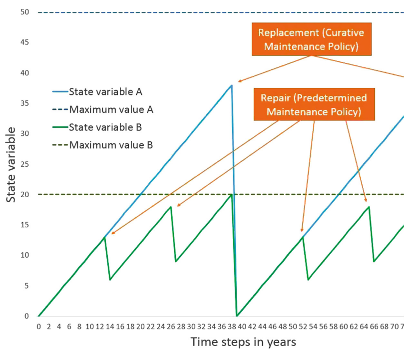


Fig. 3. Ageing curve (case 2 - situation 1)

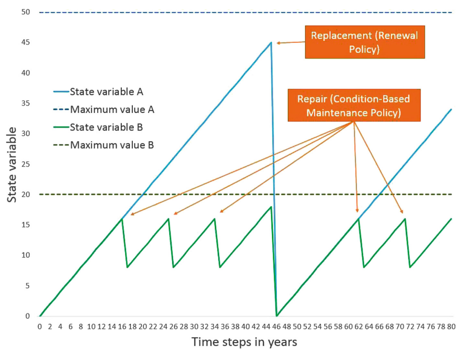


Fig. 4. Ageing curve (case 2 - situation 2)

For each situation:

- Situation 1: the predetermined maintenance policy is correctly applied. It doesn't prevent failure in this case because maintenance operations are not done often enough. As a result, life duration of 20 years of oil has been reached by the cable twice in 80 years of simulation.
- Situation 2: a condition-based maintenance policy is correctly applied, based on the evaluation of the analysis of oil-dissolved gases in the cable. As a result, no failure occurs in 80 years of simulation, and the cable was then replaced as planned by the renewal policy when it reaches 45 years old.

This example is interesting because no one can say which one of the two situations is the best: there are more failures in the first scenario, but maintenance operations cost more in the second scenario. The choice between those two scenarios is a matter of risk that the decision maker is willing to take.

Experts should be able to provide the following data in order to model this case:

- The distribution function of the life duration of each component of the cable, given normalized environmental and operating conditions. This function allows the model to get the maximum values of the state values (30 years and 50 years in this example);
- The values of $\alpha_B(t)$ for each workforce actions (action of oil degassing in this example).

5 Future Improvements of the Ageing Model

This ageing model can be enhanced by considering the actual operation of the asset instead of a temporal average. In this case, operating instantaneous conditions (switching, voltage, current…) act as another kind of punctual action. However, that would require frequent load flow computations.

Another way to make the model either more generic or more accurate could be to add a third state variable, possibly coupled with the two previous ones. This has not yet proved useful for network asset modeling. To the present, representing two assets (such as OHL towers and OHL conductors) has proved a better strategy.

6 Conclusion

The ageing model described in this paper allows the user to consider the dynamic effect of maintenance and the environment on the ageing of a physical asset, with a reasonable computing cost.

Examples given here clearly show the dynamic interactions between maintenance policies and renewal policies. The combination of this interaction coupled with other interactions like human resources, network constraints, and budgets, gives an indication of the complexity of the whole system. It shows the necessity to use simulation tools like AIO to optimize the strategies for maintenance and renewal policies, while controlling system reliability.

Acknowledgments. We would like to thank Serge BLUMENTAL from RTE and Hedi BEN AMOR from Cosmo Tech for their deep contribution in the creation of this ageing model. We also like to thank all the people from Cosmo Tech and RTE who work hard together so that AIO can succeed.

References

Boche, A., Hamadi, M., Fabre, M.: Bilan des analyses de gaz dissous 2009–2014. Écarts par rapport au bilan de 2008 (2014)

Catrinu-Renstrøm, M.D., et al.: D.52 pathways for mid-term and long-term asset management. GARPUR, s.l. (2016)

Doyen, R., Douilhet, R.: VAM, an adaptive open source software to take into account the effect of maintenances and ageing. In: 20ème Congrès de maîtrise des risques et de sûreté de fonctionnement, Saint-Malo, France, 11–13 October 2016 (2016)

Hill, A.V.: The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J. Physiol. **40**, 1–7 (1910)

International Electrotechnical Committee (TS/SC 56): Dependability management – part1: guidance for management and application. International Standard IEC (2014)

Lacroix, T., Stevenin, P.: Asset investment planning: a system driven approach on electrical transmission systems. IEEE Reliability (2016)

Lannoy, A., Procaccia, H.: Evaluation du vieillissement industriel-Méthodologie. Techniques de l'Ingénieur (2009)

Welsch, E.: Faisabilité du dégazage de l'huile de la liaison oléostatique 225kV, La Mouche – Saint Amour (2000)

Approach to Digital Asset Management Service Development

Toni Ahonen(✉), Jyri Hanski, and Teuvo Uusitalo

VTT Technical Research Centre of Finland Ltd., Tampere, Finland
toni.ahonen@vtt.fi

Abstract. Digitalization blurs industry boundaries and disrupts businesses. It offers opportunities for and challenges to asset management related to, for instance, condition monitoring and machine prognostics. In this paper, we propose a top-down approach to digital asset management service development. In this approach, the targets of the services are defined according to strategic guidelines, and customer knowledge is effectively exploited at all levels of the development work. The approach combines activities such as roadmapping, and understanding the customer's decision-making context and business model development.

1 Introduction

Digitalization blurs industry boundaries, disrupts businesses, and offers many opportunities for and challenges to asset management. It offers tools to develop new types of services and distribution through new channels and business logics. Novel digital technologies such as platforms and advanced analytics have the potential to disrupt service design, delivery, and value capture. In this paper, digitalization refers to the new digital technologies that are currently transforming industries. Major transforming digital technology is the Internet of Things (IoT). According to Gubbi et al. (2013) IoT can be defined as an interconnection of sensing and actuating devices that provide the ability to share information across platforms through a unified framework, and with the help of innovative applications, IoT can present a common operating picture. Seamless large-scale sensing, data analytics, information representation, and cloud computing enable the efficient use of IoT. In many manufacturing companies, there is an ongoing shift from the commonly used product development models (e.g. Cooper et al. 2001; Ulrich and Eppinger 2004) and intuitive approaches (Aurich et al. 2006) in service development to more customer need-driven service development approaches. Aurich et al. (2004) describe three service development strategies for manufacturing companies: (1) intuitive service development that is mainly liability driven, (2) systematic service development practices that result in service-enhanced products, and (3) integrated development of products and services that aims at providing the customer with an individual solution consisting of inseparable physical and non-physical components. Companies are moving at varying paces towards systematic or even integrated service development.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 118–127, 2020.
https://doi.org/10.1007/978-3-030-48021-9_14

Agile development approaches, originating from the information systems community (agile software development) (e.g. Abrahamsson et al. 2002), have become increasingly popular also in the manufacturing sector. Some of the main reasons for adopting agile methods are the increasing complexity, data, and service intensity of manufacturing companies and their products. Additionally, iterative and flexible processes with strong customer, employee, and supplier input (i.e. cornerstones of agile methods) are crucial traits of service development (Panesar and Markeset 2008). However, individual experiments may remain disconnected from the strategic focus of companies unless there is a holistic approach in place.

In addition to agile development practices, there is a need for systematic approaches that support the development of services. Therefore, our research question is: "How can we support the development of digital asset management services?" To answer this question, we propose a top-down approach for digital asset management service development. The approach combines activities such as roadmapping, understanding the customer's decision-making context, and business model development. Roadmapping produces guidelines for the service development process. It provides a path for reaching the strategic intent of the company and defines the desired role in the customer's business. The decision-making context connects the service to decision situations the customer faces at operative, tactical, and strategic decision-making levels. Business model development focuses on listing the available business model archetypes, selecting the most suitable ones, and tailoring them to the needs of the company.

2 Roadmapping for Digital Asset Management Service Development

The development of new digital asset management services includes strategic-level decisions. To support these decisions, companies need to understand the development of their business environment. Methods such as scenario planning and roadmapping (e.g. Ahlqvist et al. 2010; Phaal et al. 2004) support strategic decision-making and further the development of the service offering. Paasi et al. (2008) describe a roadmap that includes four phases: 1) state-of-the-art of market, technology, and products and services, 2) determination of drivers and bottlenecks, 3) formulation of visions goals, and 4) determination of intermediate goals.

Determination of the drivers and bottlenecks in a roadmap and using this knowledge as a guideline for developing a service portfolio ensures that the services follow the strategic objectives of the company and create customer value. While the relationship of the services to the customer's processes needs to be considered from the very beginning, it is important to understand both problems and improvement (optimization) opportunities. The following tasks are proposed to support this phase, with specific emphasis on the use of digital technologies in the asset management context.

- Identify bottlenecks of the production systems in terms of, e.g. availability, reliability, maintainability, maintenance support performance, quality, performance, and resource efficiency. In order to better understand the potential, provide a link to the key performance indicators applied by company's customers.

- Identify new opportunities related to how assets are managed to achieve new value and gain knowledge of the desired changes or the increase in the automation level of the processes.
- Mapping of digital technologies and applications to identify new application opportunities and definition of high-level use cases to meet the requirements derived from the phases above.

The formulation of vision and goals repositions a company with respect to digital services based on knowledge of drivers and bottlenecks. Decisions are made based on how large a role a company is willing to take. These decisions, in turn, result in decisions about different service types, e.g. product-related services, production support services, asset optimization and risk sharing, and asset and risk management, as described in Ahonen et al. (2017).

3 Understanding the Customer's Decision-Making Context

Manufacturing and service companies often develop solutions from the perspective of what they think would be valuable to their customers. However, their view of customer needs and limitations might differ from the view of various users of the solution. Companies should consider which customer's decision situations their asset management services support. The need for considering these decision situations is highlighted by the novel digital services that have the potential to benefit several decision situations at many organizational decision-making levels. Many solutions that would make sense from the perspective of a service provider remain unsuccessful because of limited consideration of the customer's decision-making context. As an example, providing solutions that increase the availability of production systems that have a wide maintenance window or increasing the reliability of non-critical systems may not be highly valued by the customer.

Therefore, companies need to know what kind of decision-making situations their customers face and in what kind of situations they can benefit from the services offered (Kunttu et al. 2016). Thus, companies need to understand the customer's decision-making context. Kortelainen et al. (2017) categorize asset management service-related decision situations into decisions for operation performance, maintenance, and investment planning. An alternative categorization connects decision situations to the customer's organizational levels. Table 1 presents a categorization of the decision-making context according to three organizational levels.

Operational-level decision situations typically occur frequently, and decisions must be made quickly. Decision situations at the tactical level are more complex and have greater impact on business performance than decisions made at the operational level. Strategic-level decision-making situations are typically highly complex, have long-term implications for business profitability, and are supported by various decision-support methods (e.g. scenarios and technology forecasts).

Table 1. Understanding key elements of the customer's decision-making context.

	Operational level	Tactical level	Strategic level
Customer's objectives	Daily operation and maintenance	Development of current assets or functions to improve business profitability	Changing current assets or functions to improve business profitability
Customer's constraints	Selected maintenance strategy, maintenance windows, criticality of equipment	Company-level strategic choices, maintenance windows, and criticality of equipment	Operating environment, and company-level vision, values, mission, and strategic choices
Example of user or decision maker	Operators, maintenance technicians, managers	Management	Leadership
Example of decision situation	Focusing maintenance actions on right assets and problems; Timing of maintenance actions; Resource planning	Replacement investment, maintenance strategy formulation	Moving in to new business, market, or technology; Investment in a new production line/plant
Business impact of decisions	Seldom crucial	Moderate to major influence on profitability of business	Major influence on profitability of business
Information required for supporting decision situations	Information on asset condition, type, location, and availability	Information on asset performance, decision alternatives, production bottlenecks, quality, influence of the maintenance tasks	Information on trends and developments in the operating environment, bottlenecks, and improvement potential of current processes
Potential digital services for the decision situation	Remote monitoring and control, predictive analytics of performance, descriptive analytics for causes and effects, failure detection, elimination of unplanned breakdowns	Identification of decision alternatives, selection of an optimal choice for the current situation, and multi-criteria risk analyses	Scenarios and roadmaps, technology forecasts, identification of decision alternatives, multi-criteria risk analyses

4 Asset Management Business Models

Osterwalder and Pigneur (2010) have defined "business model" as the rationale for how an organization creates, delivers, and captures value. Three main dimensions of a firm's business model are value creation, value proposition, and value capture (Clauss 2017). Traditionally, manufacturers have focused on producing a physical product and capturing value by transferring ownership of the product to a customer through sales. Digitalization allows a radical change of this traditional business model. Through access to product data and the ability to anticipate, reduce, and repair failures, the manufacturer has the ability to affect product performance and optimize service. New business models can, e.g. supplement the traditional ownership model by offering new service efficiencies to customers. Another possibility is to offer the product as a service (Porter and Heppelmann 2015).

Digital products and services require rethinking the value proposition. It is necessary to consider the whole lifecycle of the products and services. New ways of selling performance instead of products are possible. Pricing can be based on shared benefits. Data analytics services provide information that can be used in improving the efficiency of processes. However, costs related to collecting, storing, and analyzing data must be taken into account. Digitalization provides new opportunities for continuous cash flow and revenue from lifecycle services. Customer relationships change from individual transactions towards continuous partnerships. To reach customers globally, digitalization provides new opportunities for direct channels (Ahonen et al. 2018).

Rethinking the concept value proposition is a requirement for a successful digital service business. We categorize the digital service concepts for asset management according to their value proposition as follows (further discussion on the services in (Kortelainen et al. 2017; Kortelainen et al. 2017):

1. Provision of a service platform and enabling technologies. This service type includes technological capabilities from measurements to user interfaces. These include collaboration platforms to allow combining data from different sources. The provision of standard reports for customers' O&M function is included in the services, e.g. at an agreed on fixed fee.
2. Analytics-based services. This service type includes analytics models and related capabilities for diagnostics, problem solving, predictive maintenance (condition-based maintenance), and predictive performance management. The provision of required information on different organization and asset-hierarchy levels is in focus. The use of relevant data and knowledge produced internally is also considered (timely and cost-efficient actions based on internal use of digital data and analytics). "Data as a service" (DaaS) and "Information as a service" (IaaS) levels, identified in Kortelainen et al. (2017) according to the DIKW hierarchy (Data, Information, Knowledge, Wisdom), are linked to this service type.
3. Asset performance services. The focus of these services is on the management of customers' assets to maximize the value created in customers' processes. This calls for technological capabilities and analytics models, however, maintaining, operating, and investing in the asset fleets in an optimal manner are at the core of this service. The definition of best practices as an element of the service, for instance,

requires combining analytics capabilities and domain knowledge related to customers' production and business environments. This service type may be further divided into sub-categories, depending on the level of partnership, the exact role of the service provider, and the contract (from product-related services to holistic asset and productivity management). The provider may guarantee a desired performance at a certain cost level: manage the operations, maintenance, and performance of the asset system (e.g. cost per performance contracts). "Knowledge as a service" (KaaS) and "Wisdom as a service" (WaaS) levels (Kortelainen et al. 2017) are linked to the asset performance service type.

5 Developing Digital Asset Management Services

In this paper, we have presented methods and frameworks to support the development of digital services. In Table 2, we position the methods and frameworks within the model for service information requirements (McFarlane and Cuthbert 2012). The methods and frameworks are positioned according to the main elements of service development. Additionally, we present questions that companies should consider when developing digital asset services.

Table 2. Framework for supporting development of digital asset management services (applied from McFarlane and Cuthbert 2012)

Main elements of the service	Methods and frameworks	Key questions
Service need	Roadmapping	What are the drivers and future needs? What are the foreseen technological, market etc. developments? When do the developments take place? What are the most crucial developments for us?
	Understanding customers' decisionmaking context, decision situations and decision-making levels	Are the customer requirements understood? What are the future customer needs? What are the customers willing to pay for? What potential decision-making situations should be supported by refining the relevant data? Can we provide new understanding with the data for, e.g.: - maintenance strategy and implementation - overall equipment efficiency development

(*continued*)

Table 2. (*continued*)

Main elements of the service	Methods and frameworks	Key questions
		- investment assessments and decision-making improvements in energy efficiency
Service offering	Business model formulation	Which capabilities are needed? Who are we working with? Has risk assessment been satisfactorily completed? What service support is required? Is the resourcing within cost? From where and how are we found?
Service operation	Business model deployment	How should the system be operated to optimize value creation? How do we measure the value created by the services? Do we understand the supported functions? Is the system reliable and resilient? Is the related legislation and regulation fully understood? Do we need to improve the service infrastructure?

6 Discussion and Conclusions

In this paper, we proposed a top-down approach to digital asset management service development. The motivation is to support agile development practices with a more systematic methodology. The approach is currently being validated in a research project with participating companies.

We contemplated perspectives that should be considered when developing digital services for asset management. The approach presented in the present paper included roadmapping and scenario planning, understanding the customer's decision-making context and business model development. The integration of these elements with agile development is presented Fig. 1.

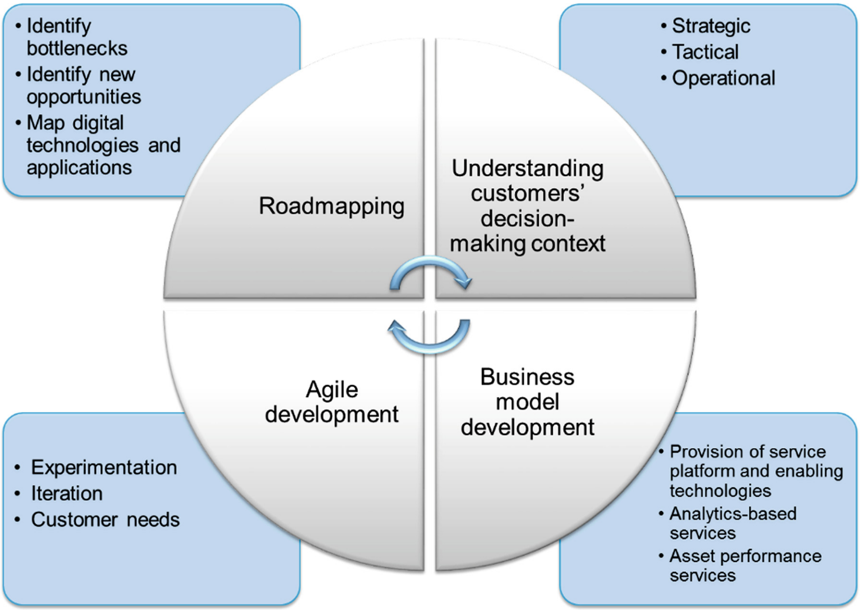


Fig. 1. Approach to digital asset service development.

Roadmapping and scenario planning produce the strategic framework for the service development process. This process results in a development path that shows the means for reaching the strategic intent of the company, and the company's desired role in their customer's business is defined.

The decision-making context connects the service to decision situations at different decision-making levels (strategic, tactical, and operative). Understanding the key elements of the customer's decision-making context supports the development of the digital asset management service.

Rethinking the value proposition is the core task of business model development for digitalized asset management services. For instance, the use of modern capabilities for building analytics for the entire lifecycle of assets or the evolution of digital twins to cover not only R&D and training phases but also to provide lifetime support are drivers for business model transformation.

More information about the implementation and benefits of the service development elements are presented in Ahlqvist et al. (2010) regarding roadmapping, Kortelainen et al. (2017) regarding customer's decision context and Osterwalder and Pigneur (2010) regarding business model development.

Acknowledgements. The authors gratefully acknowledge Business Finland for funding the research project. The companies involved in the research are gratefully acknowledged for their funding and participation.

References

Abrahamsson, P., Salo, O., Ronkainen, J., Warsta, J.: Agile software development methods: review and analysis. Technical Research Centre of Finland, no. 478. VTT Publications, Espoo (2002). https://doi.org/10.1076/csed.12.3.167.8613

Ahlqvist, T., Bäck, A., Heinonen, S., Halonen, M.: Road-mapping the societal transformation potential of social media. Foresight **12**(5), 3–26 (2010). https://doi.org/10.1108/14636681011075687

Ahonen, T., Hanski, J., Hyvärinen, M., Kortelainen, H., Uusitalo, T., Vainio, H., Kunttu, S., Koskinen, K.: Enablers and barriers of smart data-based asset management services in industrial business networks. Presented at 12th World Congress on Engineering Asset Management & 13th International Conference on Vibration Engineering and Technology of Machinery, pp. 2–4 (2017)

Ahonen, T., Hanski, J., Uusitalo, T., Vainio, H., Kunttu, S., Valkokari, P., Kortelainen, H., Koskinen, K.: Smart asset management as a service (2018). http://www.vtt.fi/sites/smartadvantage/PublishingImages/publications/SMACC_SmartAssetManagement_julkaisu_032018_web.pdf

Aurich, J.C., Fuchs, C., DeVries, M.F.: An approach to life cycle oriented technical service design. CIRP Ann. **53**(1), 151–154 (2004). https://doi.org/10.1016/S0007-8506(07)60666-0

Aurich, J.C., Fuchs, C., Wagenknecht, C.: Life cycle oriented design of technical product-service systems. J. Clean. Prod. **14**(17), 1480–1494 (2006). https://doi.org/10.1016/J.JCLEPRO.2006.01.019

Clauss, T.: Measuring business model innovation: conceptualization, scale development, and proof of performance. R&D Manag. **47**(3), 385–403 (2017). https://doi.org/10.1111/radm.12186

Cooper, R.G., Edgett, S.J., Kleinschmidt, E.J.: Portfolio Management for New Products. Perseus Publishing (2001). https://doi.org/10.1016/S0737-6782(97)90046-1

Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gen. Comput. Syst. **29**(7), 1645–1660 (2013). https://doi.org/10.1016/j.future.2013.01.010

Kortelainen, H., Hanski, J., Kunttu, S., Kinnunen, S.-K., Marttonen-Arola, S.: Fleet service creation in business ecosystems – from data to decisions, no. 309. VTT Technology (2017)

Kortelainen, H., Happonen, A., Hanski, J.: From asset provider to knowledge company - transformation in the digital era. Presented at 12th World Congress on Engineering Asset Management & 13th International Conference on Vibration Engineering and Technology of Machinery, Brisbane, 2–4 August 2017 (2017)

Kunttu, S., Ahonen, T., Kortelainen, H., Jantunen, E.: Data to decision-knowledge-intensive services for asset owners. In: Proceedings of EuroMaintenance (2016)

McFarlane, D., Cuthbert, R.: Modelling information requirements in complex engineering services. Comput. Ind. **63**(4), 349–360 (2012)

Osterwalder, A., Pigneur, Y.: Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. Wiley, Hoboken (2010)

Paasi, J., Valkokari, P., Maijala, P.: Managing opportunities, risk and uncertainties in new business creation - working report. VTT Technical Research Centre of Finland (2008). http://virtual.vtt.fi/virtual/proj3/innorisk/Innorisk_progress_report_VTT2008.pdf

Phaal, R., Farrukh, C.J.P., Probert, D.R.: Technology roadmapping—a planning framework for evolution and revolution. Technol. Forecast. Soc. Change **71**(1–2), 5–26 (2004). https://doi.org/10.1016/S0040-1625(03)00072-6

Porter, M.E., Heppelmann, J.E.: How smart, connected products are transforming companies. Harvard Bus. Rev. **93**(10), 96–114 (2015)

Panesar, S.S., Markeset, T.: Development of a framework for industrial service innovation management and coordination. J. Qual. Maint. Eng. **14**(2), 177–193 (2008). https://doi.org/10.1108/13552510810877674

Ulrich, K.T., Eppinger, S.D.: Product Design and Development. McGraw-Hill/Irwin (2004)

Risk Management of Complex Systems: Understanding the Difference Between Systematic and Systemic Failures

Jon Tømmerås Selvik[1](✉) and Jean-Pierre Signoret[2]

[1] University of Stavanger, Stavanger, Norway
jon.t.selvik@uis.no
[2] Project Leader of ISO/TR 12489:2013, Pau, France
j-p.signoret@orange.fr

Abstract. When dealing with risk related to complex systems, and for safety systems in particular, a main task is to identify and assess events that could threaten reliability performance of the system. With respect to this task, there are two important types of failures commonly referred to, i.e. systematic and systemic failures. Both of these relate to system thinking somehow, however, the meaning and importance of the two concepts should be distinct. In this paper, we study this distinction and relevance to complex systems by addressing available definitions of these given in international standards and giving some qualitative examples. Despite the exact definition of the two concepts may vary, a main conclusion is that it is only when dealing with complex systems, that it is appropriate to use both concepts for the purpose of risk management.

1 Introduction

When attempting to describe system reliability and risk using mathematical models, a starting point is the relationships between the components defining the system. These relationships are important for how they contribute to the overall system reliability performance, and the components could for example be ranked based on their contribution, using different importance measures (see e.g. Kuo and Zhu 2012; Natvig 2010; and IEC 61078:2016). However, the reliability contribution is typically not depending only on one single component. Typically, as is the situation when dealing with complex systems, there might be interdependencies influencing the system performance. Such fragility (dependencies) in the systems is a main reason that systemic failures occur (Venkatasubramanian 2011). Hence, in risk management, a key is to understand how failure events influence system reliability performance.

According to Johansen and Rausand (2011) 'complexity' refers to "*a state of difficulty in determining the output of a system based on knowledge about individual inputs and given our current knowledge base*." Translated to the system reliability model; despite the way the individual components interact or dependencies in complex systems is relevant to the understanding of the system vulnerabilities, it may be challenging to capture these interactions in a mathematical model.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 128–136, 2020.
https://doi.org/10.1007/978-3-030-48021-9_15

The issue of how to capture the dependencies should be given high attention when dealing with for example safety systems, as the understanding of this is a key to appropriate reliability modelling and calculation of these systems, as pointed out in the ISO/TR 12489:2013 (see also Rausand 2014).

Furthermore, to capture possible failure events that could cause critical system failures in safety systems, one should see beyond the traditional barrier thinking, where an assumed set of individual and independent components protects the system from going down (collapse). Holland (2014) refers to a property of complexity called 'emerging', which could be translated into the same meaning as the ISO/TR 12489:2013 places into 'holistic' or the *Aristotle principle*, i.e. that "*the sum is greater than the sum of its parts*". The main idea being that, there could be possible failure events that is not explainable through only focusing on the individual system components. Some sort of system understanding, or perspective, is needed to see beyond the "parts".

Related to the 'system', the following triplet are relevant to the above need for failure concepts covering a system perspective:

- System (level) failure
- Systematic failure
- Systemic failure

While 'system failure' or 'system level failure' is intuitively simple to understand as any failure occurring at a system hierarchy level (see e.g. ISO 14224:2016; for example, on a 'gas export' system in the oil and gas industry), the other two is not as straightforward, and lacks adequate clarification in risk and reliability literature. It is repeatedly in the ISO/TR 12489:2013 pointed out, that the analyst should be careful not to confuse the two concepts. Both relate to the 'system' somehow. But in a separate way, and with room for different interpretations. A main objective of this article is to provide some clarification regarding the distinction between these concepts based on the guidance given in key international standards on the issue.

The following of the article is structures as follows: in Sects. 2 and 3 we give a brief overview of alternative ways of interpreting the concepts 'systematic failure' and 'systemic failure' in the context of complex systems. Then, in Sect. 4 we discuss the relationship and difference between the two, where different examples are used for clarification. Finally, in Sect. 5, we give some concluding remarks.

2 What Is a Systematic Failure?

As expected, the concept 'systematic failure' is widely applied across different industries and sectors and is often linked to failures having a common cause. Traditional textbooks on system reliability commonly refer to 'systematic failures' as an opposite of random failures, i.e. those being non-random, (see e.g. Rausand and Høyland 2004; Gentile and Summers 2006).

Based on the international standards (from the International Electrotechnical Committee, IEC, and the International Organization for Standardization, ISO), there are currently two alternative ways of defining the concept. We refer to the two as the 'new definition' and the 'old definition'.

Starting with the 'new definition', the natural place to search for this, is in the IEV – the International Electrotechnical vocabulary, which includes a recognized glossary of key terms; the dependability standard IEC 60050-192:2015 issued by IEC. This define the 'systematic failure' concept as a: "*failure that consistently occurs under particular conditions of handling, storage or use.*" Hence, according to the definition, a systematic failure can then be reproduced by deliberately applying the same conditions, as these conditions influence the failure occurrence. However, the aspect that the failure is reproducible, does not make it systematic. The conditions may come from, as the IEC notes, errors in specification, design, manufacture, installation, operation or maintenance of the item. The definition is widely used across different industries. It is for example used within oil and gas (in in both the ISO 14224:2016 and the ISO/TR 12489:2013) and railway (IEC 60050-821:2017).

For the 'old definition', there are several international standards, including the systems and software engineering standard ISO/IEC 15026-1:2013 and the safety standards ISO 26262-1:2011, ISO 25119-1:2010 and ISO 13849-1:2015, all developed before the 'new definition' issued in 2015, which still are using the old definition of 'systematic failures' given below from the IEC 60050-191:1990: "*failure related in a deterministic way to a certain cause, which can only be eliminated by a modification of the design or of the manufacturing process, operational procedures, documentation or other relevant factors.*" The same definition is also used in nuclear industry (IEC 60050-395:2014/AMD1:2016), which refer to the widely used IEC 61508:2010 (cf. part 4) produced for reliability analysis of safety instrumented systems as source.

Reading the two definitions, a key issue is whether the "*consistently occurs under particular conditions*" is given the same interpretation as the "*related in a deterministic way to a certain cause*", where focus is shifted from failure causes to conditions. It is not obvious that the change from "*in a deterministic way*" to "*consistently*" imply any different meaning.

Consider the following example. Say that some car manufacturer by mistake produces a poor batch of sensors, which cause the engine to stop when the speed reaches 150 km/h, and then delivers cars with these installed. When the cars later enter the stores and the streets, the reliability of the car is obviously reduced because of this. The car will have a dormant fault that could cause a failure and accident under the right conditions. The failures occur because of the sensor issues, and this could then be traced directly to the conditions of use. The products installed on the car then have a consistent effect on the reliability performance related to all the car with these sensors installed that are in use under the same conditions (situations). In normal driving conditions there is no problem but in case the driver infringe the speed limitations, the car will systematically stop. This could be linked to the failure causes, i.e. here the sensor fabrication failure. Meaning that systematic failures is considered as the opposite to random failures in both definitions. See also examples in Rausand (2014, pp. 66–67).

If we look outside the two definitions and consider the meaning of the word 'systematic', for example as it is defined in common dictionaries. One finds the following (Oxford dictionaries, n.d.): "*Done or acting according to a fixed plan or system; methodical. For example, a systematic search of the whole city*". A systematic failure is implying that there is some system behind the way the failure occurs (i.e. it not random), as is given in the 'new definition' by the relationship between the "*conditions of handling storage or use*" and the failure events. This is necessarily a deterministic relationship, as is also suggested from the old definition in the IEC 60050-191:1990, as the change in conditions could lead to failure events. It is why the systematic failure can be removed by modifying the design.

3 What Is a Systemic Failure?

If we search for a definition of 'systemic failure' in the IEC 60050-192:2015, it is not to be found. The IEV does not define this concept. However, a safety-related ISO definition is given in the ISO/TR 12489:2013: "*failure at system level which cannot be simply described from the individual component failures of the system.*" Such failures are in the technical report used synonymous to 'holistic failures'. A main point is that they focus strictly on the system level. Furthermore, the definition indirectly distinguishes between failures that can and cannot be described from a simple combination of the individual components failures. Consequently, not all failure events causing system downtime, are systemic. The component relationships must also be considered. It makes the concept valid only to complex systems.

If we look up 'systemic' in dictionary, the above distinction is not as obvious. The Oxford dictionaries (n.d.), as one example, define the word (adjective) 'systemic' as: "*relating to a system, especially as opposed to a particular part*". The dictionary also provides an example that "*the disease is localized rather than systemic*". Based on this understanding, the failure of a system could itself be sufficient to label the failure as systemic. The dictionary definition merely implies that a system perspective is taken, i.e. that the failure is on system level.

Currently, no other ISO or IEC definitions of the concept 'systematic failure' are available to challenge the ISO TR/12489 definition. However, there exist a related definition in another ISO technical report on financial services that could challenge the understanding, i.e. on the concept 'systemic risk' (ISO/TR 13569:2005); "*risk that the inability of one of the participants to meet its obligations, or a disruption of the system itself, could result in the inability of other system participants or of other financial institutions in other parts of the financial system to meet their obligations as they become due*". Although given for 'risk', the definition suggests basically that, the failure of one or several components could cause other components and systems to fail, which are not including any reference to complex systems or that the relationship (dependencies) between the components matter. In fact, the core of this definition is that one could have a sequence of failures due to so-called 'Domino effects' that eventually cause the system to go down (collapse). Within reliability theory, such failures are referred to as 'cascading failures'; meaning that the failure of one component or element is causing others to fail (the Domino effects). However, such sequence of failures does not necessarily mean that such failures conflict with the basic idea of 'systemic failure', stating that the system failure cannot be deduced directly from the component failures.

The following trading example captures the systemic aspect: say that one participant sells a high number of shares, and then the price of these shares decreases a little bit. Next, some other participants monitoring the activity detect this decrease, and to avoid losses, sell their own shares, causing the price to decrease even more: causing many or all participants to decide to sell their shares, and the price goes to zero and the system collapse. It is easy to see that this could happen within high frequency trading when the computers take the same decisions at the same time. The failure is labelled as systemic, when there are interdependencies influencing the reliability performance of the components (elements) and of the overall system. The challenge is to understand such underlying patterns and reflect these in analysis and decision-making. Further examples are provided in the following section.

4 Examples and Discussion

As should be quite clear given the previous sections, a failure labelled as 'systemic' is different from a failure being 'systematic', although both relates somehow to system thinking. 'Systemic' has to do with the reliability performance of the system, not the system aspect in causes or conditions leading to the failure. In fact, systematic failures do not necessarily influence the system at all. For example, a systematic failure could lead to several components going down, but where there is no critical failure at the system level.

We will consider a 'photocopier' example to clarify the difference between the two concepts. It is a system that is complex yet simple to understand. Say in an office area there are two photocopiers: A and B. One reoccurring challenge is that poor paper quality causes paper jams, which then leads to photocopier down time. What would then be a systematic and systemic failure here?

Systematic failures are related to events that under the given condition (in a deterministic way) influence the functioning of the entire or parts of the system. Here the probability of a paper jam is not related to the photocopier itself, but to the probability of the mistake (adding poor paper). If both A and B are fed with the same poor type of paper, in this situation, the systematic failure is also a common cause and common mode failure.

For systemic failures, the overall system must be considered; both A and B. If A fails, the load of B is doubled (and vice versa). Hence, the probability of failure of A, P_A, depends on what happens on B. It is not possible to calculate the probability of failure of the overall system just by combining P_A and P_B as independent components. What then if A and B are different product types and B accept paper of lower quality than A? Then, in case of mistake, a systematic failure will occur on A, but not on B. The systematic failure on A, however, participates to the systemic failure of the overall system (A and B). The probability of failure cannot be obtained by the combination of the probability of failure of the components.

The example above, although not going the reliability specifics, clearly express the importance of interdependencies and what is the difference between the two concepts. A similar example of systemic failures may be given on cars and parking sensors. Despite a generally high reliability of such sensors, studies (e.g. Kidd et al. 2014) indicate that cars having such sensors are not necessarily involved in significantly less parking accidents compared with cars without. Experimental evidence concerning the effectiveness of reverse parking sensors has been mixed on the issue (Keall et al. 2017). A main reason for this, is obviously that, people behave and drive differently when having such equipment installed. Similarly, for the effect of rear view cameras, the US National Highway Traffic Safety Administration (NHSTA 2006) concluded that: "*the true efficacy of rear-view camera systems cannot be known without assessing drivers' use of the systems and how they incorporate the information into their visual scanning patterns (i.e., looking behind over the shoulder and glancing at mirrors).*" See also study in Hurwitz et al. (2010). Hence, when assessing the reduced risk of car parking accidents by introducing 'new' sensor technology, there is a need to consider the effects on the entire system, including how the technical system influences the driving behaviour. However, that accidents occur due to e.g. the parking sensors not reading the correct distance (reading error), would be a systematic failure issue.

For larger systems, the modelling might be more challenging, in particular systemic failures could be challenging to deal with as they could be well hidden by the complexity of the system. Nevertheless, the key is to understand how interdependencies influence the performance of the system and may cause critical events.

Consider another example, a critical event taking place at the, an incident occurring in a gas-graphite nuclear power plant in France in the late sixties, which clearly highlighted the need to capture systemic aspects in risk analysis. In this incident, a part of the core melt due to an operator error in the feeding of the reactor. Such an event had already been anticipated by the installation of an ashtray behind the reactor, and this ashtray was designed to melt above 1400 °C. Unfortunately, the mix between the melting core and the substance of the ashtray formed a eutectic, and this eutectic was melting at 800 °C. This unexpected systemic effect lead to the melting of the ashtray, resulting in radio-active substance falling to the bottom of the containment. More than a year was needed in cleaning the containment before restarting the reactor. It was a failure that could not be analysed in terms of core melting and ashtray failure independently from each other: the whole 'core melt and ashtray' system had to be considered. For further details about the 1969 Saint-Laurent-des-Eaux incident; see e.g. IRSN (2015).

The above examples show the importance, in situations of complexity (i.e. presence of non-linearity), to address systemic failures in risk management. And demonstrate why it is important to distinguish between such failures and those that are simply 'system level failures' in risk management.

Within medicine and patient health care, a way of capturing such dependencies is identified through the SEIPS model (Systems Engineering Initiative for Patient Safety; see Carayon et al. 2016). This model highlights the need for addressing the whole system by connecting the work system (i.e. the environment, organization, person, tasks and technology and tools) to the processes and outcomes. It is a framework for understanding the influence the different elements of the working system such as new technology may have and their relationships. The principles of the SEIPS model is relevant and applicable also to other complex systems where it is just as relevant to capture the effects of both systematic and systemic failures.

While the systematic failures may be described focusing on latent failure conditions and failures of the system elements, it should be clear that the systemic failures must be strictly addressed from a system perspective. Consequently, the risk management should be based on an approach that captures the interactions between the system components, such as for example through e.g. the classic STAMP model (System-Theoretic Accident Model and Processes), where, as is the key, the use of the models allows for; *"integrating non-linear relationships reflecting the behavioural dynamics controlling the behaviour of the entire technical and organizational structure over time"* (Leveson 2004).

5 Concluding Remarks

The concepts discussed in the current paper, all provide essential information about failures in relation to risk management when dealing with complex systems. However, as identified partly through study of internationally accepted ISO and IEC documents, the concepts may be attributed with distinct meanings. Where the definition of 'systematic failure' seem to be widely agreed upon, the 'systematic failure' concept appear to lack appropriate foundation in risk theory; e.g. might be mixed with 'system level failure'. Furthermore, we find that the concept of systemic failure, although it should be considered appropriate for only complex systems, is in management of 'systemic risk' also applied for systems that are not that. Besides, there exist a variation in definitions outside the ISO and IEC domain, for example within finance and safety (e.g. Santos-Reyes and Beard 2001), which further reflects a need to produce ISO and IEC guidance on 'systemic risk' along with 'systemic failures' in a consistent way. Particularly, the role of interdependencies should be recognized when managing the systematic and systemic failures and associated risks. A key is to understand the ability of frequently used models to address the full spectre of elements influencing the reliability performance of the system. In the paper, the SEIPS and STAMP models are mentioned, as two models that can capture systemic aspects in a risk management of complex systems.

References

Carayon, P., Hundt, A.S., Karsh, B.-T., Gurses, A.P., Alvarado, C.J., Smith, M., Brennan, P.F.: Work system design for patient safety: the SEIPS model. Qual. Saf. Health Care **15**(Suppl I), i50–i58 (2016)

Gentile, M., Summers, A.: Random, systematic and common cause failures: how do you manage them? Process Saf. Prog. **25**(4), 331–338 (2006)

Holland, J.H.: Complexity: A Very Short Introduction. Oxford University Press, Oxford (2014)

Hurwitz, D.S., Pradhan, A., Fisher, D.L., Knodler, M.A., Muttart, J.W., Menon, R., Meissner, U.: Backing collisions: a study of drivers' eye and backing behaviour using combined rear-view camera and sensor systems. Inj. Prev. **16**, 79–84 (2010)

IEC 60050-191:1990: International electrotechnical vocabulary: Chapter 191: Dependability and quality of service. International Electrotechnical Commission, Geneva, Switzerland

IEC 60050-192:2015: International electrotechnical vocabulary - part 192: Dependability. International Electrotechnical Commission, Geneva, Switzerland

IEC 60050-395:2014/AMD1:2016: International electrotechnical vocabulary - part 395: Nuclear instrumentation: Physical phenomena, basic concepts, instruments, systems, equipment and detectors. International Electrotechnical Commission, Geneva, Switzerland

IEC 60050-821:2017: International electrotechnical vocabulary - part 821: Signalling and security apparatus for railways. International Electrotechnical Commission, Geneva, Switzerland

IEC 61078:2016: Reliability block diagrams. International Electrotechnical Commission, Geneva, Switzerland

IEC 61508:2010: Functional safety of electrical/electronic/programmable electronic safety-related systems - all parts. International Electrotechnical Commission, Geneva, Switzerland

IRSN - Institut de Radioprotection et de Surete Nucleaire (France): information note on accidents which affected nuclear reactors of the Saint-Laurent-des-Eaux in 1969 and 1980 (INIS-FR-15-0588) (2015). (in French)

ISO 13849-1:2015: Safety of machinery - safety-related parts of control systems- part 1: general principles for design. International Organization for Standardization, Geneva, Switzerland

ISO 14224:2016: Petroleum, petrochemical and natural gas industries - collection and exchange of reliability and maintenance data for equipment. Edition 3. International Organization for Standardization, Geneva, Switzerland

ISO 25119-1:2010: Tractors and machinery for agriculture and forestry - safety-related parts of control systems - part 1: general principles for design and development. International Organization for Standardization, Geneva, Switzerland

ISO 26262-1:2011: Road vehicles - functional safety - part 1: vocabulary. International Organization for Standardization, Geneva, Switzerland

ISO/IEC 15026-1:2013: Systems and software engineering - systems and software assurance - part 1: concepts and vocabulary. International Organization for Standardization, Geneva, Switzerland

ISO/TR 12489:2013: Petroleum, petrochemical and natural gas industries - reliability modelling and calculation of safety systems. International Organization for Standardization, Geneva, Switzerland

ISO/TR 13569:2005: Financial services - information security guidelines. International Organization for Standardization, Geneva, Switzerland

Johansen, I.L., Rausand, M.: Complexity in risk assessment of sociotechnical systems. In: Proceedings of the 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference, pp. 2274–2283. Curran Associates, Helsinki (2011)

Keall, M.D., Fildes, B., Newstead, S.: Real-world evaluation of the effectiveness of reverse camera and parking sensor technologies in preventing backover pedestrian injuries. Accid. Anal. Prev. **99**, 39–43 (2017)

Kidd, D.G., Hagoski, B.K., Tucker, T.G., Chiang, D.P.: The effectiveness of a rearview camera and parking sensor system alone and combined for preventing a collision with an unexpected stationary or moving object. Hum. Factors **57**(4), 689–700 (2014)

Kuo, W., Zhu, X.: Importance Measures in Reliability, Risk, and Optimization: Principles and Applications, 1st edn. Wiley, Hoboken (2012)

Leveson, N.: A new accident model for engineering safer systems. Saf. Sci. **42**, 237–270 (2004)

Natvig, B.: Multistate Systems Reliability Theory with Applications. Wiley, Hoboken (2010)

NHSTA – The National Highway Traffic Safety Administration: Vehicle backover avoidance technology study. Report to congress. National Highway Traffic Safety Administration, U.S. Department of Transportation, Washington DC, USA, November 2006

Oxford dictionaries. Definition of systematic. Oxford University Press (n.d.). https://en.oxforddictionaries.com/definition/systematic

Oxford dictionaries. Definition of systemic. Oxford University Press (n.d.). https://en.oxforddictionaries.com/definition/systemic

Rausand, M.: Reliability of Safety-Critical Systems - Theory and Applications. Wiley, Hoboken (2014)

Rausand, M., Høyland, A.: System Reliability Theory - Models, Statistical Methods, and Applications. Wiley, Hoboken (2004)

Santos-Reyes, J., Beard, A.N.: A systematic approach to fire safety management. Fire Saf. J. **36**, 359–390 (2001)

Venkatasubramanian, V.: Systemic failures: challenges and opportunities in risk management in complex systems. AIChE J. **57**(1), 2–9 (2011)

Sensors and Process Monitoring Models Applied for Pre-salt Petroleum Extraction Platforms Applications

Luigi Galotto Jr[1(✉)], João O. P. Pinto[1(✉)], Cristiano A. Quevedo[1(✉)], Herbert Teixeira[2], and Mário C. M. Campos[2]

[1] Federal University of Mato Grosso do Sul - UFMS, Campo Grande, Brazil
{luigi.galotto, joao.pinto, cristiano.andrea}@ufms.br
[2] CENPES Research Center of PETROBRAS, Rio de Janeiro, Brazil
{herbertteixeira, mariocampos}@petrobras.com.br

Abstract. This work presents a modelling methodology for sensors and equipment condition monitoring developed during a research project to enhance dependability of pre-salt petroleum extraction platforms. The methodology aims to improve the capability of the auto-associative models applied for sensors monitoring in the last decades in nuclear power plants, chemical industry, refineries, gas transport and processing plants. However, actual operation problems or fault in equipment also may lead to false measurement error detection. This problem observed in the previous applications motivated the development of the improved method able to detect measurement errors and fault conditions in the process or equipment. This improvement has been obtained adding data (real or simulated) of the different conditions of operation, including the fault conditions (undesired data in the previous methodology). Therefore, the models become able to make accurate sensor estimation, even under fault conditions in the monitored process, and they also give a proper fault diagnoses about the measurement instruments and the process reducing false alarms compared to the traditional approaches. Also, some modelling challenges were observed during the development such as optimization of parameters, memory size and computing complexity. The methodology is demonstrated using simulated a process of a Petroleum Platform application. The achieved results showed a possible methodology to improve or replace the traditional approaches in the past application.

1 Introduction

Several works have been done in the last decades in order to improve the reliability of the industrial processes (Kristjanpoller et al. 2016), in controllers (Markus and Ping 1997), actuators (Sharif and Grosvenor 1998) and sensors (Hines and Garvey 2006). Usually all equipment can be analysed using the processes measurements, but fault detection in sensors showed to have an extra complexity because these are the providers of the measurement. Sensors validation is an important subject in sensor based monitoring and control systems. In petroleum and natural gas industry, as in many others, reliable measurements are key to guarantee to the plant optimum and safe

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 137–144, 2020.
https://doi.org/10.1007/978-3-030-48021-9_16

operation point and, in some cases, trustable billing. Therefore, many years of Research and Development projects have been performed in order to improve diagnostic tools especially for the industry of petroleum and gas (Galotto et al. 2015), exploring different techniques using analytical redundancy in measurements. Usually, the most accepted memory based methods such as the Auto-Associative Kernel Regression (AAKR) (Galotto et al. 2006) and the Auto-Associative Multivariate State Estimation Technique (AAMSET) are applied exploring the analytical redundancy among sensors related with the process, equipment and set points. These methods brought important contribution to several applications in control, optimization and maintenance since the correct operation of each measurement instruments could be verified avoiding operation problems and false alarms due to measurement errors. The Auto-Associative Kernel Regression (AAKR) showed to have the best generalization in many problems and it is also easy to train or adapt to new data, since it is a memory based non-parametric technique.

Since most of the process diagnosis techniques are dependent on the correct measurements and also the correct process operation is necessary to avoid false alarms in sensor fault detection, it may be observed an intrinsic dependency of all diagnostic techniques. It makes attractive the idea to get an integrated tool to perform monitoring of different parts in the process. This work proposes a methodology to create AAKR models to both tasks: sensors and process monitoring. The described process monitoring may be extended on fault detection of actuators, pipes, pumps, compressors or any other structured that can be modelled in the process.

2 Evaluated and Modelled Processes

The simulated evaluation of the methodology has been made using models of some real processes. These models are necessary due to the possibility to create any kind of fault conditions, allowing the validation of the proposed methodology. The main modelled system is an oceanic platform for petroleum exploration. The components of this platform will be briefly explained and a specific flow process will be focused to explain the methodology.

2.1 Petroleum Exploration Platform

The petroleum exploration platforms of the pre-salt oil, was chosen to be a critical process with direct application of the developed methodology. These platforms operate far from the continent in complex and extreme conditions, making high dependability processes mandatory. For ideal operation, all sensors and actuators must work correctly; otherwise, it may reduce the quality and the efficiency of the processes or even the instability of the control loops. The fault detection and diagnosis allow the improvement of the safety, reliability and the maintainability and asset management of the platform operation.

A representation of the dynamics of a platform is composed basically of: oil wells in the ocean floor, manifold, separation stage and storage. The oil from the ocean well modelled in this project is sent to the manifold. Such equipment used in the collection has the function of distributing oil production to the high-pressure separation systems. In the computational system in Fig. 1, the manifold is represented by a multiplexer, and has the function of directing the obtained products to the separation stages. The separators are modelled with demultiplexers with flow control of oil and water, and gas pressure control.

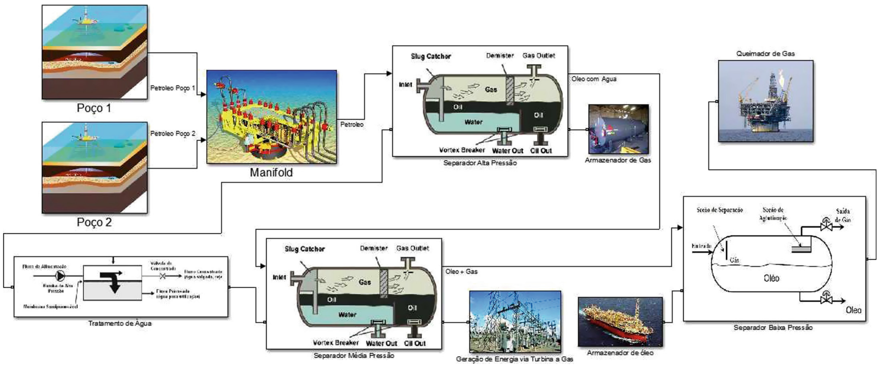


Fig. 1. Integrated models of the platform.

As a result of the first separation process performed by the high-pressure separator are obtained water, oil and gas. The oil of the first separation stage is sent to the second separator, called the medium pressure separator. Similarly, the second separation stage also results in water, oil and natural gas. Finally, the oil produced from the second separation stage is still sent to a third separation stage, called the low-pressure separator. The materials resulting from the low-pressure separator are oil and gas for exportation (transport to mainland). The water is treated and used for internal consumption of the platform. The gas from the second stage is sent to a dehydration section, and is used for power generation on the platform, and the surplus is sent to the flares. The generation of energy is carried out through natural gas turbines. The Fig. 1 shows the complete model of the platform, where the blocks are subsystems containing specific diagrams representing the each described part of the platform.

The detailed explanation of all subsystems is not possible in this paper. However, it is enough to describe one specific process of the oil flow, which was chosen to be detailed and was used to explore the presented methodology.

2.2 Oil Flow Process with Ball Valve

The valves are the most usual actuators in oil and gas industry processes and this is the main reason of the preference to use this model in paper. The detailed simulation model with ball valve is presented in Fig. 2. This model allows getting upstream, downstream and differential pressures measurements, and also the flow measurement and the *MV* command signal to the ball valve.

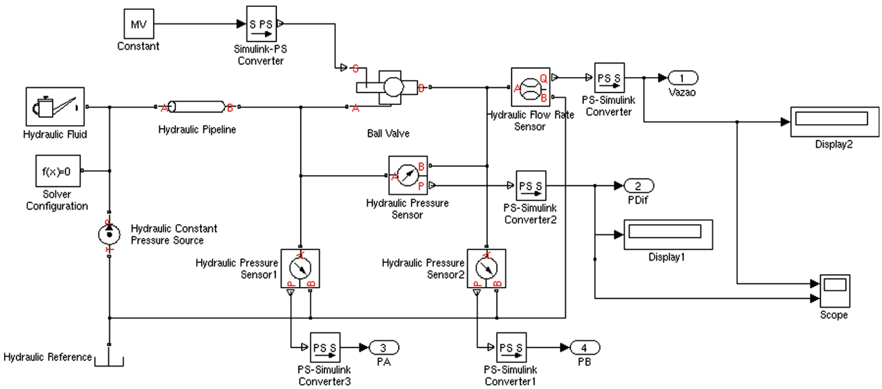


Fig. 2. Test model with Simulink/SimScape.

This model was used to generate the data in 4 different operating conditions: normal, partial fault of the valve opening 50% of the command, with complete valve failure in the opening and valve failure to close. The data generated for these fault conditions are shown in Fig. 3.

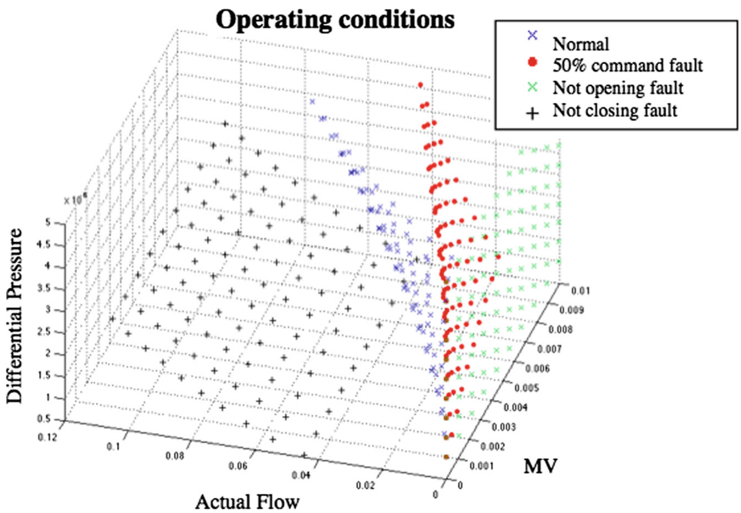


Fig. 3. Simulated data with the oil flow process model.

3 Proposed Methodology

3.1 Auto-Associative Models Review

The diagnostic method (George et al. 2006) was designed to be as close as possible to the current methodology of software tool presented in (Galotto et al. 2015) developed in previous researches.

The inputs $\boldsymbol{X}$ and outputs $\boldsymbol{Y}$ and the estimated outputs of an auto-associative model are organized in matrix form as shown in Fig. 4.

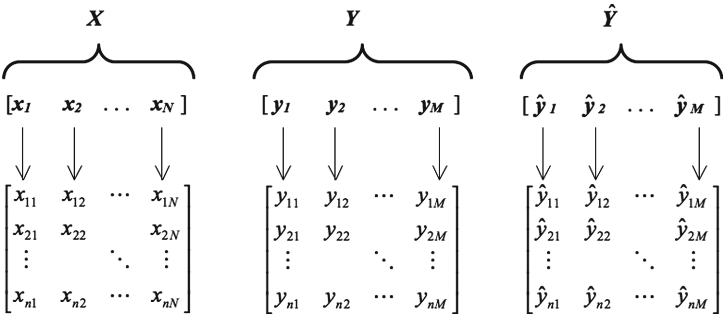


Fig. 4. Notation of input, output and estimation matrices, respectively.

In Fig. 4, n is the number of samples, N is the number of input signals and M is the number of signals at the output. These data would come from a history of data obtained from a period of observation of the process. Considering that a part of these n samples are used to represent the model memory ($\boldsymbol{X_{data}}$ and $\boldsymbol{Y_{data}}$), for AAKR models, each new query reading will provide a new estimation according to $\widehat{\boldsymbol{Y}}=\frac{\sum \boldsymbol{Y_{data}} \cdot \boldsymbol{X_{data}}, \boldsymbol{X}}{\boldsymbol{X_{data}}, \boldsymbol{X}}$.

Equation 1, which is a weighted sum of the $\boldsymbol{Y}$ of the memory, by means of the weights provided by the function $W(\boldsymbol{X_{data}}, \boldsymbol{X})$.

$$\widehat{\boldsymbol{Y}}=\frac{\sum \boldsymbol{Y_{data}} \cdot \boldsymbol{X_{data}}, \boldsymbol{X}}{\boldsymbol{X_{data}}, \boldsymbol{X}} \tag{1}$$

The function $W(\boldsymbol{X_{data}}, \boldsymbol{X})$ is the so-called kernel function, which can be different functions as long as they obey the following properties:

1. $W(\boldsymbol{X_{data}}, \boldsymbol{X})$ is not negative.
2. $W(\boldsymbol{X_{data}}, \boldsymbol{X})$ has the maximum value where $\boldsymbol{X_{data}} = \boldsymbol{X}$.
3. $W(\boldsymbol{X_{data}}, \boldsymbol{X})$ decreases monotonically with the distance $\|\boldsymbol{X_{data}} - \boldsymbol{X}\|$.
4. $W(\boldsymbol{X_{data}}, \boldsymbol{X}) = W(\|\boldsymbol{X_{data}}, \boldsymbol{X}\|)$ (radially symmetric).

Among the most known functions are: Uniform, Triangular, Epanechnikov, and Gaussian used in this work. For sensor estimation problem, each sample in memory $\boldsymbol{X_{data}}$ and $\boldsymbol{Y_{data}}$ correspond to steady state of the process under normal operating conditions (without failures in the process or measurement instruments). Therefore, this memory must undergo a rigorous selection process to avoid incorrect estimates regarding the process.

3.2 Combined Equipment Diagnosis

For the equipment diagnosis problem, memories must be added under fault conditions, and each memory must be previously classified so that the condition can be inferred for each sample. One way to accomplish this task is to use new columns in $\boldsymbol{Y_{data}}$. Each of these classification columns would correspond to a process state (i.e., normal, fault 1, fault 2). These columns would be filled with binary data, as dummy variables normally used in multivariate regressions, as illustrated in Fig. 5. Thus, using the same Eq. 1, estimates would be given for each of these classification columns with values between 0 and 1, with 1 being the highest possibility of belonging to the corresponding classification and 0 being the least possibility. Therefore, the equipment diagnostic methodology is more effective than the sensor estimation problem, because each estimation will be classified accordingly the nearest neighbour in the memory. The drawback of this method is clearly the increase of the memory and the computational complexity of memory based models that could be overcome with parametric models.

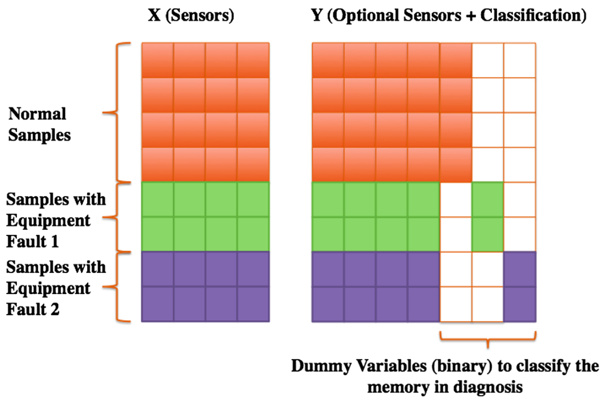


Fig. 5. Illustration of the memory matrix including the equipment diagnosis.

4 Results with the Test Data

For the valve example of Sect. 2.2 with the data in memory, each new generated data was diagnosed with the presented methodology. Figure 6 shows the mean simulation results in the 4 different cases also used to generate the memory: normal (case 1), opening 50% of the command (case 2), complete failure in the opening (case 3) and failure in the closure (case 4). The mean value of the estimates with optimized bandwidth presents the diagnostic result for each case, where it is possible to verify that the correct diagnosis is clearly identified in the 4 cases, proving the validity of the methodology.

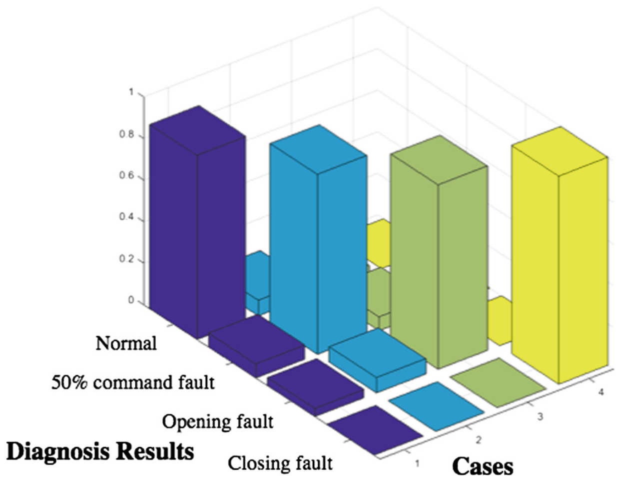


Fig. 6. Diagnosis results for different cases.

5 Conclusions

The presented methodology allows sensors and equipment fault detection simultaneously. It can be implemented increasing samples and variables in auto-associative models. The results were obtained using AAKR and data of oil flow process with simulated faults in a ball valve. But the increase of memory may lead to computational problems that could be overcome with different modelling techniques, which could be deeply analysed in future works. This methodology let also open possibilities of future evaluations in different applications, comparisons and combinations with different diagnostic methods.

References

Galotto, L., Brun, A.D.M., Godoy, R.B., Maciel, F.R., Pinto, J.O.P.: Data based tools for sensors continuous monitoring in industry applications. In: 2015 IEEE 24th International Symposium on Industrial Electronics (ISIE), pp. 600–605 IEEE (2015)

Galotto, L., Pinto, J.O.P., Ozpineci, B., Leite, L.C., Borges, L.E.: Sensor compensation in motor drives using Kernel regression. In: 2007 IEEE International Electric Machines & Drives Conference, pp. 229–234. IEEE (2006)

Hines, J.W., Garvey, D.: Development and application of fault detectability performance metrics for instrument calibration verification and anomaly detection. J. Pattern Recogn. Res. **1**(1), 2–15 (2006)

Kristjanpoller, F., Crespo, A., López-Campos, M., Viveros, P., Grubessich, T.: Reliability assessment methodology for multiproduct and flexible industrial process. In: Risk, Reliability and Safety: Innovating Theory and Practice. CRC Press (2016)

Markus, K., Ping, Z.: An embedded fault detection, isolation and accommodation system in a model predictive controller for an industrial benchmark process. Comput. Chem. Eng. **32**, 2966–2985 (1997)

Sharif, M.A., Grosvenor, R.I.: Fault diagnosis in industrial control valves and actuators. In: IMTC/98 Conference Proceedings. IEEE Instrumentation and Measurement Technology Conference. Where Instrumentation is Going, pp. 770–778. IEEE (1998)

Part II: Asset Economics and Costs

Intelligent Asset Management Budgeting and Investment Decisions for the Portfolio of Health Services Assets

Abe Nezamian(✉) and Ian Burns

Aurecon, Melbourne, Australia
{abe.nezamian, isn.burns}@Aurecongroup.com

Abstract. Assets are aging and budgets are tight, yet expectations are higher than ever. The challenge is that as the industry moves to more detailed levels of the approach, the budgets, the authority to take actions, and the performance metrics all start to become fragmented by organisational sub-units, asset service level requirements, regional requirements and compliance to regulatory and safety requirements. While potentially serving to motivate and provide realistic goals, it starts to undermine the value of a holistic asset management approach. If the decision making is not optimised across business asset portfolios and operational units, or even within business units, the value is diluted when all interactions between the assets and other factors are not adequately considered. The move from the industrial age to the information age has meant that the skyline of an organisation's physical assets has changed dramatically. This paper provides an overview of a digitalised and automated methodology tailored to address the challenges on this important theme that will shape the next decade and will drive the transformation to the new era of Asset Management. The methodology is developed through organising steps and activities by evaluating business processes, asset condition and asset degradation to generate asset investment plans based on actual and predictive asset performance long into the future. This allows authorities and health asset operational agencies to effectively optimise costs, risks, opportunities and performance.

The world is facing increasing urbanisation while, simultaneously, critical urban infrastructures have become a magnet for talent and a driver of economic growth. In Australia, economic modelling found that in 2011 (the Census base year), $187 billion of investment was derived from infrastructure. By 2031 this is projected to increase to $377 billion. Direct Economic Contribution demonstrates a pointer to areas of infrastructure demand and need. It presents an indication as to where reforms might lead to increased efficiency across our existing infrastructures (Australian Infrastructure Audit 2015). Australian capital cities contributed $854 billion to the economy in 2011, this is projected to grow by 90 percent to a contribution of $1.6 trillion in 2031. Australian major cities are vital economic engines – but unless action is taken, growing asset management cost and effectiveness threatens to cost Australians $53 billion by 2031 as the population increases to 30.5 million. Demand-side strategies, aimed at redistributing demand in time, space or mode, are equally important in making the most of existing capacity. They often present a cost-effective alternative to increasing

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 147–155, 2020.
https://doi.org/10.1007/978-3-030-48021-9_17

capacity, and have the potential to deliver better environmental outcomes, improved public health and more prosperous and liveable cities (Strategic Infrastructure 2014).

While health systems nowadays are challenged to reduce costs to match financial realities, the opportunity to reduce the cost burden associated with managing assets remains relatively neglected. About 95% of a hospital's clinical asset base consists of ventilators, infusion pumps, telemetry units and other mobile workhorse devices. The amount of such equipment per hospital bed has jumped 62% over the past 15 years, but average utilisation is only about 42%; a modern hospital's asset inventory is now both pricier than ever and often underutilised (Strategic Infrastructure 2014).

Many national health organisations recognised the need for application of new technology and organisational re-structure as part of a healthcare Asset Management (AM) transformation. These organisations such as the UK National Health Service (NHS) have a corporate responsibility to account for the stewardship of its publicly funded assets. This includes the provision, management and operation of an efficient and safe estate that supports clinical services and strategy (Health Building Note 00-08 Part A 2014). This corporate responsibility is carried by all accountable officers, directors with responsibility for estates and facilities and their equivalents, chairs, chief executive officers and non-executive board members. Together they have a responsibility to enact the principles set out in execution of the transformation program, provide leadership and work together to implement the necessary changes to provide a safe and efficient high-quality healthcare estate. These organisations have important contributions to make in delivering savings and reducing running costs. The transformation program should be developed to meet the challenges of funding these organisations such as the NHS in the future and will form part of the government's drive to increase the efficiency of the public-sector estate. Accordingly, this will be a significant step change in the way the health assets are to be managed (Health Building Note 00-08 Part A 2014). The presented methodology has been developed in part to address the challenges these organisations are facing to achieve the objectives of the Health AM transformation and required reforms.

Following increases in national demands on healthcare facilities and services, healthcare asset and facilities management has gradually matured to become an established research and development topic. Maintenance management and performance management can be reviewed in a wider context, and the main domains of healthcare are a hot topic of discussion. The five salient topics included in healthcare facilities asset management may be categorised in maintenance management, performance management, risk management, supply services management, and development (Shohet and Lavy 2004). These five core domains are interrelated, and can be integrated using information and communications technology, which provides the desired environment required for the challenging decision making and development prevalent in healthcare facilities asset management.

Technology has been incorporated by the critical asset operators for many years (Nelson 2018). However, the pace at which this adoption takes place is increasing rapidly as disruptive digital technologies have the potential to solve major AM challenges Therefore the critical health infrastructures will be transforming into a 'smart and digital assets' (Nelson 2018). Typically, it is in that stage that organisations recognise the need to understand the condition of their asset base. This condition information is paramount to understanding and building forward plans for ongoing asset maintenance and renewals, and understanding the financial implications of maximising the value of the asset base, in climates where scarce financial and human resource allocation is becoming more and more important to an organisation satisfying its business objectives.

The fact is that a small proportion of organisations have demonstrated that they have this capability, and the internal competency to achieve this. The asset performance model is connected to Artificial Intelligence and a Machine Learning Model for Case-Bases Reasoning decision making. This paper presents a digitised and automated methodology to develop Asset Management Plans (AMPs) based on condition of the asset, asset performance models of asset subclasses and critical components for intensive and critical infrastructure assets. This paper also provides an overview of the challenges and discusses various aspects of the asset condition data collection and assessment challenges representing risk to long-term integrity, level of service, safety and compliance. The paper highlights a structured, digitised and automated asset condition audit for assessment and maintaining the functional integrity and serviceability of the facility during the operating life using Artificial Intelligence.

1 Artificial Intelligence

In recent years the connectionist model (i.e. Artificial Neural Network (ANN)) based methodology for a Case-Based Reasoning (CBR) system design has become applicable for monitoring of asset performances. Special emphasis is laid on how the ANN can advance CBR technology by building an ANN-based CBR system, or integrating itself as a component within a CBR system. A simple ANN models proposed for constructing a CBR system and for solving some special issues involved in a CBR process is highlighted in this paper. The main characteristics of the model are developed based on each asset class performance requirement and degradation mechanisms, and the advantages and limitations of different models as per each asset class are discussed. Also, future research directions are outlined.

As a significant branch of Artificial Intelligence (AI), CBR has received more and more research attention. In the last few decades, CBR has grown from a quite new area to a subject of major influence. Much work has been dedicated to this topic, including its basic principle (Reisbeck and Schank 1989; Kolodner 1993), methodologies (Aamodt 1994; Leake 1996) and applications (Watson 1997). CBR systems have also been used to solve a wide range of problems. Examples of the applications of the CBR systems include medical diagnosis (Bichindaritz et al. 1998; Schmidt and Gierl 1998), time series prediction (Nakhaeizadeh 1993), product design (Main and Dillon 1999; Krampe and Lusti 1997) and planning (Arts and Rousu 1996), etc.

However, although CBR is simple in principle and has been successfully used in engineering problem-solving, it still lacks a generally theoretically sound framework. Consequently, most CBR systems could not complete their reasoning process and propose a solution to a given task, without intervention of domain experts or system managers (Watson 1997). Many CBR systems, for example, actually act as case retrieval and proposal systems, while case adaptation and case update are performed by human experts. As such, although all the asset subclasses share the same basic principle in common, it may be difficult to compare and contrast one asset class/subclass system with another, so that the performances of different asset class systems can be estimated and verified on a common benchmark based on the business strategy and service level requirements.

2 Case-Based Reasoning System

Classical knowledge or rule-based decision support systems draw conclusions by applying generalised rules, step-by-step, starting from scratch. Although successful in many application areas, such systems have met several problems in knowledge acquisition and system implementation. Inspired by the role of reminding in human reasoning, CBR systems have been proposed as an alternative to rule-based systems (prescriptive AM strategy), where knowledge or rule elicitation is a difficult or intractable process. This approach focuses on how to exploit asset performance, instead of rules, in problem-solving, and thus improving the performance of decision support systems and to be related to current condition of the asset. In brief, reasoning in CBR is based on experience or remembering, and historical performance of the asset considering the current condition as a base line.

In CBR, a primary knowledge, stored in memory, is not compiled from rules but a set of structured cases. These cases represent an experience or lesson that, under which asset operation environment and to what degradation mechanisms, some specific mitigation options have been derived in the past to achieve the goals of maintaining service levels and the effects after the solutions have been applied. When a new problem is presented to the system, a new solution is proposed by retrieving the most relevant old cases and adapting them to fit new situations, which is then introduced as a new case into the system in order to enrich the case base. Therefore, a case-based reasoner solves new problems by adapting mitigation options within an effective timeframe that were used to solve old problems. A successful use of CBR depends upon addressing issues of how to acquire, represent, index, retrieve and adapt existing historical performance cases. However, the most important operations involved in a primary reasoning process of CBR are case retrieval and case adaptation, together with case similarity assessment and case adaptation criteria. The solution, corresponding to these previous cases, is adapted to propose a solution for the given problem based on the adaptation criteria. It is obvious that the criteria for both case similarity assessment and case adaptation are inextricably linked, which means that the solution suggested is intimately connected to the similarities and differences between new and old cases. More precisely, such a statement may be described as

$$Suggested\,Solution = A\left[S\left\{C^{old}\right\}C^{New}\right] \tag{1}$$

where $\{C^{old}\}$ and C^{New} denote a set of old cases and a new case, respectively. $S(\cdot,\cdot)$ is the similarity assessment criterion, and A, an operator, is defined by the adaptation criteria.

Therefore, in essence, the CBR reasoning process is a pattern matching and classification process. Developing case similarity assessment criteria and case adaptation criteria is a central open challenge for CBR.

In a symbolic description model-based CBR system, the case adaptation process is often performed by a rule-based system (prescriptive AM program and based on vendor recommendation), which makes the CBR system design re-confront the knowledge acquisition problem for rule-based systems, and consequently, contradicts the important motivation for using CBR in problem-solving. As a result, many CBR systems currently work primarily as a proposal system, leaving the case adaptation to be undertaken by field experts or adoption of a reactive AM model. On the other hand, for a quantitative description model-based CBR system, the ANN development may become more flexible compared with symbolic description/qualitative model-based systems. This is because a lot of mathematical approaches and optimisation techniques are available for defining, synthesising and analysing the case similarity assessment and case adaptation criteria. In particular, the connectionist approach and several related ANN models have been suggested for the CBR system design. As is well known, ANN approaches have many appealing characteristics such as parallelism, robustness, adaptability and generalisation. Neural networks have the potential to provide some human characteristics of problem-solving that are difficult to describe and analyse using the logical approach of expert systems. More importantly, learning and problem-solving can be incorporated naturally and effectively by using a network.

3 Intelligent Asset Management Lifecycle Management Program

Figure 1 below gives a high-level outline of the proposed process in development of an intelligent asset lifecycle management program. This methodology is being continually refined as it is applied to asset lifecycle management of various asset classes where we have applied its fundamental principles to a number of hydrocarbons, production, manufacturing and health facilities lifecycle management in Australia and worldwide for major operators. It will bring value to the asset owners/operators through the ability to cost effectively develop various asset integrity management and risk mitigation plan options in application of Artificial Intelligence, CBR and ANN. while optimising their reliability, safety, accessibility, functionality, availability and maintainability of the asset.

The following considerations have been included in the development of an automated and digitalised AM Lifecycle management methodology to adopt artificial intelligence for asset systems, sub-systems and critical components.

- Risk Assessment
- Condition Evaluation and Relevant Threats, Bad Actor Identification
- Original Equipment Manufacturers (OEMs), Operational and Functional Obsolescence and Operational Vulnerabilities
- Degradation Mechanisms/Failure Modes/Reliability and Availability
- Asset Integrity and Maintenance Managements

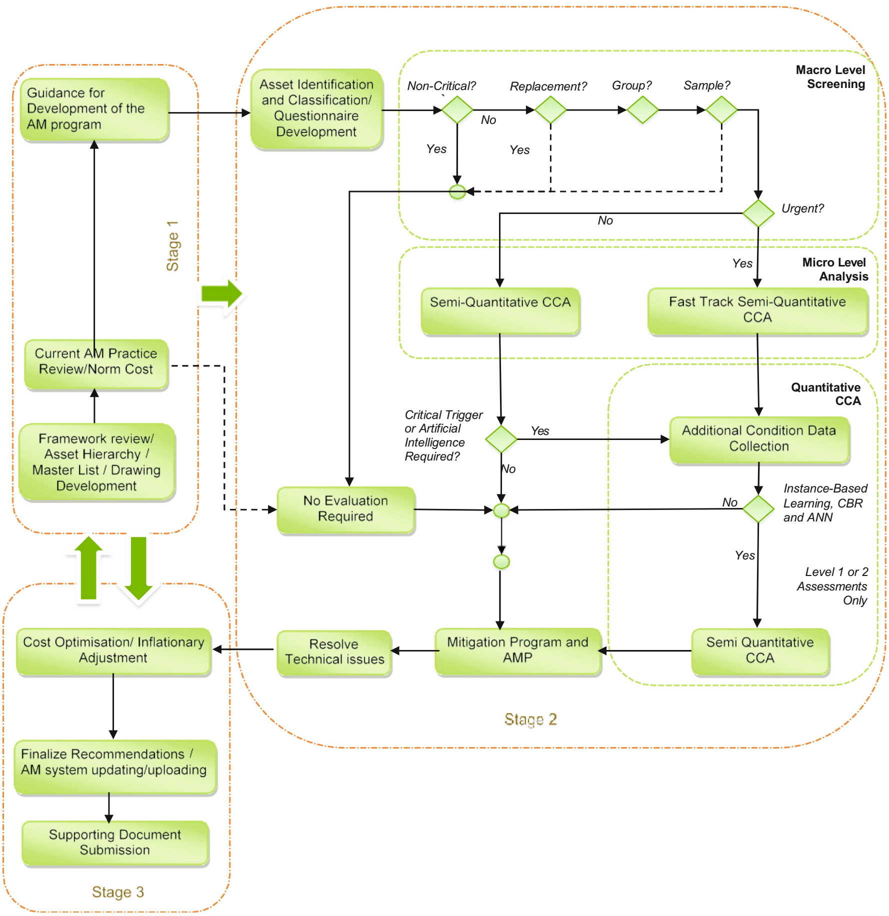


Fig. 1. Lifecycle management program methodology

- Historical performances/Functionality requirements/Current conditions

The consideration will have aimed at providing the desired outcome to the following organisation goals:

a) Fulfil the operational obligations with regard to AM costing parameters and program framework
b) Fulfil the regulatory requirements with regard to economic lifespans and Master list development
c) Fulfil the operation costing needs by evaluating Repair & Replacement proposals and verifying condition of assets
d) Identify key elements and address immediate actions
e) Finalise the AM program, considering annual increment value and OEM recommendations.

3.1 Data Collection (Functional/Design, Historical and Current)

One of the vital parts of this methodology is the efficient collection of the data required for the determination of the current condition and remnant life of the equipment. Design, operational and maintenance data and current condition data are important parameters for determination of the remnant life. The majority of this information can be collected from the Operation and Maintenance (O&M) team, maintenance records, integrity and engineering assessment reports, other inspection reports, fault history, site O&M personnel, equipment Original Equipment Manufacturers (OEMs) or vendors and site surveys. The maintenance records are used to identify the frequency/ occurrence of failures and maintenance work carried out on particular equipment in a chronological manner.

3.2 Asset Identification

After collection and collation of all functional, design, historical and current condition data, all assets sub-systems and critical components within the scope of assessment have to be identified, classified and listed. The asset identification will be conducted using a master asset register, or similar list, of equipment based on an asset owner's records. The asset hierarchy is produced and documented according to ISO 14224:2006 (2006). A 'Functional Hierarchy' is developed to align the relationship between maintainable components and each functional location within the asset. This describes the component structure of the asset item by assigning functional locations to each component. This hierarchy provides an overview of equipment that belongs together functionally, and shows the physical relationship between equipment, instruments, valves etc. This will also enable the application of an appropriate methodology.

3.3 Remnant Life and Current Condition Assessment

Qualitative, Semi-Quantitative or Quantitative factors are considered for the appropriate assets for Remnant Life Assessment (and thus lifecycle management program development) for each type of asset classes. For a small subset of equipment, the remaining life is determined using only qualitative factors and criteria where trend-able degradation mechanisms are not available (i.e. electrical equipment such as central processors or remote terminal units). In this case, obsolescence and availability of spares determined the continued use of this equipment and their remaining useful life. The assessment will be carried out based on a scoring system which is tied into the quality of data or the history of performance of the respective piece of equipment or asset. The first step in the determination of remnant life of a particular item is scoring the particular degradation mechanisms (using instance-Based Learning, CBR and ANN) and other remnant life parameters falling under:

- Design and Functionality Data
- Historical Data
- Current Condition Data

3.4 Asset Lifecycle Management Program Strategy

An asset lifecycle management program and associated recommendations will be developed for each asset based on inputs from the current condition and remnant life assessments. Option evaluation also considers the criticality of the equipment/system, machine learning, Health, Safety and Environment (HSE), operational disruptions, and support functions), nature of damage, costs associated with operations, and costs and effectiveness of maintenance, inspection and monitoring programs. During development of the lifecycle management or mitigation options, OEMs and Vendors can be approached to obtain further information about the respective equipment, such as spare parts availability, as part of baselining with prescriptive rules, the number of years it shall be available, level of technical support, potential for continued support, replacement options (e.g. new design) and formal declaration of obsolescence. These are critical in the asset life extension management. AMPs options may be assigned one of the following recommendations: RUN; REPAIR; REHABILITATE; RE-DESIGN/ MODIFY; RE-RATE; REPLACE or RETIRE.

4 Conclusions

An overview of a tailored digitalised and automated methodology was presented in this paper to address the challenges on this important theme that will shape the next decade and will drive the transformation to the new era of Asset Management. The organising steps and activities are articulated by evaluating business processes, asset condition and asset degradation to generate asset investment plans based on actual and predictive asset performance long into the future. This allows national authorities and health asset operational agencies to effectively optimise costs, risks, opportunities and performance. Further research need to develop A Task-based Knowledge Management from an organisational and a strategic perspective applying the principles in the first instance to other business units with the intention of producing a more structured decision-making process for board level debate and to maintain the organisational knowledge (Zanner et al. 2008).

References

Australian Infrastructure Audit: Australian Government (2015)

Aamodt, A.: Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun. **7**(1), 39–59 (1994)

Arts, R.J., Rousu, J.: Towards CBR for bioprocess planning. In: Advances in Case-Based Reasoning: Proceedings of the 3rd European Workshop on Case-Based Reasoning, pp. 16–27. Springer (1996)

Bichindaritz, I., Kansu, E., Sullivan, K.M.: Case-based reasoning in CARE-PARTNER: gathering evidence for evidence-based medical practice. In: Advances in Case-Based Reasoning: Proceedings of the 4th European Workshop on Case-Based Reasoning, pp. 334–345. Springer (1998)

Reisbeck, C.K., Schank, R.C.: Inside Case Based Reasoning. Lawrence Erlbaum, New York (1989)

Health Building Note 00-08 Part A: Strategic framework for the efficient management of healthcare estates and facilities, UK Government (2014)

Shohet, I.M., Lavy, S.: Healthcare facilities management: state of the art review. Facilities Manag. **22**(7/8), 210–220 (2004). https://doi.org/10.1108/02632770410547570. ISSN 0263-2772

ISO 14224:2006: Petroleum, petrochemical and natural gas industries – Collection and exchange of reliability and maintenance data for equipment (2006)

Kolodner, J.: Case-Based Reasoning. Morgan Kaufmann, Boston (1993)

Krampe, D., Lusti, M.: Case-based reasoning for information system design. In: Proceedings of the 2nd International Conference on Case-Based Reasoning, pp. 63–73 (1997)

Leake, D.B. (ed.): Case-Based Reasoning: Experiences, Lessons, and Future Directions. AAAI Press, MIT Press, Cambridge (1996)

Main, J., Dillon, T.S.: A hybrid case-based reasoner for footwear design. In: Case-Based Reasoning Research and Development: Proceedings of the 3rd International Conference on Case-Based Reasoning, pp. 497–509. Springer (1999)

Nakhaeizadeh, G.: Learning prediction of time series: a theoretical and empirical comparison of CBR with some other approaches. In: Topics in Case-Based Reasoning: Proceedings of the 1st European Workshop on Case-Based Reasoning, pp. 65–76. Springer (1993)

Nelson, B.: Connecting Remote Critical Assets. Remote Magazine, A Webcom Communications Publication (2018)

Schmidt, R., Gierl, L.: Experiences with prototype designs and retrieval methods in medical case-based reasoning systems. In: Advances in Case-Based Reasoning: Proceedings of the 4th European Workshop on Case-Based Reasoning, pp. 370–381. Springer (1998)

Strategic Infrastructure: Steps to Operate and Maintain Infrastructure Efficiently and Effectively. World Economic Forum (2014)

Mitchell, T.M.: Machine Learning. MIT Press/McGraw Hill, Cambridge (1997)

Watson, I.: Applying Case-Based Reasoning: Techniques for Enterprise Systems. Morgan Kaufmann, Burlington (1997)

Zanner, A., Burstein, F., Linger, H.: A task-based knowledge management case for addressing merger and acquisition risk. In: 19th Australasian Conference on Information Systems, Christchurch, New Zealand, 3–5 December 2008 (2008)

Impact of Circular Economy on Asset Management – Lifecycle Management Perspective

Jyri Hanski(✉) and Pasi Valkokari

VTT Technical Research Centre of Finland Ltd., Tampere, Finland
jyri.hanski@vtt.fi

Abstract. Sustainability of assets and processes is a major strategic focus area for many companies. To achieve the strategic goals of sustainability, several companies have begun to discuss how their businesses are supporting circular economy. Digital technologies and data are considered key enablers of circular economy. Circular economy solutions include e.g. minimizing the use of resources, closing resource loops and improving durability and lifetimes of assets. This paper discusses the impact of circular economy on asset management. The focus is especially on the different representations of lifecycles in production systems. We discuss the challenges and opportunities related to generic, circular and hierarchic production system lifecycles. We present main points to be considered when developing circular economy solutions from the perspective of asset and lifecycle management.

1 Introduction

Companies adopting asset management (AM) principles have realized that maintenance is not just a source of costs but also of value. Sustainability of assets and processes is a major strategic focus area for many companies. However, the focus of asset managers is often on carbon neutrality and being "green" instead of resource consumption (e.g. Stahel 2008). Efficient consumption of resources is crucial for reducing the greenhouse gas emissions and ensuring the sufficiency of natural resources. To achieve the strategic goals of sustainability, several companies have begun to discuss how their businesses are supporting circular economy (CE). CE is a system which creates value by minimizing waste, energy and the use of natural resources (Geissdoerfer et al. 2017).

CE solutions can be categorized to several archetypes (e.g. Sitra 2017). Product life extension pursues increasing the value from the invested resources, providing as long as possible useful life and maximizing profitability over lifecycle of assets.

Product as a service aims at retaining the ownership of an asset and offering it to customers as a service. The company offering the product has an incentive to optimise the utilization and lifecycle of the asset. Sharing platforms provide a means to connect asset owners with individuals or companies that would like to use them to boost asset profitability. In renewability category, renewable, recyclable or biodegradable inputs are used as substitutes for linear ones. Finally, resource efficiency and recycling aim at finding value in all material streams.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 156–163, 2020.
https://doi.org/10.1007/978-3-030-48021-9_18

The role of data and information technology in creating new business is increasing. Intelligent assets have the potential to change the way value is created and enable better resource efficiency for societies (Ellen MacArthur Foundation (EMF) 2016). EMF (2016) introduce three main forms of asset knowledge that enable value creation: location, condition and availability of assets. These knowledge types are important enablers for many CE solutions. Many CE solutions are supported by digital platforms and enhanced information management systems and practices. Information management systems ensure the availability, reliability and transparency of the solutions to the relevant actors and stakeholders. Nevertheless, the emphasis of CE literature is typically on material flows. There are, however, some exceptions. Information sharing is crucial part especially in sharing economy platforms and ecoindustrial parks (e.g. Ghisellini et al. 2015). Industrial ecology approach includes information flows in the analysis of industrial ecosystem and its environment (Erkman 1997). Data sharing is seen crucial in sharing economy based solutions, but the information and life cycle data management practices and challenges are rarely highlighted in CE. Therefore, our paper discusses the impact for lifecycle data management when CE principles are followed in industrial environments. The focus is especially on the different representations of production systems lifecycles and their effect on the CE solutions. We discuss the challenges and opportunities related to generic, circular and hierarchic production system lifecycles.

2 Circular Economy and Asset Management

CE aims at decoupling economic growth from resource consumption (Gregson et al. 2015). Strategies and solutions that extend the lifecycle and conserve value of materials - keeping durable materials in closed-loops by reuse and adaptation - offer much potential. Ness and Xing (2017) present an extensive list of principles and solutions associated with CE. These include closed-loop approaches, resource efficiency and productivity, resource effectiveness, optimization of use of goods and assets, asset management, longevity and durability, design for longevity and adaptability, sufficiency. sharing, waste management, GHG emissions, industrial economy and symbiosis, industrial metabolism, material flows, lifecycle assessment, input-output analysis and carbon neutral planning.

As highlighted by Ness and Xing (2017) CE and AM have many similarities. In the context of this study, we follow the definition of ISO 55000 (2014) for asset management "the coordinated activity of an organisation to realize value from assets". CE and AM both aim at the optimisation of resource or asset value over full lifecycle. As a main difference, CE aims at extending the use of assets and resources after the initial lifecycle in similar or different use situations to maximise the efficiency of resource usage. Additionally, the replacement of non-renewable with renewable materials is a goal of CE. AM considers multiple decision criteria including the economic perspective and risk when assessing solutions. Therefore, solutions that are viable from the CE perspective may not be viable for AM. From the perspective of AM, CE solutions may extend the lifetimes and increase the efficiency of assets, enable the recycling or reuse of components and equipment, offer new purposes for the old equipment, and cost savings and indirect quality improvements through better availability and quality of information (Hanski et al. 2016). AM threats introduced by CE solutions include complexity of the supply chain and information systems, management of new value elements and the assurance of quality in the new production ecosystem (Hanski et al. 2016).

Production system for producing petrol substitute from organic waste by Finnish energy company ST 1 is an example of CE solution that has an impact on AM. According to Valkokari et al. (2017) the solution poses various opportunities and threats for AM. Opportunities include e.g., increased delivery reliability due to several small production units, common spare part stock for fleet of production units, and accumulated fleet data that can be used to support the assurance of the machinery reliability. Threats include e.g., potential quality challenges for the end product due to heterogeneous raw material, durability of machinery and equipment is uncertain due to changing characteristics of raw material and lack of usage history data, and large number of actors in the value chain that increase the complexity of the production system and cause challenges for decision-making and leadership.

3 Circular Economy and Lifecycle Data Management

EN 13306 (2017) defines lifecycle as a "series of stages through which an item goes, from its conception to disposal". A typical lifecycle consists of acquisition, operation, maintenance, modernization, decommissioning and/or disposal phase (EN 13306 2017). Production systems can be represented by e.g., linear, circular or hierarchic lifecycles.

3.1 Generic Lifecycle

Product lifecycle management is a holistic approach for innovation, new product development and product information management from ideation to end of life. Traditionally different design for excellence approaches (DFX) illustrate the lifecycle in a linear way. For instance, IEC (2014) that gives guidelines for dependability optimization, presents the generic lifecycle as follows (see Fig. 1).

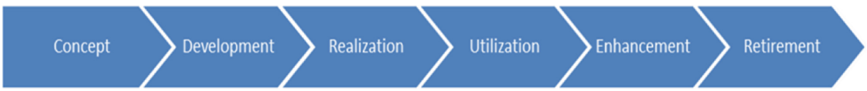


Fig. 1. Generic phases of lifecycle (IEC 2014).

This lifecycle model should be applicable for all items. It is very useful when different lifecycle management tasks and information required in these tasks are defined. The model supports very well the planning of lifecycle management activities when the item is relatively plain. From the AM perspective the utilisation and enhancement phases are usually essential.

3.2 Generic Circular Lifecycle

Generic circular lifecycles (Fig. 2) are generally used to describe CE solutions such as product life extension, product as a service, renewability, and resource efficiency and recycling. They are suitable for and utilized in e.g., describing the use of side streams, recycling of consumables and describing various AM tasks (maintenance, repair, reuse, remanufacturing and end-of-life activities). They focus predominantly on material flows. From lifecycle management perspective, circular lifecycles are more complex than linear. The extension of lifecycles, recycling and reuse requires feedback loops and connections to other production systems.

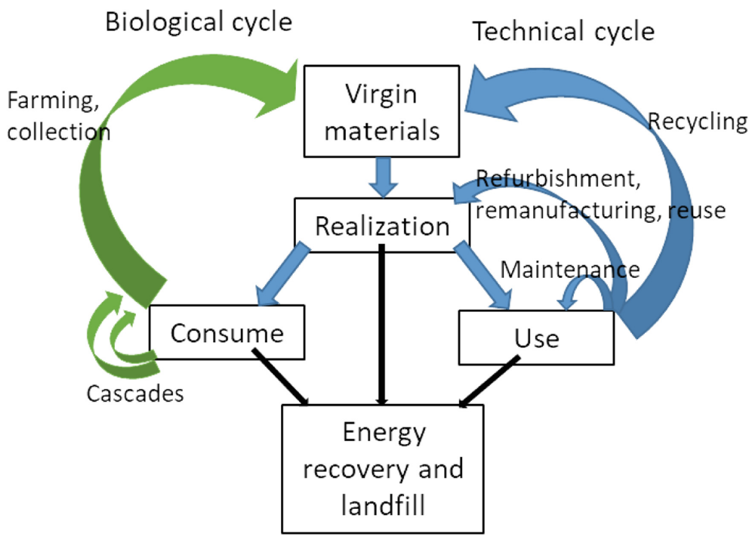


Fig. 2. Generic circular lifecycle (adapted from EMF 2013).

One practical exemplar of circular lifecycle model is illustrated in the Kingfisher's (2014), where new lifecycle of power tools is described (Fig. 3).

In this Kingfisher's case, power tools could be defined as an asset to be managed. What comes to the data that is required in planning of the AM activities, at least two focal challenges could be identified. First, besides selling tools to individual customers, model includes also a service aspect of renting tools. So to allocate effectively the AM actions for rented tools, information of customer specific usage profile (e.g. location, condition, operational environment) of the tools are needed. Second, in order to maximize harvested resources, especially location information of the assets (= tools) is important so that the incentive actions for clientele could be better addressed.

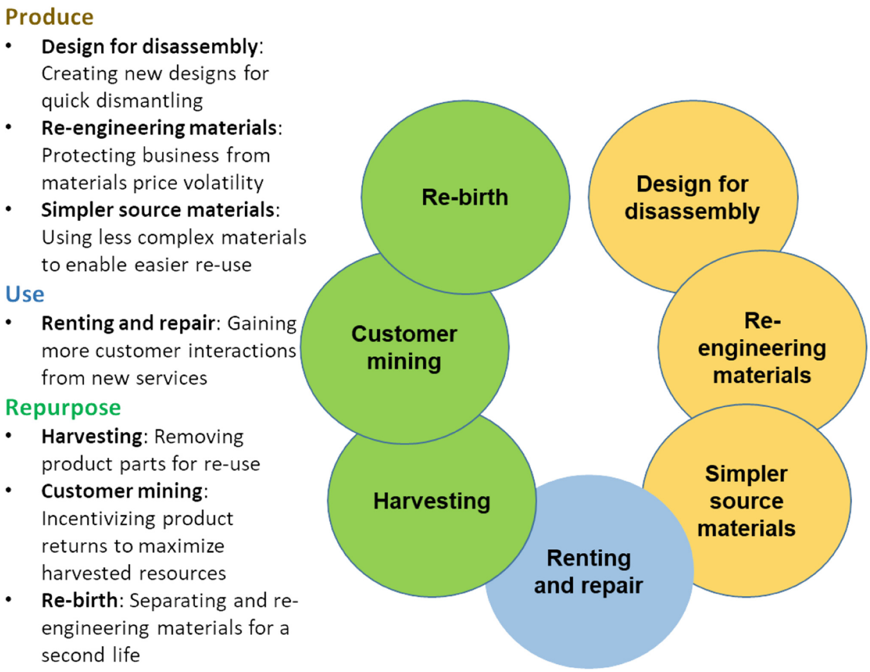


Fig. 3. Example of circular lifecycle (adapted from Kingfisher 2014).

3.3 Hierarchic Lifecycle of a Production System

A typical production system is built up on several different systems and subsystems, which are in turn built up from several components. All the elements in this hierarchical representation of production system have a lifecycle of their own. These lifecycles are of different lengths. A conceptual representation of system hierarchy is presented in Fig. 4. This hierarchic lifecycle model of production system might better

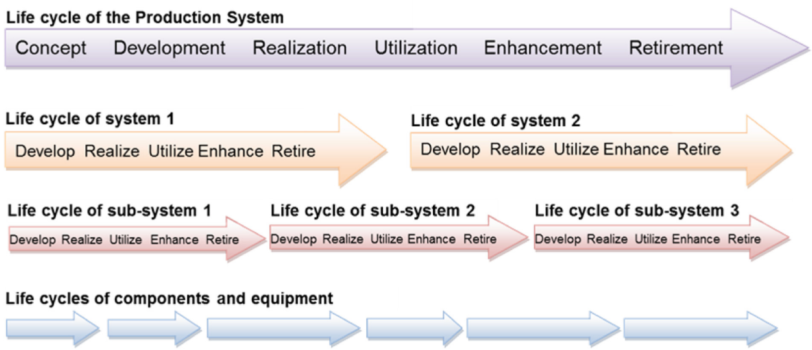


Fig. 4. Hierarchic lifecycle of a production system (adapted from Hanski et al. 2012).

support the planning AM activities of complex systems than generic lifecycle model. AM activities include effective control and governance of assets when pursuing value through managing risk and opportunity, in order to achieve the desired balance of cost, risk and performance throughout the system lifecycle.

When setting CE principles as strategic objectives, asset managers must take the production system and the supporting network of service providers into account. Furthermore, when data management plays a major role in the production system, the hierarchic representation of circular lifecycle can be used. Compared with linear solutions, CE solutions usually require more data on the usage, condition, criticality and characteristics of assets. When introducing CE into the equation, there is a need to consider the extension of lifecycles, reusing of systems, subsystems, components and materials, and planning their new lifecycles already at the system design stage. Systems, components, equipment and materials can also be utilized in other production systems and use cases.

Additionally, the optimisation of production system is more difficult as the second, third and further lifecycles should be taken into account. In order to align CE principles to AM activities there is also a need to bring in new kind of actors to prevailing value network. These include companies responsible for e.g. reverse logistics, recycling and reuse.

Figure 5 presents a schematic representation of hierarchic circular lifecycles in a production system. The figure aims to illustrate the increase of complexity when CE principles are applied to the AM activities of production system.

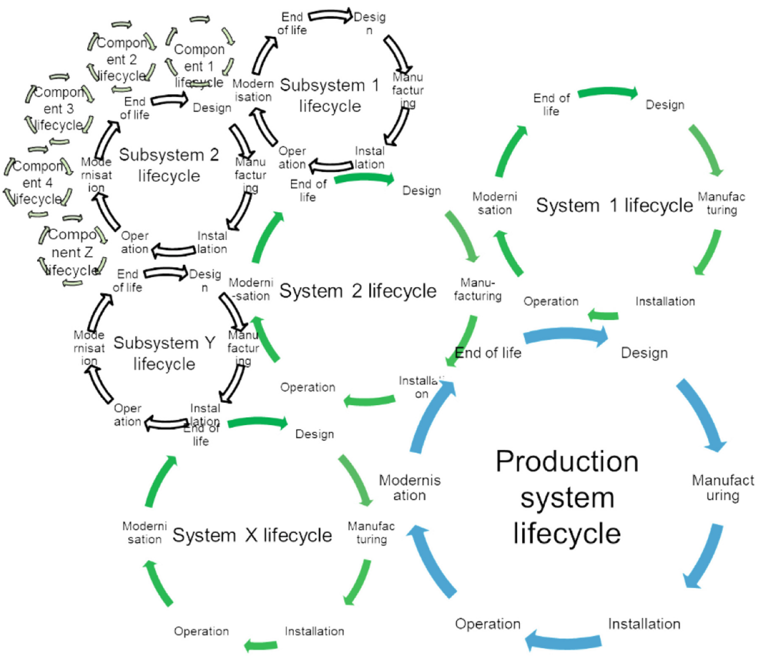


Fig. 5. Schematic representation of hierarchic circular lifecycles in a production system.

4 Discussion and Conclusions

CE solutions are mainly discussed from the perspective of opportunities in material flows and replacing fossil-based materials with renewables. AM requirements of CE models are typically not described (Hanski et al. 2016). Moving into CE solutions causes challenges and opportunities for lifecycle management in production systems. We argue that CE should also be considered from the perspective of production system required to deliver them. Following the CE principles in managing a production system poses several fundamental questions for the asset managers: how does the AM change? How are the CE principles followed from component and equipment level to the production system level? What are the new KPIs?

CE solutions are typically described using linear or circular presentations of lifecycles. In a production system, the situation is usually more complex. A production system consists of several hierarchic layers with varying information and management needs, information output and lifecycles. Generic representation of CE solutions does not sufficiently capture the challenges of AM in production systems. When current data management systems have been implemented, all the potential future use cases could not have been foreseen. For instance, reuse and recycling opportunities have not been sufficiently considered when the systems have been designed. Additionally, the data is often deficient (basic information about the asset, usage history, etc.), which is a challenge for digital services. Future research avenues include more indepth definition of CE especially in the context of AM. In addition, new digital technologies and applications such as machine learning and the digital twin pose opportunities and challenges for AM from the lifecycle management perspective.

References

Ellen MacArthur Foundation: Towards the Circular Economy: Economic and Business Rationale for an Accelerated Transition. Report (2013)

Ellen MacArthur Foundation: Intelligent Assets: Unlocking the Circular Economy Potential. Ellen MacArthur Foundation, pp. 1–25 (2016)

Erkman, S.: Industrial ecology: an historical view. J. Clean. Prod. **5**(1–2), 1–10 (1997)

Geissdoerfer, M., Savaget, P., Bocken, N.M.P., Hultink, E.J.: The circular economy – a new sustainability paradigm? J. Clean. Prod. **143**, 757–768 (2017)

Ghisellini, P., Cialani, C., Ulgiati, S.: A review on circular economy: the expected transition to a balanced interplay of environmental and economic systems. J. Clean. Prod. **114**, 11–32 (2016)

Gregson, N., Crang, M., Fuller, S., Holmes, H.: Interrogating the circular economy: the moral economy of resource recovery in the EU. Econ. Soc. **44**(2), 218–243 (2015)

Hanski, J., Valkokari, P., Kortelainen, H., Ahonen, T.: Circular economy models - opportunities and threats for asset management. In: Proceedings of Maintenance Performance Measurement and Management Conference, MPMM 2016, Luleå, Sweden, 28 November 2016 (2016)

Hanski, J., Kunttu, S., Räikkönen, M., Reunanen, M.: Development of knowledge-intensive product-service systems, p. 21. Report. VTT Technology (2012)

IEC: IEC 60300-1: Dependability management – Part 1: Guidance for management and application. International Electrotechnical Commission, Geneva (2014)

ISO: BS ISO 55000-2014 - Asset Management - Overview, principles and terminology (2014)

Kingfisher: The business opportunity of closed loop innovation (2014). https://www.kingfisher.com/sustainability/files/downloads/kingfisher_closed_loop_innovation.pdf. Accessed 16 Mar 2018

Ness, D.A., Xing, K.: Toward a resource-efficient built environment: a literature review and conceptual model. J. Ind. Ecol. **21**(3), 572–592 (2017)

SFS EN 13306:2017: Maintenance. Maintenance terminology. European Standard (2017)

Sitra: The most interesting companies in the circular economy in Finland (2017). https://www.sitra.fi/en/projects/interesting-companies-circular-economy-finland/. Accessed 3 Jan 2018

Stahel, W.R.: Global climate change in the wider context of sustainability. Geneva Pap. Risk Insur. Issues Pract. **33**(3), 507–529 (2008)

Valkokari, P., Hanski, J., Ahonen, T.: Impact of maintenance on circular economy. In: Proceedings of 2nd Annual SMACC Research Seminar, Tampere, Finland, 7 November 2017 (2017)

What Is the Economic Value of Environmental Goods - And How Can This Value Be Incorporated in Public Decision-Making?

Kari Ella Read[1(✉)] and Marika Karras[2(✉)]

[1] WSP Norge AS, Oslo, Norway
kari.ella.read@wsp.com
[2] Sweco Environment AB, Malmö, Sweden
marika.karras@sweco.se

Abstract. Nature's ecosystem and its services create the foundation of our society's existence. Yet, services such as fresh air and water, fertile soil and flow regulation have upon recent been treated as unlimited resources. The consequences can be seen in the rapid loss of biodiversity and issues related to climate change. Although environmental issues have got more attention recently, it can be challenging for decision-makers to assess what measures to choose, when and where to implement these, and how to distribute financial responsibility among stakeholders. One reason for this is suggested being that many environmental assets lack a direct monetary value.

In order to study the significance of including the economic value of environmental goods in public decision-making, Marika Karras and Kari Ella Read have in their master's thesis estimated the economic value of ecosystem services and incorporated these values in a cost-benefit analysis (CBA). The CBA was performed on measures preventing floods in urban areas through a solution called Sustainable Drainage System (SuDS), which is a solution based on mimicking the flow control found in nature. The cost aspect of the CBA was based on the investment and operational cost of these solutions, while the benefits constituted of the economic value of the ecosystem services that the system provides. These services include flood risk reduction, water and air purification properties, noise reduction, local climate regulation and an increase in recreational values.

The result of the study indicates that the suggested sustainable drainage system could be justified economically, but only by including both the monetary value of the ecosystem services in the CBA and using a discount rate of 1,4% recommended by the Stern-report (Stern 2006).

1 Introduction

In recent years, the focus on global sustainability and environmental protection seem to have increased in the developed world. Also, on a local level, the calls for investments to secure natural resources and measures to reduce the consequence of a changing climate appears to have increased. However, assessing what is a reasonable price to pay for environmental measures can be challenging, as well as allocating the financial

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 164–173, 2020.
https://doi.org/10.1007/978-3-030-48021-9_19

responsibility. A suggested reason for this is that the increase or reduction of environmental values traditionally has, if all, been communicated in other scales than monetary. And, in a society where economic measures often are the ruling criteria, values with no direct economic value can easily be underestimated or overlooked.

In areas concerning health or safety issues, valuation methods for estimating socioeconomic effects are more commonly used to facilitate decision-making processes. The methods for estimating the monetary value of utilities without direct economic values are, in other words, not new but rarely used related to environmental values.

In this paper, we present methods for monetary valuation of non-market values commonly used in other areas and present a case study where these methods are applied to environmental goods. The paper is based on the master thesis, *Cost-benefit analysis of sustainable drainage system as flood risk reduction measure in an urban area: Focusing on monetary valuation of ecosystem services*[1], conducted by Marika Karras and Kari Ella Read at The Faculty of Engineering at Lund University (2016). The aim of the paper, as well as the master's thesis, is to discuss how incorporation of environmental values in a cost-benefit analysis can affect the result, and further assess if this can contribute to public decision-making processes regarding environmental issues.

2 Monetary Valuation Methods

According to classic economic theory, the concept of supply and demand create the foundation of the market and a commodity's price on the market is supposed to reflect its value (Pihl 2014). The reasoning behind this theory is based on several assumptions, like the market must be efficient and people acting on the market must make rational decisions (Pihl 2014).

Some argue that the market value, also called the direct value method, can be used to reflect the value of some environmental goods such as food, water and raw materials (De Groot et al. 2012). However, if you value a tree by the price of timber you miss values of other attributes like the tree's ability to store carbon, reduce erosion and its drainage properties. Also, in a changing climate, some of the natural processes that make up the foundation of commodities on the market will change, which eventually will affect the price. Though, when this occurs, the changes might be irreversible and needs to be replaced with technical solutions that can become more expensive and, or, less effective than the original alternative. So, by capturing the value of benefits provided by nature in advance, this can hopefully prevent such a scenario from occurring in the first place.

To estimate the value of utilities without a market price, there are several methods to be obtained.

[1] Ecosystem services can be described as "the direct and indirect contributions of ecosystems to human wellbeing" (TEEB 2010b).

2.1 The Revealed Preference Approach

The revealed preference approaches include a cluster of methods estimating the values of certain amenities by studying people's behaviours on markets that are closely linked to the values we wish to evaluate. The Revealed Preference Method involve descriptive studies and exclude therefore problems of certain biases which otherwise can be a problem in valuation studies (The Swedish Environmental Protection Agency 2012).

Willingness to pay (WTP) is essential to this method as it gives an indication of the perceived value these services provide (Mattson 2014) (De Groot et al. 2012). Based on people's actual behaviour on the market they demonstrate how an attribute affects their willingness to pay. The methods in this cluster include The Replacement-/Substitute Cost Method, The Damage Cost Avoided Method, Production Function-Based Approaches, The Hedonic Pricing Method and The Travel Cost Method.

2.2 The Stated Preference Approach

The Stated Preference Approach involves asking people about their preferences, willingness to pay, through surveys or interviews (Mattson 2014). Within the group of stated preference methods, exists a variety of strategies for getting a broad and as accurate as possible picture of people's opinion. The most commonly used are The Contingent Valuation Method, Choice Modelling, and Group Valuation.

2.3 Benefit Transfer

Due to limited resources within a project, performing valuation studies can be difficult. In such a case, one possibility is to estimate the economic value by transferring values from other studies completed for a similar location or issue (TEEB 2010). Moreover, for values to apply to a given study, adjustments usually must be made, like adjusting to the specific location, the society in question, and time of execution (inflation).

3 Case Study of Economic Valuation of Ecosystem Services

Traditionally, the surface runoff caused by rainfall has been drained from streets and buildings through an underground pipeline system, leading straight to a recipient or sewage plant. As the areas consisting of impermeable surfaces and number of people in cities have increased, the traditional pipelines often fail to transport the excess water due to capacity overload, further resulting in floods. (The Swedish Water & Wastewater Association 2011). The consequence can be severe economic damage to people's properties and vital infrastructure.

Another contributing factor to the frequency of floods is the excepted changed precipitation patterns due to climate change, which for the Nordic countries is expected to result in an increased intensity of rainfall (IPCC 2007).

In order to deal with these changes, various solutions are available, one being an addition to the traditional drainage system called a sustainable drainage system (SuDS). The solution, which aims to prevent floods by mimicking the flow control found in nature, is becoming increasingly more popular as a contribution to climate change strategies in cities around the world.

The sustainable drainage system consists of meandering dikes, ponds, greenery and safe flooding sites within urban areas, which allows the rainwater to be delayed on its way to the sewage system. As the water is delayed, the openness of the solution means that some of the water will evaporate and infiltrate into the ground, which in turn reduce the amount of water needed to be handled by the existing pipeline. Figure 1 below illustrates a comparison of retention time in a natural area, which represents the optimal effect of a sustainable drainage system, and an urban area with many impermeable surfaces.

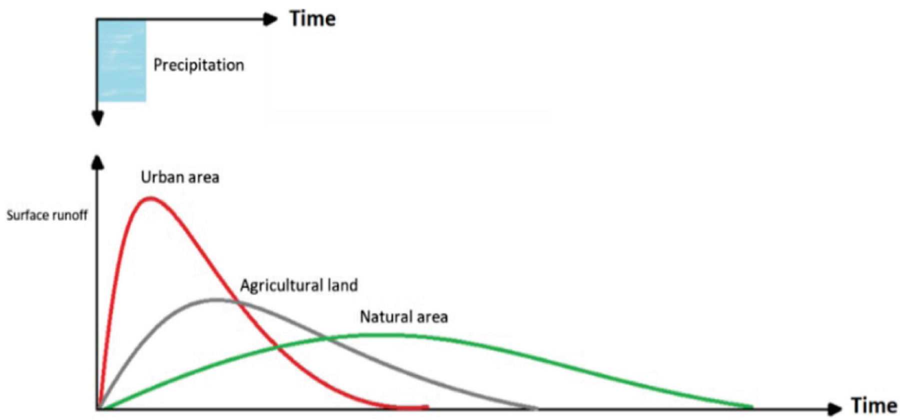


Fig. 1. Illustration of retention time in the various environment. Limitation in storm drains and low retention in urban areas cause floods, whereas the agricultural land and natural areas have a low flow of water due to infiltration, absorption by plants and evaporation (The Swedish Water & Wastewater Association 2011)

A study of the effect of sustainable drainage systems indicates that these solutions can reduce the risk of floods significantly (Theland 2015), but also provide benefits like recreational value, local climate regulation and biological cleansing of water, benefits also known as ecosystem service (Block and Bokalders 2014). This means that the SuDS, unlike conventional drainage solutions, provide benefits continuously regardless if it rains or not.

3.1 Research and Applied Method

The case study involves the implementation of a sustainable drainage system (SuDS) in the flood-prone area Soederkulla in Malmö, Sweden. Based on previous flooding events the amount of stormwater needed to be handled within the area was estimated, and from this various SuDS-solutions was suggested, including rain gardens, ponds, dry ponds, dikes, plants, green roofs, and constructed flooding sites.

The cost aspect of the cost-benefit analysis (CBA) constituted of the investment cost (implementation and resources) of the suggested SuDS, together with the annual operative cost of maintaining the efficiency of the system.

The benefit aspect of the CBA included the increase of ecosystem services to the area due to the proposed measures. The effect of these was quantified and further evaluated in economic terms partly by applying economic valuation methods for commodities with no direct market value, and partly by benefit transfer from other valuation studies (See Table 1 for more details).

Table 1. Summary of methods used to evaluate the value of various ecosystem services provided by the sustainable drainage system.

Ecosystem service	Description and applied monetarizing method
Flood control	The economic benefit in the form of flood control provided by the suggested system was estimated based on the Damage Cost Avoided Method. By using site-specific data collected from three large flooding events in the area in 2007, 2010 and 2014, each event was classified by a specific reoccurrence time based on a Dahlström and Hernebring-model (Dahlström 2010; Hernebring 2015). This, combined with information on the pay-out from the four largest insurance companies in southern parts of Sweden at each event, gave an estimate of the annual risk reduction in economic terms
Recreational value	The sustainable drainage system causes an increase of greenery in the area, which studies have indicated to have a connection with an increased recreational value (The Swedish Society for Nature Conservation 2011). The recreational value provided by the solutions was estimated by benefit transfer from several hedonic pricing studies studying the effect of urban environmental/green attributes influence on housing prices (The Danish Business Authority 2014) (Zhou et al. 2012). Based on the increase in green attributes contributed by the solutions suggested, and the quality of these, the added value of the recreational value was estimated
Noise reduction	The value of the SuDS effect on noise reduction was estimated based on template values for reductive noise measures recommended by The Swedish Transport Administration, ASEK (2016). The expected decrease in traffic noise resulting from the suggested measures was based on studies of green areas impact on perceived noise (The HOSANNA project 2013). Combined, this gave an estimate of the annual economic benefit due to the noise reduction in the area
Water purification	Drainage water contains a significant amount of pollutions like heavy metals and oil components (The Swedish Water & Wastewater Association 2016). Due to biological purification properties in parts of the SuDS, these pollutions can be reduced significantly (The Swedish Water & Wastewater Association 2011). Based on the area of the system containing these properties, the expected annual reduction of

(*continued*)

Table 1. (*continued*)

Ecosystem service	Description and applied monetarizing method
	pollution was estimated. The value of this reduction was based on the cost of water purification per cubic meter water in treatment plants in Sweden
Local climate regulation	As green areas and trees help to regulate temperatures and humidity, such areas have long been used to stabilise the local climate in cities (Clark et al. 2008)
	By calculating the isolation effect of green roofs, the reduced energy demand resulting from this was used to estimate the value of the benefit "local climate regulation"
Air purification	Due to plants ability to collect and store pollutions, the green solutions cause an air purification effect (Nowak et al. 2014). Vegetation also has a diluting effect on pollutants as trees and shrubs cause turbulence around the branches further causing a decreased concentration of pollution close to the ground (Johansson 2014)
	The air purification effect caused by the SuDSs was estimated based on the total volume of trees and shrubs suggested as a part of the SuDS, combined benefit transfer from various effect studies on the topic. The value of the reduction of pollutions was evaluated based on The Swedish Transport Administration's template value, ASEK, for reduction of air pollutants related to reduced health issues caused by lung and heart conditions in cities

All parameters used in the analysis were declared in probability distributions. This allows for Monte Carlo simulations in the decision-tool @Risk, which facilitates a transparent presentation of the uncertainties in the calculations.

The expected lifetime of the sustainable drainage system was estimated to be 100 years, and two different discount rates were used, 3,5% recommended by the recommended by The Swedish Transport Administration, ASEK 2016, and 1,4% recommended in the Stern-report for climate adaption measures (Stern 2006).

4 Results

The table below presents the four simulated scenarios modelled in the case study. The upper left shows the result of the discounted cost and benefits over a 100 years' time span, with a discount rate of 1,4%, where all ecosystem services were included. The table beneath shows a similar scenario, but when only the ecosystem service "flood control" was included.

The upper right table shows the result of a scenario where all ecosystem services were included but where a discount rate of 3,5% was used. The table below shows a scenario where only flood control was included. All values are presented in a million Swedish kronor, 2016 currency.

Discount Rate 1,4 % (all ecosystem services included)

	Value	Unit
Benefits discounted	101,4	MSEK
Cost discounted	77,30	MSEK
Net present value (NPV)	24,13	MSEK
NPV confidence level 5 %	*-9,80*	*MSEK*
NPV confidence level 95 %	*63,14*	*MSEK*

Discount Rate 3,5 % (all ecosystem services included)

	Value	Unit
Benefits discounted	53,19	MSEK
Cost discounted	61,67	MSEK
Net present value (NPV)	-26,90	MSEK
NPV confidence level 5 %	*-8,48*	*MSEK*
NPV confidence level 95 %	*12,42*	*MSEK*

Discount Rate 1,4 % (only flood control included)

	Value	Unit
Benefits discounted	9,31	MSEK
Cost discounted	77,30	MSEK
Net present value (NPV)	- 67,98	MSEK
NPV confidence level 5 %	*-76,44*	*MSEK*
NPV confidence level 95 %	*-59,72*	*MSEK*

Discount Rate 3,5 % (only flood control included)

	Value	Unit
Benefits discounted	4,88	MSEK
Cost discounted	61,67	MSEK
Net present value (NPV)	-63,12	MSEK
NPV confidence level 5 %	*-26,91*	*MSEK*
NPV confidence level 95 %	*-50,77*	*MSEK*

The result of the CBA gave a positive net present value only in the case where the economic value of all ecosystem services was included, and a discount rate of 1,4% was used (recommended by the Stern-Report for climate change adaption measures (Stern 2006)).

5 Discussion

Despite a relatively high investment and maintenance cost of the sustainable drainage system, the case study indicates that the system in the specific area was socio-economically profitable, but only if the value of ecosystem services was included in the CBA, and a discount rate of 1,4% was used.

The case study exemplifies how the economic evaluation of the environmental goods, here presented as ecosystem services, can enable incorporation of such values in traditional decision-making strategies, and how this can change the output from being seemingly socioeconomically unprofitable to profitable. The result of the case study also shows how significant the choice of discount rate is for the outcome of the CBA, and that incorporation of the monetary value of environmental goods was not on its own enough to make the system profitable. However, by both including environmental values in monetary terms and using an "environment-friendly" discount rate in the public decision-making processes, this may help to optimise the resource allocation and hopefully secure these goods for the next generation.

However, there are several issues and uncertainties related to these methods that need to be addressed.

Limitations of the Scientific Method

Applying monetary valuation methods to environmental goods involves several uncertainties that are important to consider. The lack of sufficient information concerning dependencies between various ecosystems, their services, and their connection to the welfare of society are among such uncertainties.

There is also a risk that the suggested solutions will not provide the expected value, and if they do, it is possible that the value will not be an annual contribution as assumed in the case study.

The sustainable drainage system also demands more land than a conventional alternative, which in cities is related to high alternative costs that are not included in the study. The system is yet a more flexible solution compared to a conventional system as the SuDS is located above ground, which will allow for adjustments according to demand in contrast to the conventional solutions placed underground, where the dimension is more or less fixed. This flexibility also has an economic value that is not included in the case study but would increase the net present value of the SuDS.

The methods used for the economic valuation of environmental goods are also related to several uncertainties. The revealed preference approach is for instance related to uncertainty in terms of whether or not we are measuring what we think are measuring. The stated preferences, on the other hand, are often criticised as people tend to be affected by various biases when participating in such studies.

The complexity of environmental issues also complicates the use of such studies, for example in making an optimal allocation of financial responsibility. This, partly because the consequence and cause of environmental issues often are dislocated in time and location, and since many of the environmental issues, like flooding, are related to an accumulated effect of many "smaller" wrongdoings.

6 Conclusion

The result of the case study indicates that the monetary evaluation of environmental goods is an applicable method and could bring added value to decision-making processes. An evaluation of environmental goods in economic terms might also help to increase the awareness of the relationship between the welfare of societies and environmental goods, and thereby highlight the connection between environmental issues on a global and local level. However, there is still an information gap related to the connectivity between species, the ecosystems they interact within, and lastly their effect on us humans that are essential to consider when dealing with complex and fragile environments. More studies on the topic are needed in order to verify the cause-effect relationships and valuation methodology. This lack of sufficient information should be handled with great care to secure these assets for future generations. Wrongful use of such methods without regarding the uncertainties can in the worst case lead to a justification of decisions that can cause emergent and irreversible consequences to environmental assets.

References

Block, M., Bokalders, V.: Urbana ekosystemtjänster: Låt naturen göra jobbet. C/O City, Stockholm (2014)

Brännlund, R., Kriström, B.: Miljöekonomi. Studentlitteratur, Lund (2012)

Clark, C., Adriaens, P., Talbot, F.: Green roof valuation: a probabilistic economic analysis of environmental benefits. Environ. Sci. Technol. **42**(6), 2155–2161 (2008)

Dahlström, B.: Regnintensitet - en molnfysisk betraktelse. The Swedish Water & Wastewater Association, Stockholm (2010)

De Groot, R., et al.: Global estimates of the value of ecosystems and their services in monetary units. Ecosyst. Serv. **1**, 50–61 (2012)

Hernebring, C.: The cloudburst in Southwestern Scania 2014-08-31: with focus on consequences and in relation to rainfall statistic in Malmö. VATTEN – J. Water Manag. Res. **71**, 85–99 (2015)

Hultkrantz, L., Nilsson, J.: Samhällsekonomisk analys, en introduktion till mikroekonomi. SNS Förlag, Falkenberg (2008)

IPCC: Summary of policymakers. In: Climate Change 2007: Impacts, Adaptation and Vulnerability Contribution of Working Group—To the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge (2007)

Johansson, E.: Utvärdering för ramverk av tätortsnära ekosystemtjänster. Department of Environmental and Energy Systems Studies, The Faulty of Engineering, Lund University, Lund (2014)

Karras, M., Read, K.E.: Cost-benefit analysis of sustainable drainage system as risk reduction measure against floods in urban areas – Focusing on monetary valuation of ecosystem services. Division of Risk Management and Societal Safety, Faculty of Engineering, Lund University, Lund (2016)

Mattson, B.: Kostnads-nyttoanalys. The Swedish Civil Protection Agency, Stockholm (2014)

Nowak, D., Hirabayashi, S., Bodine, A., Greenfield, E.: Tree and forest effects on air quality and human health in the United States. Environ. Pollut. **193**, 119–129 (2014)

Pihl, H.: Miljöekonomi för hållbar utveckling, 5 uppl. red. Studentlitteratur, Lund (2014)

Stern, N.: The Economics of Climate Change - The Stern Review. Cabinet Office HM Treasury. Cambridge University Press, Cambridge (2006)

TEEB: Glossary of terms (2010b). http://www.teebweb.org/resources/glossary-of-terms/. Funnet 26 July 2018

TEEB: TEEB for Local and Regional Policy Makers., s.l. The Economics of Ecosystems and Biodiversity (2010)

The Danish Business Authority. Gevinster ved investeringer i byliv og bykvalitet, s.l. The Danish Business Authority (2014)

The HOSANNA project: Novel solutions for quieter and greener cities. European Union Seventh Framework Programme (FP7/2007-2013), Gothenburg (2013)

The Swedish Environmental Protection Agency: Summarized information about ecosystem services. The Swedish Environmental Protection Agency, Stockholm (2012)

The Swedish Society for Nature Conservation: Räkna med ekosystemtjänster - Underlag för att integrera miljövärden i den, s.l. The Swedish Society for Nature Conservation (2011)

The Swedish Transport Administration: Analysmetod och samhällsekonomiska kalkylvärden för transportsektorn: ASEK 6.0. The Swedish Transport Administration, Stockholm (2016)

The Swedish Water & Wastewater Association: P105 Hållbar dag- och dränvattenhantering - råd vid planering och utförande. The Swedish Water & Wastewater Association, Stockholm (2011)

The Swedish Water & Wastewater Association: P110 Avledning av dag-, drän- och spillvatten – funktionskrav, hydraulisk dimensionering och utformning av allmänna avloppssystem. The Swedish Water & Wastewater Association, Stockholm (2016)

Theland, J.: Funktionen av hållbara dagvattenlösningar och gröna ytor vid extrema regn. Avdelning för teknisk vattenresurslära, Lunds tekniska högskola, Lunds universitet, Lund (2015)

Zhou, Q., Panduro, T., Thorsen, B., Arnbjerg-Nielsen, K.: Adaption to extreme rainfall with open urban drainage system: an integrated hydrological cost-benefit analysis. Environ. Manag. **51**, 586–601 (2012)

Engineering and Financial Realities of Water Infrastructure in South Africa

Cornelius Ruiters(✉) and Joe Amadi-Echendu

Department of Engineering and Technology Management, Faculty of Engineering, Built Environment and Information Management, University of Pretoria, Pretoria, South Africa
corneliusruiters07@gmail.com, joe.amadi-echendu@up.ac.za

Abstract. The Government of South Africa has been the main provider of infrastructure, particularly in the water sector. Government administration and institutional structures continue to shape, influence and impose unique complexities and constraints on infrastructure investment. The country experiences a serious backlog in investments for the development and management of water infrastructure. This under-investment was estimated at more than US$50 billion. South Africa needs to find a solution to this backlog by putting in place new or fresh institutional structures and funding models for effective strategic provision of water infrastructure. Primary and secondary data were collected for the development of investment models for water infrastructure in South Africa. The research identified several new funding and institutional models for financing water infrastructure, e.g. private sector, water charges and tariffs, public-private partnerships (PPP), which are starting to get traction in the water sector value chain of South Africa.

Keywords: Engineering · Financing · Funding · Investments · Institutions · Models · Realities · South Africa · Water infrastructure

1 Introduction

Much of South Africa's water infrastructure is at a crossroads (DWA 2008, 2013; DBSA 2009; WB 2010; NPC 2013; SAICE 2017). Following decades of under-investment, vital elements of the nations' infrastructure are in serious disrepair, if not in a crisis. South Africa's infrastructure – investment sunk in water infrastructure – is struggling to cope with the cumulative demands of South Africa's economic and social growth and the vast new trade and investment opportunities emerging. There is a serious backlog in infrastructure investment, especially in water, estimated conservatively at US$50 billion, which requires immediate attention. Institutional structures – those of public entities – which have served South Africa well in decades past now appear unable, and ill-equipped, to cope with the nation's present infrastructure planning and delivery challenge. Even with large increases in tax revenues and aggressive "surpluses" of government public entities, the water infrastructure investment required to meet South Africa's present and future needs has not materialized. Simultaneously, large capital resources are accumulating in the private sector,

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 174–182, 2020.
https://doi.org/10.1007/978-3-030-48021-9_20

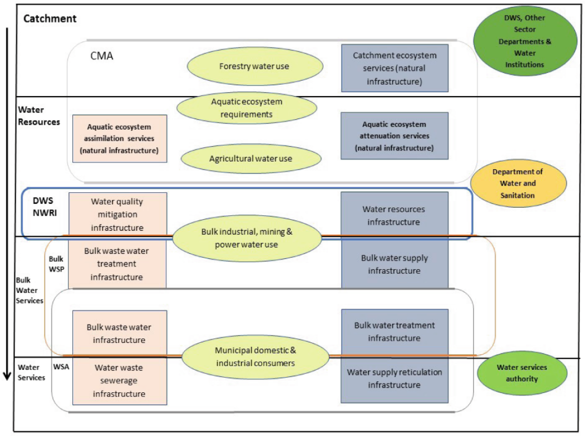
particularly in superannuation and managed funds, which could be increasingly tapped for infrastructure investment. Closing this circle – between infrastructure capital needs and private-sector capital availability – should be a priority.

The research objective for the study was based on water infrastructure investment needs for the development of a comprehensive water infrastructure investment framework and models that will inform funding and financing, integrated planning and life-cycle costing, i.e. planning and construction costs, operation and maintenance. Primary and secondary data were collected for the development of investment models for water infrastructure in South Africa. Participants included municipalities, water utilities, water entities and agencies, local and international private companies with an interest in water and sanitation infrastructure, multilateral financial institutions, commercial banks, and government departments. The methods used included: 1) surveys and questionnaires; 2) E-mail correspondence with participants; 3) One-on-one interviews with participants, and; 4) focus group discussions. Secondary data were collected from reports relating to water infrastructure and investment needs in South Africa, i.e. case studies, annual reports, data bases, research reports, theses, etc.

2 Water Infrastructure Engineering and Financial Realities in South Africa

The South African government recognizes that it simply does not have the resources required to finance and build water infrastructure as quickly and readily as everyone would like. As such, alternative delivery models are required. After considering various financing and procurement options, the government determined that Alternative Financing and Procurement (AFP) will allow South Africa to finance and implement many large infrastructure projects better and sooner, without tying up public funds that can be used for other purposes. This means the construction work could be financed and carried out by the private sector, which will assume the financial risks of ensuring that the project is finished on-time and on-budget. The completed facility would be publicly owned, publicly controlled and publicly accountable. AFP models can be selected for given projects based on an assessment against the principles articulated in the National Water Resources Strategy (NWRS) framework for planning, financing and procuring public infrastructure (DWA 2013).

Water infrastructure provision, management and investments are hierarchical (cf. Fig. 1), based on administrative and/or political boundaries. The hierarchy ranges from a national level to a local level with the responsibility for the implementation at each level (sphere) varying from the government of the administrative boundaries to a combination of private sector and various public- sector water institutions, i.e. water entities/agencies or special purpose vehicles (SPVs), water utilities or boards, and catchment (basin) management agencies.



Note: CMA = Catchment Management Agency; DWS NWRI = Department of Water and Sanitation, National Water Resources Infrastructure; WSP = Water Services Provider; WSA = Water Services Authority

Fig. 1. A new hierarchical institutional governance framework for water infrastructure in South Africa.

Infrastructure investment began to decline in the 1990s as governments increased the share of public consumption expenditure in their budgets at the expense of public investment. Fiscal policies of budget surpluses and debt reduction have reinforced this decline. Government capital expenditure as a share of GDP, which was around 7.2% in the 1970s and 1980s, has fallen to a low of 3.6% of GDP. Business leaders, politicians, professional economists, local governments, industry and community groups have increasingly expressed concern over the decline in South Africa's infrastructure investment and have stressed the need for action. The post–1994 South Africa has placed more pressure on the National Treasury (NT) and Department of Water and Sanitation (DWS) to develop alternative funding (financing and economic) analyses and models for the provision of improved national water infrastructure.

The South African Institution for Civil Engineering (SAICE) recent publication has revealed the very serious problems now facing South Africa (SAICE 2017). Rating on a scale of "A" to "D", revealed that the water infrastructure class received a D^{-}, however, indicating it was in serious condition and need urgent attention, although sufficient for South Africa's current and immediate future needs.

3 Funding and Financing (Investment) Models for Water Infrastructure

Emerging infrastructure backlog and deficient capability warrants immediate attention if South Africa is to build upon, and secure, its sustained economic growth and productivity gains. The first task is to overcome the highly visible and well- documented backlog in existing infrastructure. The second task is to establish new, forward-looking and resilient institutional frameworks to facilitate timely infrastructure investment by integrating the full range of strategic planning, management and technical expertise in South Africa's public and private sectors. The framework for water infrastructure funding and financing models that was designed to meet the challenges presented by the current and growing imbalances that exist between the supply of and demand for water in South Africa (cf. Table 1; Figs. 2 and 3). Such models as presented in Table 1, could constitute alternative and innovative water infrastructure funding models for the development of future water infrastructure projects in South Africa. The new or modified funding models could take the form of one or a combination of these models.

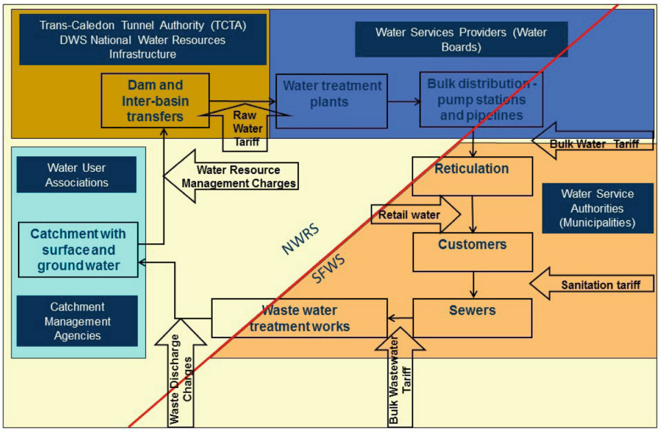
Table 1. Types of funding models with their characteristics and application for the financing of water infrastructure in South Africa.

Types of funding and financing	Application and characteristics
National Revenue Fund (on-budget) budget allocations for water resources infrastructure	Infrastructure spending includes direct expenditure on water infrastructure projects, i.e. water and waste water infrastructure projects through direct transfers
Grants (Municipal Infrastructure Grant (MIG), Equitable Share (ES); Conditional Grants) from the National Revenue Fund (on-budget) for water services infrastructure	Allocations to provinces, local government or municipalities from the National Revenue Fund (NRF), which are provided for and whose purpose is specified annually (NT 2018). Water services and sanitation (waste water) are the biggest concern in terms of backlogs in South Africa. The total capital required to meet current backlogs and projected future demand as well as undertake required rehabilitation for all municipal services in all municipalities to be US$90 billion over 10 years (including provision for escalation). These include the capital investment required by water service providers and water services authorities
Tariff models (via balance sheet) for cost recovery	Current legislation provides for a 3-tier pricing of water supply and management services, including (cf. Figs. 1, 2 and 3) (DWAF 1997, 2007): Tier-1: raw water uses from the water source or supplied from a government waterworks; Tier-2: water supplied in bulk by public water institutions, i.e. water utilities or boards, Trans- Caledon Tunnel Authority (TCTA), Komati Basin Water Authority (KOBWA); and Tier-3: water distributed to households (by water services authorities). Appropriate water pricing and charges, i.e. water-user charges, tariffs, fees or levies, for water has become a critical issue for South Africa and have substantial impacts on demand and supply. Different water tariffs are employed in South Africa to ensure revenue collection to maintain and expand the water supply infrastructure (cf. Figs. 1, 2 and 3)

(*continued*)

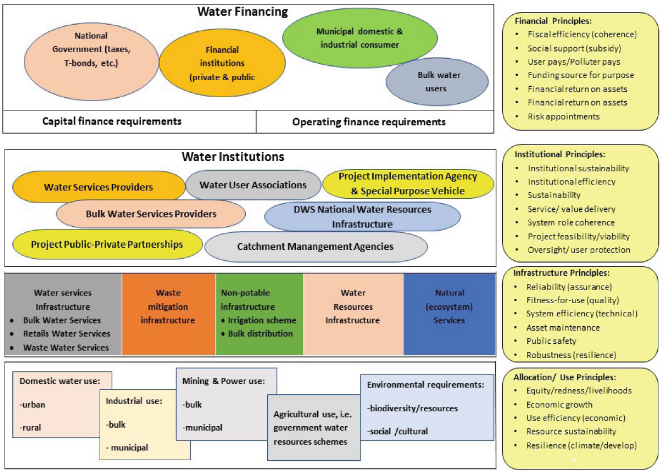
Table 1. (*continued*)

Types of funding and financing	Application and characteristics
Regional Bulk Infrastructure Grant (RBIG)	"Enabling infrastructure" required to connect water infrastructure over vast distances with bulk and reticulation systems
Infrastructure Operations and Maintenance Grants	Preventative measure for the further deterioration of water infrastructure. This augments funding for water infrastructure operations and maintenance in addition to available reserves
Raising of funds on the financial markets (off-budget) through Special Purpose Vehicles (SPV)	The use of special purpose vehicles (SPV) for raising funds from the financial markets, off-budget, for the financing of water infrastructure. These are multidisciplinary organisations specialising in project financing, implementation, liability management, developing, operating and maintaining the water resources infrastructure in the basin. (cf. Figs. 4 and 5)
Private-Public-Partnerships (PPP)	An institutional framework has been developed to guide this type of development and this both accounts for and contributes in part to the mixed experiences (Figs. 1, 4 and 5). The implementation of this framework is essential in allowing the inclusion of the private sector for the implementation of water infrastructure development projects or other forms of non-traditional funding or delivery are appropriate (cf. Ruiters 2013; Ruiters and Matji 2017)
Private sector markets	A well-structured privatisation model could be part of the water infrastructure development and operations and maintenance solutions with the injection of private capital into new water infrastructure (cf. Figs. 4 and 5)
Financial institution for infrastructure	A national infrastructure bank to help finance transformative projects of national importance, i.e. Development Bank of Southern Africa (DBSA) (Ruiters 2013; Ruiters and Matji 2017). This would overcome weaknesses in the current selection of projects by removing funding decisions from politically volatile appropriations process
Innovative and/or alternative financing models	An institutional framework has been developed to guide this type of development and this both accounts for and contributes in part to the mixed experiences (Figs. 1, 4 and 5). The implementation of this framework is essential in allowing the inclusion of the private sector for the implementation of water infrastructure development projects or other forms of non-traditional funding or delivery are appropriate (cf. Ruiters 2013; Ruiters and Matji 2017)



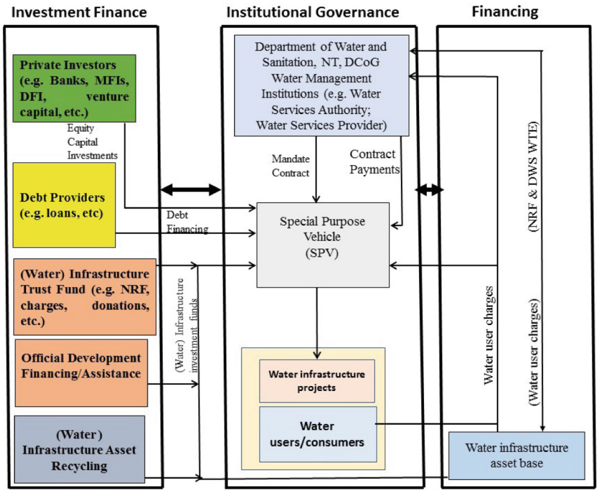
Note: TCTA = Trans-Caledon Tunnel Authority; DWS = Department of Water and Sanitation; NWRS = National Water Resources Strategy; SFWS = Strategic Framework for Water Services

Fig. 2. Water sector value chain in South Africa.



Note: DWS = Department of Water and Sanitation

Fig. 3. A new national water infrastructure investment (funding and financing) framework model for water infrastructure in South Africa.



Note: DWS WTE = Department of Water and Sanitation, Water Trading Entity; DCoG = Department of Cooperative Governance; DFI = Direct Foreign Investment; MFI = Multi-lateral Financial Institutions; NRF = National Revenue Fund

Fig. 4. A new Public-Private Partnership (PPP) as investment model for water infrastructure in South Africa.

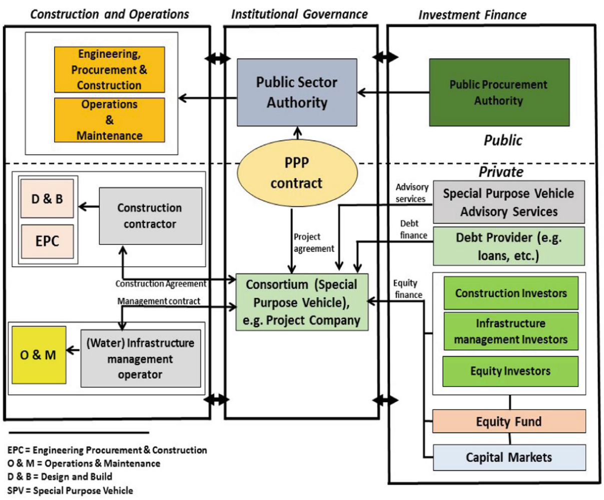


Fig. 5. A new typical traditional public model and Public-Private Partnership (PPP) as investment models for the funding and financing of water infrastructure in South Africa

4 Conclusion

A solution to the funding and financing of the water infrastructure problem in South Africa could be a combination of the models listed above. Some of the funding models are already in existence but they are fragmented and in need of serious review and reconfiguration. Although some of these investment models partially fund South Africa's public water infrastructure, the link between costs and use is not well established. Reinforcing the relationship could create stable funding vehicles that do not depend solely on general tax revenues. In the water sector, these imperatives are now greater than ever before. Furthermore, these investment models can play a greater role in meeting the investment needs of South Africa's water infrastructure and raise revenues to support sustainable water infrastructure. If there is the intention to proceed on the tenet that water infrastructure is an essential part of the nation's capital infrastructure providing a basis for economic, social and environmental development, then there should be in place funding models for water infrastructure just as how funding models exist for other capital infrastructure development, e.g. electricity, energy, transportation (roads), and telecommunications. Combining the models and addressing the regulatory environment would depend on government structure, financial markets and the political climate, to name but a few. If the water infrastructure is classified as an essential part of a nation's capital infrastructure producing goods for public benefits, then the above models should be favourable alternatives for obtaining capital financing. These models can be consolidated to create a water infrastructure funding model pool and/or (water) infrastructure trust fund. From this pool and/or trust fund, suitable model (s) can be selected for water infrastructure financing based on the implementation environment.

References

DWA (Department of Water Affairs, Republic of South Africa): National Water Resources Strategy: June 2013, 2nd edn. (2013). http://www.dwa.gov.za/nwrs2013

DWAF (Department of Water Affairs and Forestry, Republic of South Africa): Water for Growth and Development, Version 8 (2008). http://www.dwaf.gov.za/Documents/Notices/WFGD_Framework_v8.pdf

DWAF (Department of Water Affairs and Forestry, Republic of South Africa): Water Services Act of South Africa. Government Printer, Pretoria, South Africa (1997)

DWAF (Department of Water Affairs and Forestry, Republic of South Africa): Raw Water Pricing Strategy of South Africa. Government Printer, Pretoria, South Africa (2007)

DBSA (Development Bank of Southern Africa): Water Security in South Africa. Development Planning Division Working Paper Series No. 12. Development Bank of Southern Africa, Midrand, South Africa (2009)

NPC (National Planning Commission, Republic of South Africa): National Development Plan for South Africa (2013). http://www.npc.gov.za/documents

NT (National Treasury, Republic of South Africa): Estimates of National Expenditure 2017. National Treasury, Pretoria, South Africa (2018)

Ruiters, C.: Funding models for financing water infrastructure in South Africa: a framework and critical analysis of alternatives. Water SA **39**(2), 313–326 (2013)

Ruiters, C., Matji, M.P.: Funding and financing for wastewater and sanitation infrastructure in South Africa: pricing, tariffs and operational efficiency. In: Ferreira Tiryaki, G., Mota Dos Santos, A.L. (eds.) Infrastructure Investments: Politics, Barriers and Economic Consequences, pp. 183–216. Nova Publishers, Inc., New York (2017)
SAICE (South Africa Institution of Civil Engineering): SACIE infrastructure report card for South Africa 2011 (2017). http://www.saice.org.za/html
WB (World Bank): Africa's infrastructure: a time for transformation. World Bank, Washington, DC (2010)

Using Total Cost of Ownership to Compare Supplier Product-Service System Offering

Giuditta Pezzotta(✉), Gianluca Pezzotta, Emanuele Dovere, Fabiana Pirola, and Paolo Gaiardelli

Department of Management, Information and Production Engineering, University of Bergamo, Bergamo, Italy
giuditta.pezzotta@unibg.it

Abstract. Companies plan investments and acquire assets evaluating among different alternatives what is more suitable to satisfy their needs. Management choices should always be forward-looking, accurate, and coherent; otherwise, firm competitiveness in the market may be profoundly affected, or, even worse, its survival may be severely put to the test. In this context, implementing a robust and refined decision-making process is essential. Total Cost of Ownership (TCO) is a reliable approach that can be used to estimate the convenience of an investment in the long-term since it takes into account all costs associated with the entire lifecycle of an asset. The concept of TCO is nowadays well known and widely studied in the literature, and can be used for several purposes, not exclusively related to economic benefits of a managerial decision. Notably, it can be used as a decision making parameter to be evaluated in the selection of the kind of Product-Service System (PSS, e.g. maintenance, training) the customer could buy from a supplier. In particular, a logical process to support the PSS provider to compute TCO is proposed. This would allow the suppliers to compare different PSS offering and propose to their customers the most appropriate and convenient one, as a combination of the asset and a set of services to guarantee to customers cost and operational efficiency in the long-term. The development process is validated by using real data related to alternative PSS scenarios a supplier can propose to the customer.

1 Introduction

The concept of Total Cost of Ownership (TCO) has been analysed widely in the academic literature. TCO nomenclature is used to identify "all costs associated with the acquisition, use, and maintenance, of a good or service" (Ellram and Siferd 1993). Estimating the TCO is essential when evaluating the cost associated to a specific asset along with its lifecycle, whatever the firm operates as a service provider, as a retailer, or as a manufacturing company (Burt et al. 2003). Paying too much attention only to purchasing price is myopic, not addressing significant ownership or post ownership costs (Burt et al. 2003), and it may affect the choice of an investment without adequately analysing the costs and benefits in the long-term perspective. Trends like increased emphasis on the provision of product service systems (PSS), supplier base rationalisation, increased global competition, and a growing recognition

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 183–191, 2020.
https://doi.org/10.1007/978-3-030-48021-9_21

on the significance of purchasing expenditures, have implied greater attention on TCO implications (Saccani et al. 2017). However, despite its potentiality, firms are often still reluctant to adopt a TCO approach in investment perspectives (Visani et al. 2016; Saccani et al. 2017), due to managers' unfamiliarity of the concept, lack of information, unwillingness to share data (Milligan 1999), difficulty to estimate correctly asset TCO (Meckbach 1998), or, again, because they focus primarily on purchase price, wrongly believing that reducing purchasing costs also reduced indirect costs (Avery 1999). Modern technology and dissemination of knowledge can help break down these barriers: several TCO systems are available on the market (Ersten 1997) and current computer systems let organisations grasp data and organize them in a more flexible and usable way (Ellram 2002). Notably, TCO can be used as a decision making parameter to be evaluated in the selection of the kind of product and the related combination of service (namely Product-Service System (PSS), e.g. maintenance, training) the customer is willing to buy from a supplier along the product lifecycle. In this perspective, a logical process, based on well-known approaches, to support PSS providers in computing TCO in order to use it to compare different PSS offering along the asset lifecycle. This would allow the supplier to propose to their customer the most appropriate and convenience PSS offering, as a combination of the asset and a set of services to guarantee to customers cost and operational efficiency in the long-term. The paper is therefore organised as follows: in Sect. 2 state of the art is summarized while in Sect. 3 the logical flow is presented. Section 4 summarizes the application case used to test the logical flow while Sect. 5 closes the work and proposes further research prospects.

2 State of the Art

Correct estimation of TCO requires firstly to have clear in mind which costs incur in the long-term perspective. Academics have proposed several approaches to model TCO. Ellram and Siferd (1993) suggest categorising cost drivers in six main classes: quality, management, delivery, service, communication and price. Ellram (1993) proposes a model of TCO for which three main categories of costs have to be identified: pre-transaction costs, transaction costs and post-transaction costs. Burt et al. (2003) classify three main cost areas, which are acquisition costs, ownership costs and post-ownership costs. Anderson et al. (2009) identified in the acquisition, operating and disposal costs the principal expenditures. From a global taxonomic point of view, two main types of TCO models exist: standardised models, that can be used for a variety of purchases; and unique models, that are related to a specific purchase (Ellram 1994), which involve reasonably a deeper level of complexity and detail. The appropriateness of using a standard model varies case by case, since "there is no standard approach for successful TCO use and implementation" since it is company dependent (Ellram 1994). Two types of stakeholders with different perspectives can be identified in the TCO analysis: users (industrial equipment or plant owners/managers) and suppliers (industrial equipment or plant builders/providers). As stated by Roda and Garetti (2014), having analysed several publications, both users and suppliers are interesting regarding the potentiality of having a tool able to calculate the TCO. However, while TCO implementation advantages are deeply analysed in literature from the customer point of view, little research has been done on the possibility of using TCO also from the supplier perspective, and results are not always unanimous regarding its utility for suppliers (Rosenback 2013). Despite this, TCO can be useful when selling product-service proposals of high priced assets that can guarantee cost savings in the long-run (Brown 1979): in this perspective, buyers evaluate during the purchasing factors such operating and maintenance costs, and length of services along the lifecycle. Consequently, suppliers must consider these issues in all the phases of the asset lifecycle. Similarly, Stremerch et al. (2001) highlights that customer firms comprehensively evaluate during the product purchase the provider service offering due to the high monetary value, high perceived complexity and involving long-term mutual commitment; in this perspective, providers must support their customer in the purchase selection providing a detailed TCO estimation of their alternative PSS offering. Through a case study, Rosenback (2013) identified how industrial suppliers could use TCO to show the superiority of their offering; this concept is similarly reprised by Wynstra and Hurkens (2005), which affirm that customers prefer to purchase higher priced and valued offerings if TCO savings in the long-term perspective are considerable. However, no definite guidelines useful for suppliers to estimate asset TCO seem to exist in the literature.

3 Developed Methodology

Adapting classical methodologies to compute TCO, Woodward (1997) and Burt et al. (2003) and integrating them with concepts taken maintenance and operations management literature, a logical process that suppliers can follow to estimate correctly asset TCO has been developed. The developed approach is based on the evaluation of both costs and performance of the asset (Roda and Garetti 2015). In our understanding, to provide accurate TCO evaluation of the TCO of the PSS offering, availability, reliability and maintainability have been considered.

Steps of the process are resumed in the following figure (Fig. 1):

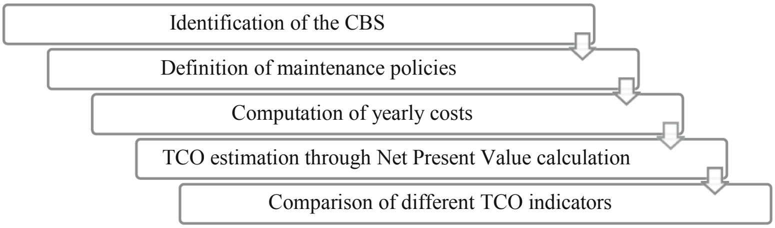


Fig. 1. Process for suppliers to estimate asset TCO

The first step regards the identification of the Cost Breakdown Structure (CBS), so the definition of all cost drivers incurring in the asset lifecycle. Four main categories of costs can have been identified (Woodward 1997; Burt et al. 2003; Ensparantza et al. 2006; Marquez 2007): acquisition costs, operating costs, maintenance costs and disposal costs. Acquisition costs (or initial capital) refers to all the costs related to the acquisition of the asset and its installation and commissioning (Woodward 1997). Operating costs includes all necessary cost items to run the asset during its life cycle while maintenance costs include all the expenditures which are related to restore the asset to its functionality after failures, or to prevent it from breakdowns (Woodward 1997; Ensparantza et al. 2006). Disposal costs incur at the end of an asset's working life, and they are deducted from the residual value of the asset itself (Woodward 1997). It is relevant to state that the level of detail and the form of the CBS varies consistently according to the asset under examination (Roda and Garetti 2015); therefore, the cost structure should be refined case by case.

The second step concerns the definition of service policies, so how supplier intends to provide services, mainly related to the maintenance activities. Often suppliers can provide a different type of service and service related contracts to customers, and adopt a specific maintenance framework could impact drastically on the TCO and asset performance. Since maintenance actions are strictly related to the replacement or repair of entities, it becomes essential to predict components' state and failures' behaviour during the asset lifecycle. In this phase, the four maintenance functions (reliability function, failure probability distribution function, failure probability density function and failure rate) of components and Mean Time Between Failures have been calculated. This implies that suppliers must also be able to estimate the functions in question, from a statistical point of view. Maximum Likelihood Estimation (MLE) and Median Rank Regression (MRR) are the most widely known approaches to estimate functions' parameters, by merely having historical data concerning time to failure of items (Abernethy 2000; Marquez 2007; Genschel and Meeker 2010).

The third step of the process refers to the computation of yearly expenditures related to the asset. Analytical formulas to compute the monetary values of each cost driver are suggested. In particular, maintenance expenditures are computed only for the most critical components, whose maintenance schedule can have a substantial impact in the long-term perspective, identifying criticality level using qualitative or risk assessment techniques (Marquez 2007).

The following step regards discounting estimated yearly costs on the base period and summing them, finding the Net Present Value (NPV), which is the measure of the overall TCO by using an appropriate discount rate, estimated through the Weighted Average Cost of Capital of the company (Watson and Head 2010), or by asking industrial accountants (Woodward 1997).

The fifth step is the comparison of TCO obtained from different scenario implementing different PSS offering and related different maintenance policies. Comparing different scenario has the aim to identify the one that guarantees a win-win economic agreement for both the supplier and the customer. Since the customer can select the solutions that minimise the asset cost along the lifecycle, while the supplier can use the cost estimation to evaluate pricing policy and contract specifications better.

4 Application in Real Case

The process explained in Sect. 3 has been applied in a real supplier company, namely Alpha, a leader in the industry of numerical control machines, with the aim to support the sale of services to the customer. In particular, TCO has been estimated for a multi-axis workstation utilised to work composite materials and light alloys, namely "Machine A", which is predicted to have ten years of useful life. As TCO model inputs vary according to the customer to which the asset is sold. Therefore, the presented calculation focus on a specific customer located in Italy. Following precisely the process showed in Sect. 3, the CBS has been defined as the following:

- acquisition costs (Purchase price, training and delivery costs),
- operating costs (Energy and set up costs),
- maintenance costs (Corrective maintenance, preventive maintenance, components and technicians' transportation costs),
- disposal costs (Second-hand market value).

In the second step, the PSS offering has been defined to perform a comparison of the TCO in different PSS scenarios. The company offers to the customers mainly maintenance related services: corrective maintenance (CM), and time-based preventive maintenance (PM). To estimate parameters of basic maintenance functions and MTBF, starting from historical time to failure (TTF) of components, software R has been used. Only most critical components, identified with risk assessment techniques, have been included in the analysis. Also, confidence intervals for components' reliability functions and MTBF have been calculated. Therefore, several maintenance-related service scenarios have outlined. Three scenarios have been defined for CM, in which all components are replaced at the medium value and lower and upper bound of MTBF confidence level, namely "estimated", "pessimistic" and "optimistic". Moreover, since "pessimistic" and "optimistic" scenarios are not likely to happen, further CM scenarios have been identified by adopting a Design of Experiment methodology. For PM, five scenarios have been created, in which components are replaced at the 90%, 80%, 70%, 60% and 50% level of reliability. Scenarios based on three components that could have a substantial impact on the final TCO have been identified. Ten years of fully production time (FPT) for scenario has been calculated. It has been then possible to compute a ratio between every PM configuration FPT and every FPT obtained in different corrective and realistic maintenance scenarios. Doing an average of all the computed ratios related to a specific PM policy, it has been possible to evaluate if adopting a proactive PSS offering based on PM could impact positively on system performance. Results are the following (Table 1):

While it is clear to see that providing 90% reliability level PM creates a counter-productive situation of over-maintenance, it is also possible to see that implementing a service-based contract related to PM provision of at least at the 80% of reliability increases the performance. In the third step of the process, yearly costs have been computed, using specific analytical formulas for every identified cost element and maintenance strategy. In a context of PM, two levels of risk have been delineated: a medium risk profile, for which monetary risk costs are calculated on the basis of

Table 1. Average FPT gain or loss of adopting a specific PM policy rather than a CM.

PM policy	Average FPT gain/loss concerning CM frameworks
PM: 90% reliability	−0,22%
PM: 80% reliability	+0,12%
PM: 70% reliability	+0,20%
PM: 60% reliability	+0,21%
PM: 50% reliability	+0,20%

estimated CM costs; and high risk profile, for which, instead, risk costs have been estimated on the basis of CM costs in the realistic scenario with the highest value of 10 years annualized maintenance costs. Once yearly costs have been found, the NPV/TCO for each scenario has been computed. The appropriate rate of return for the company has been established at 10%. An example of the computations performed for each scenario is reported in the appendix.

In the next two tables, a comparison of the TCO obtained in the different scenarios are reported (Tables 2 and 3):

Table 2. TCO of "Machine A" in the corrective maintenance scenarios

	CM: Optimistic scenario	Realistic scenario with lowest TCO	CM: estimated scenario	Realistic scenario with highest TCO	CM: Pessimistic scenario
TCO (10 year NPV)	564.255 €	582.556 €	589.702 €	600.144 €	669.051 €

From the results, it is clear that, whether the risk profile is assumed, proposing to the customer a maintenance-related service implementing a PM at 80% level of reliability or below it is more economically convenient for the customer than adopting a pure CM strategy. This is true also for different customers, empirically, in fact, it results that if the customer is more distant, then benefits of adopting a PM strategy are even more consistent, since, in this case, plant shutdown and transportation costs associated with CM scenarios are higher. In this perspective, the customer can only be incentivised in signing service contracts with the supplier. In that way, supplier bonds the customer in the long-term perspective, with the possibility that he/she will acquire again from the supplier itself in the future. Moreover, delicate and essential feedbacks from the customer can be acquired during the partnership, to improve the PSS and related performance.

Table 3. TCO of "Machine A" in the preventive maintenance scenarios (medium risk)

	PM: 90% reliability	PM: 80% reliability	PM: 70% reliability	PM: 60% reliability	PM: 50% reliability
TCO (10 year NPV) Medium risk	622.730 €	587.193 €	577.965 €	575.106 €	576.776 €
TCO (10 year NPV) High risk	623.586 €	588.904 €	580.532 €	578.528 €	581.054 €

5 Conclusion

TCO is a powerful instrument that however it is still at a primitive stage in terms of diffusion by companies, even if benefits of using this approach are undiscussed and deeply underlined in literature. Advantages of estimating TCO also exist for suppliers, especially for those that sell expensive assets and related maintenance services. Firstly, it can be useful during negotiations with customers, which always appreciate having a detailed overview of the asset they are buying. Providing exhaustive information on total cost, and particularly its deviation when adopting different maintenance policies, could be used to create competitive advantage concerning competitors that do not provide a TCO estimation of their assets. Within this approach, the supplier can demonstrate that through a PM contract benefits for the customer are both monetary and from a performance point of view. Therefore, suppliers, especially those that provide maintenance services of their assets, must have clear in mind how to estimate TCO. The process proposed in this paper goes in this direction by proposing a series of steps that suppliers could follow when computing TCO. The present research has several limitations, among others the more relevant one is that it is based on deterministic computation. Further development will be related to moving from deterministic to stochastic approach by use of cost engineering methods.

Appendix

Example of NPV/TCO calculation in the estimated corrective maintenance scenario

t = 0	AC			OP		MC			DC	Cash flow
	Purch.	Shipp.	Training	Energy	Set up	Prev.	Corr.	Transp.	Second hand market value	
Y1	250.000 €	1.000 €	6.300 €							257.300 €
Y2				38.328 €	8.750 €		0 €	0 €		47.078 €
Y3				38.328 €	8.750 €		0 €	0 €		47.078 €
Y4				38.178 €	8.768 €		7.974 €	1.700 €		56.619 €
Y5				38.052 €	8.785 €		6.162 €	1.700 €		54.699 €
Y6				38.068 €	8.785 €		5.504 €	1.700 €		54.057 €
Y7				37.768 €	8.820 €		15.366 €	3.399 €		65.353 €
Y8				38.328 €	8.750 €		2.372 €	1.133 €		50.584 €
Y9				37.614 €	8.838 €		18.590 €	3.966 €		69.007 €
Y10				38.328 €	8.750 €		1.155 €	567 €		48.800 €
NPV				38.068 €	8.785 €		6.637 €	2.266 €		55.756 €
Disc. costs	250.000 €	1.000 €	6.300 €	234.298 €	53.918 €	0 €	35.229 €	8.957 €	0 €	589.702 €

References

Abernethy, R.B.: The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life. Safety, Survivability, Risk, Cost and Warranty Claims, North Palm Beach (2000)

Anderson, J.C., Narus, J.A., Narayandas, D.: Business Market Management: Understanding, Creating, and Delivering Value. Pearson Prentice Hall, Upper Saddle River (2009)

Avery, S.: MRO costs have big impact on profits! Purchasing **126**, 95–96 (1999)

Brown, R.J.: A new marketing tool: life-cycle costing. Ind. Mark. Manag. **8**(2), 109–113 (1979)

Burt, D.N., Dobler, D.W., Starling, S.L.: World Class Supply Management: The Key to Supply Chain Management. Irwin/McGraw-Hill, Boston (2003)

Ellram, L.M.: A framework for total cost of ownership. Int. J. Logist. Manag. **4**(2), 49–60 (1993)

Ellram, L.: A taxonomy of total cost of ownership models. J. Bus. Logist. **15**(1), 171 (1994)

Ellram, L.M.: Total cost of ownership. In: Handbuch Industrielles Beschaf- fungsmanagement, pp. 659–671. Gabler Verlag (2002)

Ellram, L.M., Siferd, S.P.: Purchasing: the cornerstone of the TCO concept. J. Bus. Logist. **14**(1), 163 (1993)

Genschel, U., Meeker, W.Q.: A comparison of maximum likelihood and median-rank regression for Weibull estimation. Qual. Eng. **22**, 236–255 (2010)

Meckbach, G.: Study reveals organizations blase about TCO. Comput. Canada **24**(43), 15–16 (1998)

Márquez, A.C.: The Maintenance Management Framework: Models and Methods for Complex Systems Maintenance. Springer, London (2007)

Roda, I., Garetti, M.: The link between costs and performances for total cost of ownership evaluation of physical asset. State of the art review. In: 2014 International ICE Conference on Engineering, Technology and Innovation (ICE), pp. 1–8 (2014)

Rosenback, M.: Antecedents and obstacles to total cost of ownership analysis in industrial marketing: a case study. In: 29th IMP Conference, Atlanta, Georgia (2013)

Saccani, N., Perona, M., Bacchetti, A.: The total cost of ownership of durable consumer goods: a conceptual model and an empirical application. Int. J. Prod. Econ. **183**, 1–13 (2017)

Stremersch, S., Wuyts, S., Frambach, R.T.: The purchasing of full-service contracts: an exploratory study within the industrial maintenance market. Ind. Mark. Manag. **30**(1), 1–12 (2001)

Visani, F., Barbieri, P., Di Lascio, F.M.L., Raffoni, A., Vigo, D.: Supplier's total cost of ownership evaluation: a data envelopment analysis approach. Omega **61**, 141–154 (2016)

Watson, D., Head, A.: Corporate Finance: Principles and Practice. Pearson Education, London (2010)

Woodward, D.G.: Life cycle costing–theory, information acquisition and application. Int. J. Proj. Manag. **15**(6), 335–344 (1997)

Wynstra, F., Hurkens, K.: Total cost and total value of ownership. In: Perspektiven des Supply Management, pp. 463–482. Springer, Heidelberg (2005)

Reducing and Reinvesting Working Capital in Business Ecosystems

Matti Rissanen(✉), Miia Pirttilä, Lotta Lind, and Timo Kärri

Industrial Engineering and Management,
Lappeenranta-Lahti University of Technology, Lappeenranta, Finland
{matti.rissanen,miia.pirttila,
timo.karri}@lut.fi, lotta.m.lind@gmail.com

Abstract. Issues related to supply chain finance, also known as supplier finance or reverse factoring, have been in focus of both academics and practitioners after the financial crisis. In the broadest perspective of supply chain finance, it concerns all supply chain objects: the flow of goods, information and finance; processes; current and fixed assets; and personnel involved in the supply chain. This way, supply chain finance connects with the assets management of business ecosystems. The focus of schemes in this area is to support the working capital management of the key suppliers of the buyer. The schemes are implemented between the tier 1 supplier and the buyer. In this paper, we study working capital management in an information and communications technology business ecosystem producing desktop computers, spreading the study from the focal actor to the component suppliers in the furthest tier. We show the potential that exists to release capital from operational issues to more profitable purposes.

Keywords: Working capital management · Business ecosystem · Supply chain finance · ICT industry

1 Introduction

Williamson (1985) categorizes assets besides fixed and variable parts further based on the degree of specificity, wholly specific and nonspecific. Most of the previous asset management literature discuss the specific fixed assets like site specify, physical assets specify or human assets specificity. Our interest direct to specific variable (current) assets that can be considered including dedicated assets in Williamson classification i.e. inventories, accounts receivable and accounts payable. Asset items accounts receivable and payable are highly specific, the products or services are sold to specific buyers. The amount of bill is in accounts receivable for the supplier of products or services and the same amount is registered to a buyer's accounts payable.

The flow of financial resources in supply chains gained attention during recent years (Hoffman and Kotzab 2010; Wuttke et al. 2013; Gelsomino et al. 2016). The financial crisis of 2008 acted as a beginning for scholars to study and companies to manage the financial flow in addition to the material flow of a supply chain (Seifert and Seifert 2008). Large or otherwise powerful companies took advantage of delaying payments to their suppliers in order to improve their working capital management i.e.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 192–199, 2020.
https://doi.org/10.1007/978-3-030-48021-9_22

the greater proportion of assets is financed by suppliers than before the financial crisis. The first tier suppliers of these large companies passed the problem forward to their suppliers, enforcing them to extend payment terms and so the effect continued further upstream. The government of the United Kingdom solved the liquidity problem of small and medium sized suppliers by introducing a reverse factoring program in order to finance the supply chains (Seifert and Seifert 2008). This program offers early payments to British suppliers at a cost based on buyers' credit rating and the powerful buyer can implement strategy that in other context would be squeezing suppliers. To protect European business, the European Union introduced a late payment directive that aim to limit payments to supplier max 60 days.

Business ecosystems are characterized by supplementary functions between individual companies and coevolution in a direction set by the focal actor (Moore 1998; Adner 2017). The companies are drawn together through a platform (a technology, tool or service) provided by the focal actor for the ecosystem partners to use to improve their own performance as well as the performance of the ecosystem (Iansiti and Levien 2004). Members of the ecosystem should jointly participate in the development of the platform to improve the competitiveness of their customers. Information and communications technology (ICT) ecosystems and industry as whole are renowned for requiring continuous innovation and increasing production capacity. For this reason, ICT companies collaborating should think of their networked environment as an ecosystem consisting of multilateral relationships rather than a traditional supply chain mainly consisting of bilateral relationships. An example case of a failed ICT ecosystem is IBM (see Iansiti and Levien 2004), who limited innovation and extracted value from its partners only to fail when more open strategies were offered by companies like Microsoft and Apple.

The relationships within a supply chain can be broken into dyadic ones whereas the relationships in business ecosystems are multilateral (Adner 2017). The positions of different companies in supply chains are divided into upstream (supplier) and downstream (buyer) (ibid.) but in business ecosystems, the multilateral relationships have to be described in other manners. In fact, the positions and boundaries of an ecosystem are impossible to precisely define since it can take part in multiple industries but rather the ecosystem should be described by recognizing the organizations that are the most critical to the business of the focal actor (Iansiti and Levien 2004).

In a global context, ensuring the health of a supply chain from a financial point of view by directives and programs may not be a sustainable solution – instead, it rather hides and shifts the original problem without solving it. In this paper, we study the working capital management of a global ICT ecosystem with Dell as its focal actor. We demonstrate the potential to release capital by improving the cycle time of working capital. The demand for semiconductor products and other electronics increase through the rise of Internet of Things and other technological advances. To meet this demand companies in ICT sector have to continuously innovate and invest in new plants and decrease the risk of disturbances in business ecosystem to ensure production. The aim of this paper is to demonstrate through an example of how companies operating collaboratively in a business ecosystem by using working capital management can reduce tied up capital. In addition, we introduce some possibilities on how to reinvest the capital in more profitable targets.

2 Research Design

The focus of this study is on a business ecosystem in ICT industry. Dell was selected as the focal actor for the ecosystem to be studied, because Dell operates on negative cycle time of operating working capital (Bruun and Mefford 2004; Hofmann and Kotzab 2010; Lind et al. 2016). This indicates that the suppliers and buyers of Dell favour it by offering long payment time and paying fast. Dell has no need to arrange the finance of its specific current assets under study. We started the research process by positioning the suppliers listed by Dell (2018) to the ecosystem (Fig. 1). The companies, who had data available and were included in the study, are presented in solid colour. In addition to using the supplier list of Dell, some companies were added to the ecosystem to ensure that with the ecosystem companies we would be able to build a working desktop computer. The ecosystem presented in Fig. 1 is by no means a full representation of Dell's ecosystem but rather a part of it, which can produce a desktop computer. In addition, we dismantled a display with Dell's brand to find more about the suppliers, since the list Dell provides does not go much into detail on what each company supplies. Some potential relationships are presented in Fig. 1 as lines between the companies. The nature of these links is not in the scope of this study but they could include transfer of materials, information, influence and funds (Adner 2017).

In the ecosystem, the suppliers are set in four tiers based on the importance of the supplier in providing the desktop computer. In the first tier are companies that provide the computational technology: chipset, motherboard, central processing unit and graphical processing unit. Intel stands out in this tier for being a key partner to Dell as well as well as providing most of the chipsets used in Micro-Star International's (MSI) motherboards. MSI also acts as an original design manufacturer for Dell. Adding computer parts delivered by companies in the second tier, we have all the parts necessary in the computer case (power supply unit, memory, storage, cooling). Third tier includes the providers of peripheral devices (keyboard, mouse, display) and parts for supplementary functionalities (audio card, network card, optical drive). Fourth tier consists of companies that provide components and/or manufacturing for all the previous tiers.

Finally, the research sample presents Dell's business ecosystem within ICT industry including figures from 20 companies for each year in the 2014–2016 period. The used methodology can be described as archival research (Moers 2007). Data for the study was collected from public sources and Thomson ONE database. Public sources being financial statements and annual reports published by the studied companies. Due to the fact that some of the public sources used languages the authors could not reliably translate, some data was collected from the Thomson ONE database.

The working capital management in the ecosystem is studied by the cash conversion cycle (CCC) developed by Richards and Laughlin (1980). The CCC consists of three components: cycle times of inventories (DIO), accounts receivable (DSO) and accounts payable (DPO). Relative profitability is measured by the return on total assets (ROA%). The definitions and calculations for each measure are shown in Table 1.

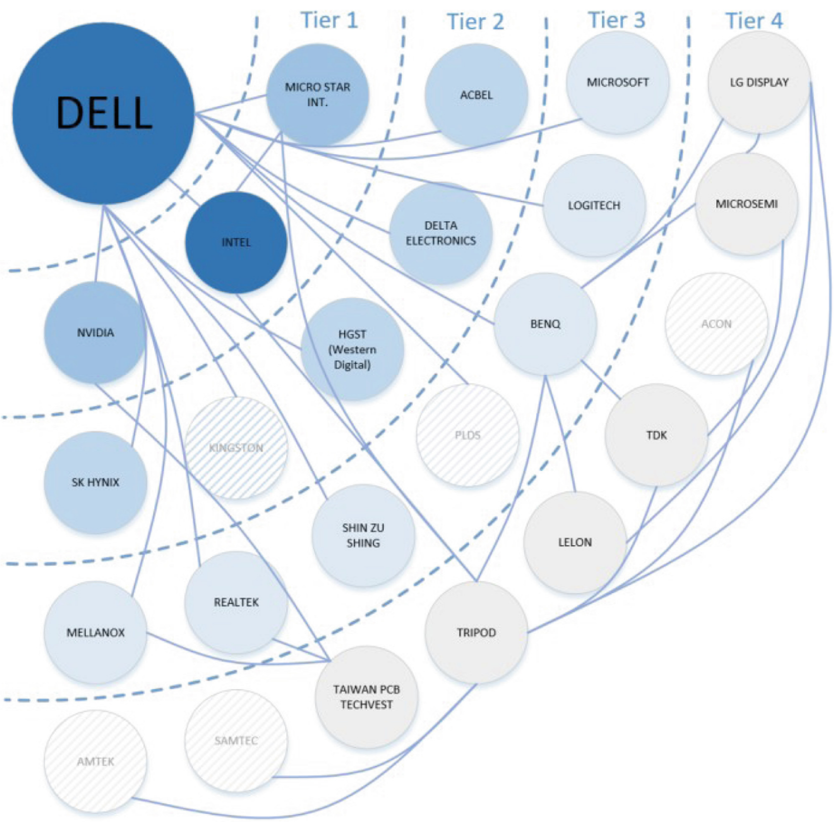


Fig. 1. Structure of the present study ecosystem with the companies of sample

Table 1. Measurement methods of working capital management and profitability

Variable	Description	Definition
DIO	Cycle time of inventories	(Inventories/Sales)*365
DSO	Cycle time of accounts receivable	(Accounts receivable/Sales)*365
DPO	Cycle time of accounts payable	(Accounts payable/Sales)*365
CCC	Cash conversion cycle	CCC = DIO + DSO − DPO
ROAROA	Return on total assets	Earnings before interest and taxes/Total assets

3 Results and Analysis

To analyse the state of working capital management in Dell's ecosystem we calculated the CCC and its components for each year from 2014 to 2016. The return on assets indicates the profitability of the ecosystem. Figures are presented in Table 2 for each company included in the sample and for the four tiers they are placed in.

Table 2. The average figures of CCC and its components and ROA of the years 2014–2016

The name of company	DIO	DSO	DPO	CCC	ROA
DEL ϒ /	**13**	**45**	**89**	**−30**	**−2%**
Tier 1	**43**	**45**	**35**	**53**	**13%**
Intel	32	47	16	63	21%
Micro Star International ϒ /	63	52	66	49	9%
Nvidia	34	37	24	47	9%
Tier 2	**43**	**69**	**52**	**60**	**9%**
ACBEL /	45	87	81	51	4%
Delta Electronics /	43	84	64	62	9%
HGST (Western Digital) ϒ /	46	39	47	38	6%
SK Hynix	37	67	16	88	16%
Tier 3	**35**	**74**	**48**	**60**	**12%**
BenQ /	46	89	78	57	2%
Logitech ϒ /	42	29	46	26	5%
Mellanox	33	53	27	58	37%
Microsoft /	11	77	29	59	12%
Realtek ϒ /	42	44	50	36	6%
Shin Zu Shing /	35	149	57	127	10%
Tier 4	**44**	**90**	**41**	**93**	**5%**
Lelon	55	122	27	150	8%
LG Display /	33	56	41	49	1%
Microsemi	60	57	24	92	3%
Taiwan PCB Techvest /	32	125	69	87	6%
TDK	50	77	43	84	8%
Tripod /	37	103	44	96	5%

The results of CCC calculations reveal that the companies in fourth tier tie up more working capital in comparison to the other tiers. We identified two specific working capital strategies applied by companies in the ecosystem. Dell, the largest direct seller of PCs in the world, has achieved negative CCC by maintaining long payment terms with its suppliers. Besides Dell, four other companies (ϒ in Table 2) in the sample collect the payments of their customer faster than they pay for their suppliers, in other words, their DPO is longer than their DSO. It can be assumed that this is their strategy for managing trade credit. Suppliers finance the purchases of their buyers by offering generous payment terms. The second strategy is for companies to use supply chain finance as a means to finance their inventories, i.e. their DPO is longer than the DIO (/ in Table 2). Thirteen companies of the sample have managed to negotiate favourable payment terms from this point of view so it can be considered a popular strategy, for suppliers to be financing the inventories of buyers, in this ecosystem. Those four companies that postpone payments to their suppliers can finance part of their inventories using supplier finance, and Dell does not have to tie up any money to finance its operational working capital. After this analysis, we have seven companies left that did not apply either of the two working capital management strategies. These companies pay their suppliers relatively fast. They favour their customers offering relative long payment periods and their cycle time of inventory varies from a month to two months. Those whose DIOs are long, the work in process ties up relatively much working capital.

To analyse how much money is tied up in this ecosystem, we analyse the figures of Dell and the four tiers in Dell's ecosystem. The amounts are presented in millions of USD. FED's rates were used for those companies whose figures were in other currencies than USD. The scenario is formed as follows: for each company, the shortest cycle time of each CCC component in their tier was selected as a new cycle time. In the calculations, it was assumed that this shortening of the DIO, DSO and DPO did not affect the sales revenue or EBIT of companies. Calculations with the new cycle times were done for each year of the observation period. The scenario can be considered too general because it does not consider the specific characterises of companies. Our aim is to demonstrate the possibilities of the business ecosystem have to reduce its tied up working capital hence how a company would do that is not at scope. The new value of a working capital item (inventories, accounts receivable and payable) was calculated as following:

$$\textit{value of working capital item} = \frac{\text{(tier's min cycle time * firm's sales revenue)}}{365} \quad (1)$$

This scenario leads to an ecosystem where buyers pay their bills fast to their suppliers and they can in favour give their own customers a rise in payment time from 26 days to 54 in 2016. Cycle time of inventories in each tier varies from 10 days to 43 days in 2016. The value of each working capital item and total assets is presented as the sum of the tier and ROA% is the average of the tier. Shortening the cycle time of working capital components, like set in the scenario, decreases the capital tied up in the ecosystem from three to six percent: 2016: 12 415; 2015: 11 808 and 2014: 16 832 millions USD. The ROA% does not increase remarkable because the total amount of assets are huge in comparison with operational working capital. The improvement of ROA% can be achieved by reinvesting releasing capital to a more profitable target. The scenario means an increase of 1.07%, 3.56%, 14.48% and 14.35% in ROA in tiers 1 to 4 respectively. In Table 3 we present more detailed figures using the year 2016 as an example. The results of scenario suggest that the business ecosystem could release 12 Milliard USD from operational working capital to be reinvested to targets that support the business ecosystem's goal.

Table 3. Actual figures of the ecosystem and results of the scenario in millions of USD, the year 2016

		Inventory	Accounts receivable	Accounts payable	Working capital	Total assets	ROA %
2016	Dell	2 538	9 420	14 422	−2 464	118 206	−2.8%
	Tier 1	6 481	6 254	3 328	9 407	122 255	13.9%
	Tier 2	4 900	6 288	4 025	7 163	64 364	7.4%
	Tier 3	2 760	18 873	7 468	14 164	199 002	11.5%
	Tier 4	3 673	7 227	4 333	6 566	42 426	5.4%
	BEC total	20 352	48 061	33 577	34 837	546 253	
Scenario	Dell	2 538	9 420	14 422	−2 464	118 206	−2.8%
	Tier 1	5 637	6 046	2 816	8 868	121 715	13.9%
	Tier 2	4 789	4 203	1 645	7 347	64 548	7.5%
	Tier 3	2 376	6 371	6 252	2 495	187 333	12.6%
	Tier 4	3 171	5 567	2 561	6 177	42 036	6.1%
	BEC total	18 511	31 607	27 696	22 422	533 8388	

4 Discussion and Conclusions

This paper studied working capital management in the business ecosystem formed around a focal company in the ICT industry. The operational working capital is considered as a dedicated asset in this study. The analysis confirmed the findings of previous studies, which have shown that with the help of its business partners, Dell is able to maintain payment terms that enable operating with negative working capital (Bruun and Mefford 2004; Hofmann and Kotzab 2010; Lind et al. 2016). The scenario conducted in the paper showed that the business ecosystem could release working capital by better collaboration. Benchmarking the cycle times within the tiers and taking the best practices into use could be a way to reduce working capital. Of course, in reality, such scenario would be difficult to realize and would not end up in win-win situations for every actor, but from the perspective of a business ecosystem, it would result in more effective financial flows.

In terms of working capital management, Dell is a very strong player in the ecosystem. The results indicate that its success is enabled by the other partners in the ecosystem. Judging by the customisability of Dell's desktop computer products, they do not limit the innovation of their ecosystem partners in the way that lead IBM's similar ecosystem to fail. The whole ICT industry thrives on innovation and apparently, that is a good enough platform for Dell's ecosystem to succeed. However, Dell has strong competitors and a strong innovative platform from one of them could quickly weaken Dell's position in its own ecosystem.

In recent years, academic research on working capital management has concentrated on different solutions of supply chain finance, including financial institutions, with the aim of ensuring the sufficient financial flows between the different actors in the ecosystem. However, collaboration in working capital management at the ecosystem level could also be conducted through negotiations. Instead of feeding banks with the interest costs and service fees of supply chain solutions, ecosystems could collaborate in order to release excess working capital which is unproductively tied up in the payment terms and inventories, and use it for example to strategic investments which would benefit the whole ecosystem long term.

Acknowledgements. The authors would like to thank the Finnish Cultural Foundation for the financial support of Miia Pirttilä.

References

Adner, R.: Ecosystem as structure: an actionable construct for strategy. J. Manag. **43**(1), 39–58 (2017)

Bruun, P., Mefford, R.N.: Lean production and the Internet. Int. J. Prod. Econ. **89**(3), 247–260 (2004)

Dell: Supply Chain (2018). http://www.dell.com/learn/us/en/uscorp1/cr-social-responsibility?s=corp. Accessed 4 Apr 2018

Gelsomino, L.M., Mangiaracina, R., Perego, A., Tumino, A.: Supply chain finance: a literature review. Int. J. Phys. Distrib. Logist. Manag. **46**(4), 348–366 (2016)

Hofmann, E., Kotzab, H.: A supply chain-oriented approach of working capital management. J. Bus. Logist. **31**(2), 305–330 (2010)

Iansiti, M., Levien, R.: Strategy as ecology. Harvard Bus. Rev. **82**(3), 68–81 (2004)

Lind, L., Monto, S., Kärri, T., Schupp, F.: Detecting working capital models in the ICT supply chains. Int. J. Supply Chain Invent. Manag. **1**(3), 233–249 (2016)

Moers, F.: Doing archival research in management accounting. In: Chapman S.C., Hopwood, A. G., Shields, M.D. (eds.) Handbook of Management Accounting Research, vol. 1, pp. 399–413. Elsevier, The Netherlands (2007)

Moore, J.F.: The rise of a new corporate form. Washington Q. **21**(1), 167–181 (1998)

Richards, V.D., Laughlin, E.J.: A cash conversion cycle approach to liquidity analysis. Financ. Manag. **9**(1), 32–38 (1980)

Seifert, R.W., Seifert, D.: Working Capital in Times of Financial Crisis: Three Trade Credit Strategies. Perspectives for Managers, no. 166 (2008). 4 p.

Williamson, O.E.: The Economic Institutions of Capitalism. Free Press, New York (1985)

Wuttke, D.A., Blome, C., Henke, M.: Focusing the financial flow of supply chains: an empirical investigation of financial supply chain management. Int. J. Prod. Econ. **145**(2), 773–789 (2013)

Total Cost of Ownership for Asset Management: Challenges and Benefits for Asset-Intensive Organizations

Amir Noorbakhsh[1](✉), Ian Howard[2], Brett Kirk[2], and Kerry Brown[3]

[1] WA School of Mines, Curtin University, Bentley, Australia
amir.noorbakhsh@postgrad.curtin.edu.au
[2] School of Civil and Mechanical Engineering, Curtin University, Bentley, Australia
i.howard@exchange.curtin.edu.au, brett.kirk@curtin.edu.au
[3] School of Business and Law, Edith Cowen University, Joondalup, Australia
k.brown@ecu.edu.au

Abstract. Nowadays, optimal supplier selection process is essential for the performance success of business organizations. Total cost of ownership (TCO) is a method which enables purchasing firms to access all key information for assessing and selecting suppliers. An effective supply chain process which is based on TCO tool is argued to extend purchasing decisions beyond the initial price and consider the long-term perspective of all hidden and explicit costs for carrying out business with various suppliers. All purchasing enterprises are seeking the maximum profit from the purchased products/services. In order to identify all cost drivers, both quantitative and qualitative methodological approaches have been operated and analysed in the TCO application. However, it is not clear which methodologies are better, and under what circumstances, in relation to the TCO concept. This paper examines the application of a range of qualitative and quantitative methodologies in TCO purchasing techniques. The paper outlines and reviews the strengths, limitations and barriers of both qualitative and quantitative methods. In conclusion, the framework supports optimizing decision-making based on total cost of ownership in a firm's cost management process.

1 Introduction

The main objective of this study is to illustrate the implementation of qualitative and quantitative methods in TCO application models and survey all benefits and impediments of different methods over the entire decision-making process. Total Cost of Ownership (TCO) is a global approach, long-established in different types of businesses and professions. TCO is a reliable method which generates rigorous information not only for suppliers but also for purchasers to make the best decision regarding their products and services. TCO is a philosophy which considers elements beyond just the initial price of goods or services for the better realisation of cost management. These aspects of TCO are not restricted to only buyer and seller but contains service providers

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 200–208, 2020.
https://doi.org/10.1007/978-3-030-48021-9_23

and value-added sellers. This method notes all tangible and intangible costs including acquisition, administration, possession, operation, audit, refusal, displacement, maintenance, and disposal of a product or service. It is critical to emphasise that the most cost-effective seller or service provider would not necessarily be the one that has a minimum initial price.

For identifying all cost drivers, both quantitative and qualitative methodological approaches can be implemented and assessed in the TCO function. However, better methodologies regarding TCO concepts are still not transparent. This paper examines the implementation of a range of qualitative and quantitative methods in TCO purchasing techniques. The paper outlines and reviews the strengths, limitations and barriers to both qualitative and quantitative methods. In conclusion, the framework supports optimising decision-making based on total cost of ownership in a firm's cost management process. Critical performance elements, best methods and their implementation impediments in TCO concept and application are mostly analytical and predictive.

One of the most significant benefits of TCO is improved decision-making. A TCO study can illustrate the most expensive elements. Therefore, a productive TCO model is a model based on the capable and flexible decision-making process. It can help the high-level asset managers determine priorities and make better decisions by considering both qualitative and quantitative approaches. Both qualitative and quantitative methods consider tangible and intangible elements and make a trade-off between them to help the decision-maker in finding the most capable and proper supplier or service provider. This study reviews and highlights the benefits, restrictions and obstacles of both qualitative and quantitative approaches.

2 Research Method

Thorough literature review around qualitative and quantitative methodologies in TCO application is the first stage in this research. The literature review will define the limitations and barriers to both approaches in TCO implementation. One of the primary phases of every research approach is a determination of the proper methodology. Quantitative and qualitative are two main approaches to research. Selecting the most appropriate method typically relies on the type of data. Cooper and Schindler argue that quantity is the amount of something while the quality is the fundamental specification or kind of something (Cooper and Schindler 2003). This paper assesses implementation of both approaches in TCO application.

TCO implements the Activity-Base Costing concept for identification of the cost drivers. Hence, there should be a method for those cost elements recognition which has a higher impact on total cost of products or services. A proper technique including Quantitative, Qualitative or a mix of both must be deployed in the critical cost factor identification phase. The magnitude and scope of a TCO application may, however, vary from merely introducing the most notable cost factors employing the Pareto principle (Ellram 1994) to complex techniques implementing mathematical programming (Degraeve and Roodhooft 1999; Degraeve et al. 2000, Degraeve et al. 2005b).

The general research steps for TCO application are as below (Fig. 1);

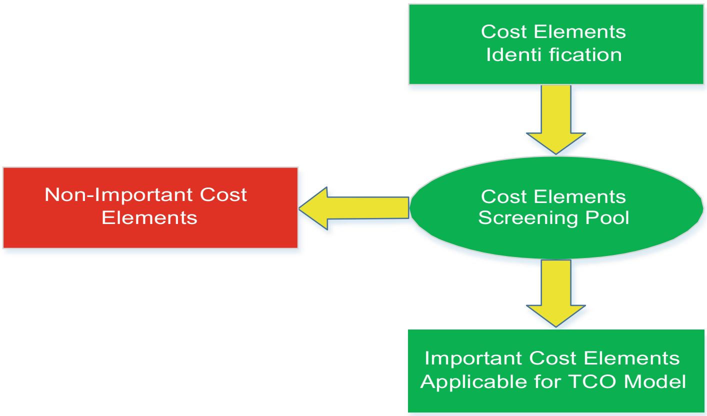


Fig. 1. Adopted cost elements identification (Ström and Bladh 2008)

1. Study TCO concept and related theories
2. Study of related business industry: Expert interview, literature, etc.
3. Study of proper approach: Qualitative, Quantitative, Combination, Case study, triangulation etc.
4. Collecting data: Qualitative, Quantitative
5. Validate the data
6. Calculation and analysis of the results
7. Identifying cost pivotal drivers
8. Prioritising cost factors
9. Identifying variables
10. Build the model
11. Test the model

3 Literature Review of Methods

Purchasing functions are a primary component of any business management approach and is one of the critical areas of the operational decision-making process which can quickly impact all divisions of a competitive business enterprise. The central part of the purchasing function is the supplier selection process which has encountered considerable evolution during the past two decades (Weber et al. 1991). The advanced global market requires qualified suppliers to produce high-quality products at reasonable price. Degraeve and Roodhooft believed that supplier selection decisions are a multi-objective decision (Degraeve et al. 2000). Vendor selection represents up to 70% of total production costs in manufacturing (Heberling 1993). The standard method of vendor selection frequently overlooked actual cost purchasing strategies including tangible and intangible cost (Mohammady Garfamy 2006).

In 1987, the Gartner Group presented the TCO analysis for the first time, and since then this concept has been utilised in the various research areas especially in automotive manufactures. Ellram was the first person to introduce the TCO concept as a supplier selection method (Ellram 1993a). She also studied other purchasing methods and compared them to TCO application and then examined TCO for vendor evaluation and selection. She contends that the total cost of ownership (TCO) is a useful buying tool for different businesses which contain both direct and indirect expenses including initial price, parts/equipment, operation, maintenance, and disposal (Ellram 1993a). Hurkens highly recommended the total cost of ownership method for specifying the hidden costs, as well as the visible expenses, of running a business with various suppliers (Hurkens et al. 2006).

Lisa Ellram classified TCO benefits as: improvement in supplier evaluation, decision-making, supplier and purchaser relationship, realisation and continuous progression (Ellram 1994). Ellram identified some limitations to TCO implementation including the complexity of TCO concept, cultural issue, education and training, data and resource restriction. She emphasised that TCO quantifies all expenses associated with the buying practices. There are different types of methods which are similar to TCO such as life-cycle costing, zero-based pricing and the cost-ratio process (Monczka and Trecha 1988; Ellram 1993b) but ignore the cost associated with the items from initiation to end (Ellram 1993a).

Degraeve et al. has researched TCO applications across different industry sectors. In his first paper, Degraeve presented a developed mathematical programming model which implemented TCO to choose the best vendor while specifying order quantities with the goal of rectifying the heating electrode purchasing system issue of a steel producer company in Belgium (Degraeve and Roodhooft 1999). He then highlighted the ignorance of purchasing performance by the typical cost management tools and proposed a decision making support system to reduce total costs (Degraeve and Roodhooft 1999). As mentioned before, TCO applies to both products and services. Hence, Degraeve applied the mathematical programming model of TCO for selecting not only airline company but also specifying companies' market share (Degraeve et al. 2004). In 2005, he built the model by combining the TCO with mathematical programming and using three case studies, and as a result, he proposed that a combination of methods provided improved functionality than those methods separately (Degraeve et al. 2005a).

Bhutta and Huq (2002) compared a total cost of ownership (TCO) with analytic hierarchy process (AHP) and highlighted benefits and restrictions of both methods. Ellram indicated that TCO and activity-based costing alliance could create an enterprise with powerful competitive benefit from data associated with supply chain costs (Ellram 1995a). Later, a logistics outsourcing structure was developed by Maltz and Ellram (1997). They built the model which encompassed non-monetary elements and claimed that concentration only on quantitative expenses could lead to insufficient decision making. In 1998, Ellram studied TCO as a strategic cost management tool through several case studies and investigated the linkage between strategic cost management (SCO) and TCO (Ellram and Siferd 1998).

Type of data is the pivotal element for selecting the most appropriate TCO method. Neuman argued that qualitative research stresses explanatory and critical approaches while quantitative research refers to objectivity and technical styles (Neuman and Kreuger 2003). Qualitative methods underline the meaning, the interpretation or analogy characterising something, on the other hand, the quantitative way considered the content and aims of its measurement (Cooper and Schindler 2003). The interview is one of the best observation assessment tools when using theories, assumptions, or problem perceptions in a qualitative study (Simon 2006).

Sarkis was one of those who conducted quantitative research on strategic vendor selection and who presented an approach based on an analytical network process-based (Sarkis and Talluri 2002). Their research was an experimental case study including various criteria with the participation of managerial decision-makers level. Ellram undertook investigations through many case studies and a literature review to highlight the problem of vendor selection (Ellram and Siferd 1998).

Case studies highlight recognition and comprehensive investigation of the facts in a specific group. Generally in a case study, after choosing the typical and critical cases, a detailed analytical examination will be implemented to reach to the mean general model or pattern among multiple typical instances (Neuman and Kreuger 2003). A developed matrix model which performing pre-arranged benchmarking for weighting vendor selection was presented by Gregory to minimise the vendor rating individuality of every component (Gregory 1986). Operation Management Strategy (OMS) employed by Chou and Chang for vendor selection criteria classification with the aim of finding a solution for strategic vendor selection issue, they utilised Simple Multi-Attribute rating technique (SMART) based on Fuzzy Algorithm (Chou and Chang 2008).

There are other approaches in qualitative and quantitative which are implemented for alternative vendor evaluation. Supplier's design performance was the central element for Araz and Ozkarahan to evaluate supplier capabilities and classified vendors regarding their overall function (Araz and Ozkarahan 2007). The preference ranking organisation method for enrichment evaluation (PROMETHEE) methodology was operated for revealing the reason for suppliers different performance.

Florez-Lopez utilised fuzzy theory for ranking the vendors based on their competency to set up the value for the customers (Florez-Lopez 2007). A combination of analytic hierarchy process (AHP) and quality management system (QMS) was developed by Chan and Chan to prioritise and select the vendors in the industry of advanced technology (Chan and Chan 2004). Sucky found that the standard vendor selection overlooked the costs associated with choosing a new supplier or switching from one provider to another provider (Sucky 2007).

The hardest part of cost quantification is coming to an accurate estimation of the hidden cost of goods or services throughout its entire lifecycle. Ellram found that indirect expenses could far outrun the initial purchase price or apparent values (Ellram 1993b, Ellram 1993a, Ellram 1994, Ellram 1995b, Ellram 1995a). It is clear that TCO encompasses an enormous variety of cost items and this can make the calculation much more complicated. Hence TCO relies on those cost drivers which are very pivotal and sensitive to the purchaser firms, that's why the implementation of TCO expect experts from both parties, purchaser and supplier, to cooperate with each other for mapping and addressing their clients' goal and expectation.

4 Discussion

The qualitative approach is generally utilised when an examination of the phenomena or the relations must be assessed. In this approach, the type of available data is soft. While the quantitative methods are not highly adaptable to changes in research, the qualitative practices are open to conversion and are more transformative. For this reason, researchers prefer to apply this approach especially when they need precise description in their research. This method is based on the researchers own judgment and perception which can influence the research questions. Lack of generalisation due to the type of collected data is one of the weak points of the qualitative method. It is also very helpful for comprehension of the phenomena and also is very flexible and adaptable in the research process (Johannesen 2004; Jakobsen and Staavi 2009).

The objective of the quantitative approach is to collect data that are measurable and to assess these received data by implementing statistical tools or programs. In this method, a generalisation of numerical and analytical outcomes of the research is favourable. Sometimes representative data are not enough and will be a challenging part for the researchers. Shortage in opportunities for a complete and comprehensive statement of the investigated phenomena is not applicable while this method is constructive with valuable outcomes if the gathered representative data is enough.

When a combination of qualitative and quantitative methods is implemented, triangulation can be helpful especially for the research where checking the outcomes should be considered more than one time. In 1994, Yin stated that there is not any restriction while working with a mix of qualitative and quantitative evidence in case studies (Yin 1994). Combination of these methodologies enriches results accuracy due to considering various views and angles. The combination method can be more reliable if the each of the methods present a similar conclusion which includes valuable outcomes such as adequate explanation, more reliable statements and more comprehensive description. The application of methodological triangulation has three classifications: qualitative assessment of a series of quantitative data, qualitative pre-arrangement of a quantitative data collection and finally implementing both methods at the same time.

Selecting the optimal data collection method is the next stage of the process after data requirements and analytical approach determination. Zikmund classified types of data in data collection stage, primary data and secondary data (Zikmund 1997). The primary data is obtained by the researcher during their study through interviews, questionnaires, observations and careful analyses while secondary data is collected by other researchers and can be noticed in different sources of information like literature, papers and reports. Collecting primary data is a time-consuming matter while more accurate and specific. On the other hand, collecting secondary data is much easier but non-specific and can be out of date. Johannesen argued that selecting the proper approach is mostly based on the research questions (Johannesen 2004). If the types of the research questions involve "why" or "how" questions, the case studies can be useful approach especially for the critical situation or events with limited control by the researcher (Yin 1994).

Qualitative approaches can contain tools for simulating and examining the judgment of decision-makers regarding issue condition and tools for deliberating about potential solutions. The set of quantitative approaches encircle a wide variety of methods. Data mining techniques are one of the useful techniques for evaluating comparable decisions made in the former studies to obtain a general model and decision guideline which may be implemented in future for improvement of the efficiency and performance of following findings. The optimisation method is another quantitative method which hand decision-makers for reducing total cost by finding the optimal solutions to economic issues, linear programming is one example of optimisation methods. An exploratory research method including surveys of companies and manufacturers within different industries and case studies of various organisations can be operated. Data collection can be implemented through interviews with purchasing experts, asset managers.

5 Conclusion

Selecting an appropriate methodology is the first step of any research approach development. Exploratory research is commonly carried out when the issue, scope or research design and framework is not well established and needs to be developed appropriately. While qualitative research is objective research, quantitative analysis is a more deductive study. In qualitative research, the hypothesis is more assessed by predetermined frameworks. In quantitative studies, research implements a detailed statistical approach. According to Creswell, for TCO cost strategies application, qualitative research which has the capability in framework development can be the better approach (Creswell 2002). There is another discussion that for TCO application and to capture a highly reliable and valid data, implementation of both qualitative and quantitative simultaneously can be a better option. Hence, the purchaser can make a trade-off between conflicting qualitative and quantitative elements for identifying a list of proper suppliers, classifying and prioritising desired suppliers and finally selecting the best supplier.

In the process of identifying activities and cost drivers, an activity-based costing approach can be constructive. TCO contains a range of cost elements. TCO by nature is a complex concept, and no need to make it more complicated by considering every non-important cost items which don't have any magnitude or any impact on final TCO model. Thus, TCO is associated with user's judgment. According to Pareto's principal, the concept is to consider the 20% of the components that responsible for the 80% of the total costs. One integral element in supplier selection is a risk of the supplier. There are several arguments regarding quantifying the supplier risk and take it to account in TCO application. Some risk data are measurable while others can be risks that may be non-trackable or very subjective. Hence a method which considers both qualitative and quantitative data is proposed for evaluating the risk of the supplier.

References

Araz, C., Ozkarahan, I.: Supplier evaluation and management system for strategic sourcing based on a new multicriteria sorting procedure. Int. J. Prod. Econ. **106**, 585–606 (2007)

Bhutta, K.S., Huq, F.: Supplier selection problem: a comparison of the total cost of ownership and analytic hierarchy process approaches. Supply Chain Manag.: Int. J. **7**, 126–135 (2002)

Chan, F.T., Chan, H.: Development of the supplier selection model—A case study in the advanced technology industry. Proc. Inst. Mech. Eng. Part B: J. Eng. Manuf. **218**, 1807–1824 (2004)

Chou, S.-Y., Chang, Y.-H.: A decision support system for supplier selection based on a strategy-aligned fuzzy SMART approach. Expert Syst. Appl. **34**, 2241–2253 (2008)

Cooper, D., Schindler, P.: Business Research Methods, 8th edn. McGrawHill, New York (2003)

Creswell, J.W.: Educational Research: Planning, Conducting, and Evaluating Quantitative. Prentice Hall, Upper Saddle River (2002)

Degraeve, Z., Labro, E., Roodhooft, F.: An evaluation of vendor selection models from a total cost of ownership perspective. Eur. J. Oper. Res. **125**, 34–58 (2000)

Degraeve, Z., Labro, E., Roodhooft, F.: Total cost of ownership purchasing of a service: the case of airline selection at Alcatel Bell. Eur. J. Oper. Res. **156**, 23–40 (2004)

Degraeve, Z., Labro, E., Roodhooft, F.: Constructing a total cost of ownership supplier selection methodology based on activity-based costing and mathematical programming. Account. Bus. Res. **35**, 3–27 (2005a)

Degraeve, Z., Roodhooft, F.: Effectively selecting suppliers using total cost of ownership. J. Supply Chain Manag. **35**, 5–10 (1999)

Degraeve, Z., Roodhooft, F., van Doveren, B.: The use of total cost of ownership for strategic procurement: a company-wide management information system. J. Oper. Res. Soc. **56**, 51–59 (2005b)

Ellram, L.: Total cost of ownership: elements and implementation. J. Supply Chain Manag. **29**, 2–11 (1993a)

Ellram, L.: A taxonomy of total cost of ownership models. J. Bus. Logist. **15**, 171 (1994)

Ellram, L.M.: A framework for total cost of ownership. Int. J. Logist. Manag. **4**, 49–60 (1993b)

Ellram, L.M.: Activity-based costing and total cost of ownership: a critical linkage. J. Cost Manag. **8**, 22–30 (1995a)

Ellram, L.M.: Total cost of ownership: an analysis approach for purchasing. Int. J. Phys. Distrib. Logist. Manag. **25**, 4–23 (1995b)

Ellram, L.M., Siferd, S.P.: Total cost of ownership: a key concept in strategic cost management decisions. J. Bus. Logist. **19**, 55 (1998)

Florez-Lopez, R.: Strategic supplier selection in the added-value perspective: a CI approach. Inf. Sci. **177**, 1169–1179 (2007)

Gregory, R.E.: Source selection: a matrix approach. J. Supply Chain Manag. **22**, 24–29 (1986)

Heberling, M.E.: The rediscovery of modern purchasing. J. Supply Chain Manag. **29**, 47–53 (1993)

Hurkens, K., Valk, W., Wynstra, F.: Total cost of ownership in the services sector: a case study. J. Supply Chain Manag. **42**, 27–37 (2006)

Jakobsen, K., Staavi, S.U.: Proposal of a global Total Cost of Ownership Model for FMC Technologies' suppliers. Universitetet i Agder/University of Agder (2009)

Johannesen, A.: Per Arne Tufte og Line Kristoffersen 2. utg. (2004). Introduksjon til samfunnsvitenskapelig metode

Maltz, A.B., Ellram, L.M.: Total cost of relationship: an analytical framework for the logistics outsourcing decision. J. Bus. Logist. **18**, 45 (1997)

Mohammady Garfamy, R.: A data envelopment analysis approach based on total cost of ownership for supplier selection. J. Enterp. Inf. Manag. **19**, 662–678 (2006)

Monczka, R.M., Trecha, S.J.: Cost-based supplier performance evaluation. J. Supply Chain Manag. **24**, 2–7 (1988)

Neuman, W.L., Kreuger, L.: Social Work Research Methods: Qualitative and Quantitative Approaches. Allyn and Bacon, Boston (2003)

Sarkis, J., Talluri, S.: A model for strategic supplier selection. J. Supply Chain Manag. **38**, 18–28 (2002)

Simon, M.: Recipes for success. Dissertation & Scholarly Research, Dubuque, IA (2006)

Ström, A., Bladh, A.: Total Cost of Ownership-Revealing the true cost of owning and operating equipment (2008)

Sucky, E.: A model for dynamic strategic vendor selection. Comput. Oper. Res. **34**, 3638–3651 (2007)

Weber, C.A., Current, J.R., Benton, W.: Vendor selection criteria and methods. Eur. J. Oper. Res. **50**, 2–18 (1991)

Yin, R.K.: Case Study Research: Design and Methods, Applied Social Research Methods Series, vol. 5. Sage Publications, London (1994)

Zikmund, W.: Business Research Methods, 5th edn. Harcourt Brace, Fort Worth (1997)

On the Expected Monetary Value of Hydroelectric Turbine Start-up Protocol Optimisation

Martin Gagnon(✉), Denis Thibault, and Michel Blain

Institut de recherche d'Hydro-Québec (IREQ), Varennes, QC, Canada
{gagnon.martin, thibault.denis, blain.michel}@ireq.ca

Abstract. Hydroelectric turbine start-up protocols have a significant influence on the fatigue reliability of such asset. Fatigue cracking, is one of the two main degradation mechanisms and often considered the most difficult to predict. Using fatigue reliability model previously developed by the authors, we study the influence of changes in the start-up protocol. Two units are studied. For the first unit, two protocols are studied and for the second unit three protocols are considered. The goal of the study is to look at the expected monetary value (EMV) as a metric able to combine the expected probability of failure with the monetary consequence of such failure. Our results shows that such metric is suitable to identify which start-up protocol is the most advantageous and why?

1 Introduction

Fatigue and cavitation are the two main degradation mechanisms of hydroelectric turbine runners. Of the two, fatigue cracking is often considered more difficult to manage since the damage cannot be monitored until significant cracks appear. These cracks upon detection are automatically considered critical and have to be repaired. Hence, the probability of fatigue failure is defined as the probability of detecting a crack during a given inspection (Gagnon et al. 2013a). This probability is a function of both the expected values and uncertainties associated with loading, defect size and material properties. When comparing monetary gains related to the optimisation of the runner blade loading during start-up, not only the failure probability but also costs and monetary impacts over time need to be taken into account. To this end, the expected monetary value (EMV) is a simple metric that enables one to combine the probability and the consequence in term of economic value. Using this metric, to identify a solution that maximizes our expected monetary gains, we can properly account for the combined cost associated with implementing a given scenario, the economic impacts of the event and the probability of this event.

In this study, we are looking at the implementation of an optimised start-up protocol with the goal of minimizing fatigue cracking probability (Gagnon et al. 2010; Gagnon et al. 2014a, b; Gagnon et al. 2016). The start-up protocols have no influence on the initial fatigue reliability of the runner but will influence the cumulative contribution of each unit start-up over time. Usually, the initial reliability is not 100% which is related to the probability of being dead on arrival due to the lack of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 209–216, 2020.
https://doi.org/10.1007/978-3-030-48021-9_24

information about the runner's initial state. To evaluate these start-up protocols, we propose a simple methodology to quantify EMV using study cases based on data from Hydro-Québec.

The paper is structured as follows. First, an overview of the reliability model with the influence of both expected stresses and uncertainty on turbine runner fatigue reliability is presented. Then, the study cases and the methodology are detailed. Finally, we will discuss the results in terms of applicability for decision making and asset management.

2 Hydroelectric Turbine Reliability

In previous work (Gagnon et al. 2013a, b; Gagnon et al. 2014a, b), we defined that, for any given point in time, the criteria for hydroelectric turbine runner fatigue reliability was that every allowed steady state operation should have a stress range below the high cycle fatigue (HCF) onset threshold. This means that, as the defect grows, the stress range should stay below the HCF onset threshold formed by the Kitagawa diagram (Kitagawa and Takahashi 1976). This limit state is illustrated in Fig. 1(left). In its simplest form, the hydroelectric turbine runner stress spectrum is composed of at least a high amplitude low cycle fatigue (LCF) component which contributes to crack propagation combined with an HCF components which should not contribute to propagation. The LCF component represents the overall stress range generated by the start to stop cycle and the HCF component represents the stress range during steady state operation. This simplified load spectrum is shown in Fig. 1(right). Notice the arrow in Fig. 1(left) which represents the movement toward the limit state of the joint distribution formed by the HCF stress range $\Delta\sigma$ and defect size a as the LCF component contributes to the crack propagation. This growth is a function of both the number of start to stop cycles the hydroelectric turbine undergoes and the LCF components stress range.

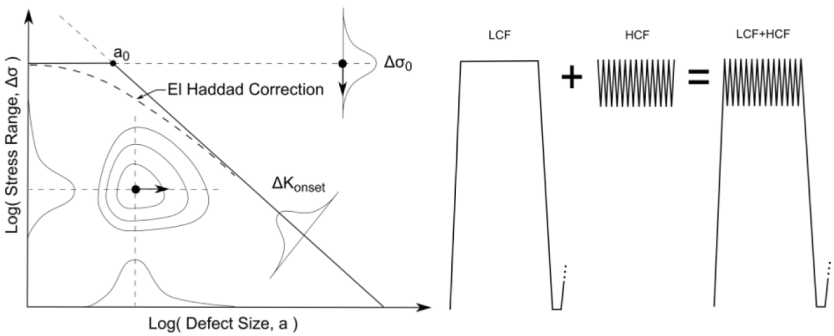


Fig. 1. (Left) Fatigue reliability limit state. (Right) Structure response spectrum.

The Kitagawa diagram combines two limits (The fatigue limit $\Delta\sigma_0$ and the HCF onset threshold ΔK_{onset}) that are joined together using the El Haddad correction (Haddad et al. 1979). The limit state is expressed as follows:

$$g(a, \Delta\sigma) = \Delta\sigma - \frac{\Delta K_{onset}}{\sqrt{\pi(a+a_0)}Y(a+a_0)} \quad (1)$$

$$a_0 = \frac{1}{\pi}\left(\frac{\Delta K_{onset}}{\Delta\sigma_0 Y(a_0)}\right) \quad (2)$$

Often, the crack growth threshold ΔK_{th} is used in place of ΔK_{onset} because it is more readily available. From Eq. 1, we obtain the probability of failure P_f by solving the following:

$$P_f = \int_{g(x)\le 0} f_X(x)dx \quad (3)$$

in which, x is an n-dimensional vector of random variables with a joint density function $f_X(x)$. This value is approximated using First Order Reliability Methods (FORM) as describe in (Gagnon et al. 2013a). The results obtained can either be expressed in terms of the reliability index β or probability of failure P_f using:

$$P_f = \Phi(-\beta) \quad (4)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Notice that from this definition, cracking will occur when the largest defect in a given volume propagates due to the largest HCF stress cycle of the stress spectrum above the limit state. Both defect size and stress cycle range cannot be known exactly and only their largest possible value is of interest hence the uncertainties around both values are modelled using an extreme value distribution. In our case, we have chosen to use the Gumbel distribution for simplicity purpose. Futhermore, the location of the limit state is also uncertain. Through this paper, the cracking probability is defined as the probability of entering the most probable region where the limit state is expected and whose parameters are defined in (Thibault et al. 2015). This is less conservative than using the lower limit of the limit state uncertainty interval which represents the limit below which we believe that failure is not possible (Gagnon et al. 2013b).

3 Start-up Protocol and Experimental Results

The wicket gates opening is the main parameter which regulates the operation during transient events. The speed governor typically has only a limited number of adjustable parameters for safety purpose. The typical adjustable parameters for the wicket gates opening sequence are: opening limit, fold back speed and fold back opening. Notice that outside those specific parameters, the governor operation could be modified by the manufacturer or plant operator at a relatively low cost. Governor actions are triggered by the rotating speed. The aim is to reach synchronous speed in an acceptable time interval for the unit synchronization as defined in IEC 61362 (IEC 2012). The opening sequence is as follows: first the gates open with a prescribed speed until they reach an opening limit then, after the rotating speed reaches the fold back speed, the wicket gates are limited to a fold back opening which is maintained until near synchronous speed for synchronization with the grid. Since fast synchronization is not the main objective, optimization possibilities are available in the start-up sequence parameters. Data measured on two different hydroelectric turbines are used in this study. For Unit 1, we compare two different start-up protocols. The blade response over the whole start-to-stop cycle is presented for each protocol in Fig. 2.

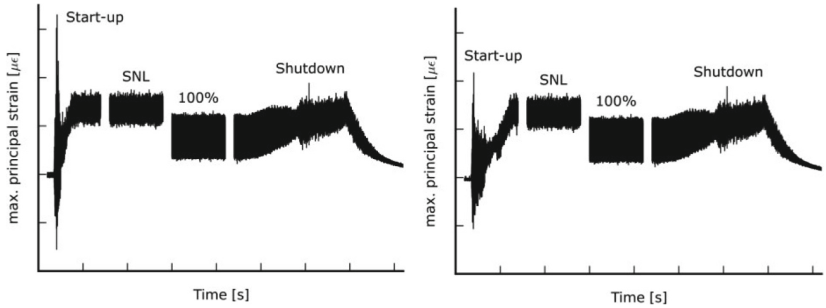


Fig. 2. Startup for Unit 1. (Left) Protocol 1. (Right) Protocol 2.

In Fig. 3, on the other hand, we show the response to three different start-up protocols tested on Unit 2. Here, the opening, rotating speed and blade response are presented for each protocol. This gives a more detailed view of the start-up protocol. Since, we consider only one LCF stress cycle in our simplified load spectrum, only the maximum peak-to-peak stress range during start-up is considered in this study.

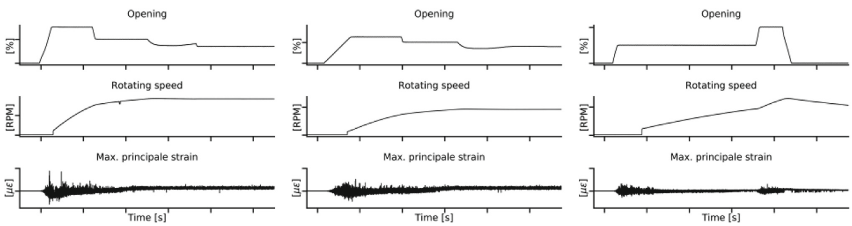


Fig. 3. Start-ups for Unit 2. (Left) Protocol 1. (Middle) Protocol 2. (Right) protocol 3.

4 Expected Monetary Value

To combine financial and probabilistic information in the context of decision making, the EMV is often used as a simple quantitative metric. This metric expresses the gain expected if an event of a given probability was repeated an infinite number of times. The EMV for a given decision is expressed as follows:

$$EMV = \sum P(Event) \times Monetary\ value\ of\ the\ event \tag{5}$$

In the case where one needs to choose between two solutions for which the sum of the implementation costs and monetary impacts are equal, we can compare the difference in probability multiplied by the combined cost and impacts of such series of events. Furthermore, if we account for the costs and impacts similarly across projects, the solutions can be compared with each other's. For start-up protocol, we have considered identical implementation costs and consequence for a given unit. Hence, for this unit, only the probability of the failure event will change. We have accounted for time by discretizing the time horizon of the study T using the conditional probabilities on a series of smaller time interval Δt. The probability of a failure event in a given time interval $[t, t + \Delta t]$ if we know that the component is functional at time t is expressed as follows:

$$P(t < T \leq t + \Delta t | T \rangle t) = \frac{F(t + \Delta t) - F(t)}{1 - F(t)} \tag{6}$$

where $F(t)$ is the probability of failure obtained in Eq. 4. To assess the expected gains from the start-ups protocol optimization, the annual EMV on a 12 years horizon was chosen. The monetary impacts details cannot be disclosed and are considered confidential. However, these monetary impacts account for availability loss due to down time and maintenance costs. Let's mention that in this study availability loss due to down time is the main contributor to the monetary impacts. Furthermore, all the results are presented as present values to account for inflation and discount rate.

5 Results

The reliability indexes obtained as a function of time are presented in Figs. 4 for each of the start-up protocol studied. Higher reliability with time is observed for protocols generating lower stress range. In the case of Unit 2, a greater improvement across the three studied solutions is observed compared to Unit 1.

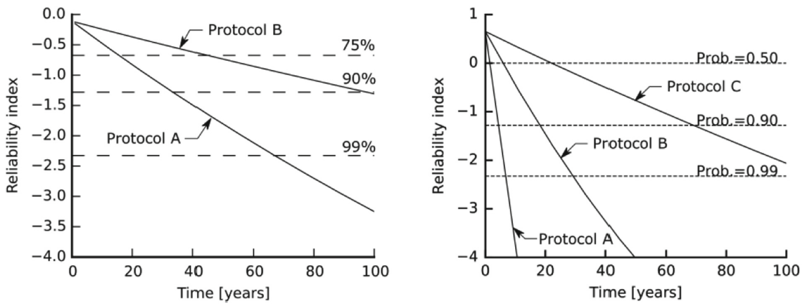


Fig. 4. Reliability indexes as à function of time. (Left) Unit 1. (Right) Unit2

From the reliability indexes, the probabilities of failure are estimated to obtain the annual conditional probabilities. Using the difference in annual conditional probability, we obtain the annual EMV of choosing a given start-up protocol compared to another one. The annual conditional probabilities over the chosen twelve years horizon are presented in Fig. 5 (left).

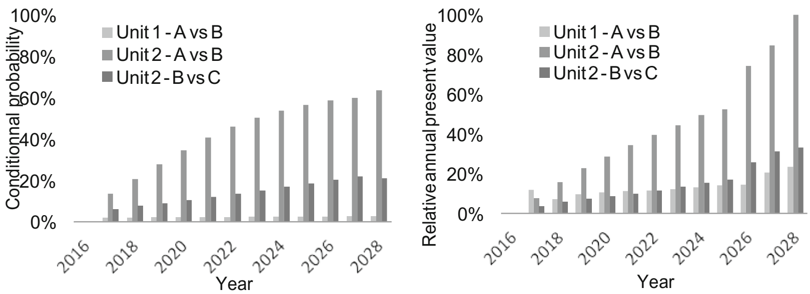


Fig. 5. (Left) Conditional probability difference. (Right) Relative annual present value.

Using the annual conditional probabilities difference in Fig. 5(left) and the failure monetary impact, we can assess the annual EMV which are presented in Fig. 5(right) as present value to account for inflation and discount rate. The results have been normalised and percentage of the maximum value are shown. We observe that for Unit 1 the protocol change from A to B as the largest initial annual EMV. However, for this unit, the annual EMV stay relatively constant with time when compared to Unit 2. This is observed because, the failure consequences for Unit 1 are higher and the annual conditional probability differences increase more slowly compered to Unit 2.

6 Discussions

We should mention that Unit 2 start-up protocol A generates unusually high stress in the runner blade which is not representative of a typical runner. Notice that since such high stresses have been observed this illustrates the large gain in annual EMV that should be expected from moving to more typical values using start-up protocol B. Nonetheless, if we look at Unit 2 start-up protocol B vs C, we observe that even if the values are small initially the probability differences increase every year which is not negligible for longer time horizon. This effect is not as important in Unit 1 for which the conditional probability differences are small and the failure monetary impacts are more important. We should mention that annual EMV do not represent a possible monetary gain realisation and be careful with the conclusions extracted from such results. Furthermore, we only account for uncertainty in the evaluation of event probability. All the financial information are deterministic. While such an approach generates an appropriate relative metric in the context of this study, it could generate false confidence if uncertainties in the economic parameters are large. The annual EMV seems only appropriate as a metric to compare projects and solutions with each other's. As an assessment of how much money could be gained or lost, the methodology only provides qualitative information: a more complete analysis taking into account probabilistic financial inputs and the quantification of risks like the probability of losing money has to be realized. Decision makers should be warned of the risks involved and the assumption made before assessing how much money should be allocated for a given solution.

7 Conclusions

The objective of this paper was to look at EMV as a metric to compare different start-up protocols by accounting for probability of failure and monetary cost. This study has showed that such metric while simple to implement is able to account for a width array of parameters. However, such study should be followed by sensitivity analysis or proper probabilistic modelling before decision making. This even more so when small differences are observed between projects.

References

Gagnon, M., Jobidon, N., Lawrence, M., Larouche, D.: Optimization of turbine startup: some experimental results from a propeller runner. In: 27th IAHR Symposium on Hydraulic Machinery, Montréal, Canada (2014a)

Gagnon, M., Nicolle, J., Morissette, J.-F., Lawrence, M.: A look at Francis runner blades response during transients. In: 28th IAHR Symposium on Hydraulic Machinery and Systems, 4–8 July 2016, Grenoble, France (2016)

Gagnon, M., Tahan, A., Bocher, P., Thibault, D.: A probabilistic model for the onset of High Cycle Fatigue (HCF) crack propagation: application to hydroelectric turbine runner. Int. J. Fatigue, **47**, 300–307 (2013a)

Gagnon, M., Tahan, A., Bocher, P., Thibault, D.: On the fatigue reliability of hydroelectric Francis runners. Procedia Eng. **66**, 565–574 (2013b)

Gagnon, M., Tahan, A., Bocher, P., Thibault, D.: Influence of load spectrum assumptions on the expected reliability of hydroelectric turbines: a case study. Struct. Saf. **50**, 1–8 (2014b)

Gagnon, M., Tahan, S.A., Bocher, P., Thibault, D.: Impact of startup scheme on Francis runner life expectancy. In: 25th IAHR Symposium on Hydraulic Machinery and Systems, 20–24 September 2010, Timisoara, Romania (2010)

Haddad, M.E., Topper, T., Smith, K.: Prediction of non propagating cracks. Eng. Fract. Mech. **11** (3), 573–584 (1979)

IEC: Guide to specification of hydraulic turbine governing systems. IEC 61362, Éd. 2.0, 2012-04, 138 p. (2012)

Kitagawa, H., Takahashi, S.: Applicability of fracture mechanics to very small cracks or the cracks in the early stage. In: ASM Proceedings of 2nd International Conference on Mechanical Behaviour of Materials, Metalspark, Ohio (1976)

Thibault, D., Gagnon, M., Godin, S.: The effect of materials properties on the reliability of hydraulic turbine runners. Int. J. Fluid Mach. Syst. **8**(4), 254–263 (2015)

Economic Life-Cycle Indicators for Public School Buildings

Filipa Salvado[1(✉)], Nuno Almeida[2(✉)], and Álvaro Vale e Azevedo[1(✉)]

[1] National Laboratory for Civil Engineering (LNEC), Lisbon, Portugal
{asalvado,ava}@lnec.pt
[2] Instituto Superior Técnico (IST), University of Lisbon, Lisbon, Portugal
nunomarquesalmeida@tecnico.ulisboa.pt

Abstract. Costs prediction throughout the life cycle of building projects is supported by the life-cycle-cost (LCC) concept, encouraged by several international and regional standards (e.g. ISO 15686-5, EN 15643-4, EN 16627), procurement guidelines and regulations (e.g. the European Directive 2014/24/EU). However, public procurers of building projects still face difficulties regarding the costs estimation and time ratios of systems and components over the life-cycle of building facilities. Public databases with the adequate quantity and quality of economic information is needed. The existing ones often present problems such as those of inadequate data granularity, incompleteness, inaccuracy or data structures with formats that make comparison and extrapolation difficult. This paper addresses these problems and proposes a standardized structure for economic data collection throughout the whole life-cycle of building as well as economic life-cycle indicators. A case study related to 130 public school buildings constructed in Portugal since 1955 is presented as to show the potential application of the proposed structure towards widespread use of the LCC concept in public procurement environments.

1 Introduction

Over their life cycle, constructed assets demand for a considerable amount of resources and trigger transformations with important economic consequences. The economic performance of buildings may be assessed at an early stage along with the decision to initiate the construction project, or at any point afterwards, e.g. when scenarios are considered for supporting decisions such as refurbish, re-new, expand, retain or demolish. Information is needed to support these assessments.

The Life Cycle Cost (LCC) approach became popular in the 70s and since then it has been influencing the Architectural, Engineering and Construction (AEC) sector (Cole and Sterner 2000; Gluch and Baumann 2014; Meckler 1977). Its development involved gradual changes (Flanagan and Norman 1989; Goh and Sun 2016; Langdon 2007), finally comprising all key phases of the life cycle of any project or facility (Ludvig et al. 2010).

The cost estimates over the life cycle of a building is considered as one of the most important and critical project management activities. The preparation of reliable and

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 217–225, 2020.
https://doi.org/10.1007/978-3-030-48021-9_25

accurate estimates, are key factors in decision making processes (Jrade and Alkass 2007). One of the main challenges in this regard lies in organizing and deriving reliable and useful information from data, which can then be distilled into applicable knowledge. This can be achieved with comprehensive indicators that are widely accepted and easily understood by all relevant stakeholders in building projects or facilities.

The present paper aims to define economic life-cycle indicators based on historical data of public school buildings. These indicators are designed to assist lifecycle costs estimates within public procurement environments.

2 LCC and Public Procurement in the AEC Sector

Recently, the European Directive 2014/24/EU, the International standard ISO 15686 and the European standards EN 15643-4 and EN 16627 have jointly established the foundations for enhancing LCC applications. But there are still challenges to overcome, namely those related to the reliability of economic data (Arja et al. 2009; Ive 2006; Jrade and Alkass 2007; Ma et al. 2016).

As a response to recent standards and European Union acts procurers need to use life-cycle cost methodologies to (IISD 2016): i) design economically, environmentally and socially preferable tender specifications (including both mandatory and optional elements that suppliers are required to meet); ii) develop indicators on which bids will be appraised; iii) justify the purchase of environmentally and socially preferable alternatives that may require higher purchasing or initial costs although delivering the best value for money across the life cycle; iv) determine the “need to purchase” and subsequently discern between the outright purchase of an asset and the option of contracting services that would fulfil this need; v) evaluate the costs and benefits of designing building and infrastructure projects as private finance initiatives; vi) establish energy performance contracting agreements (contractual agreements that oblige developers to build and refurbish in such a manner that specified levels of energy costs savings can be guaranteed); and vii) establish cooperative contracts (agreements between buyers and vendors designed to reduce the administrative burden of contracting, while leveraging volume to negotiate preferred pricing and service for members of a group or consortium).

Nowadays, procurers around the world adopt a variety of approaches, formats and cost breakdown structures in their LCC analyses, the most quoted in the literature being those established in ISO 15686-5. Procurers appeared to be using it to determine what could be counted when performing LCC analyses, while specialists were adapting it for the assessment of non-construction costs such as land, finance and user-support charges, income-related costs such as income from sales and potential third party income during operation, and life cycle costs such as construction, operation, maintenance and end-of-life disposal. However, most procurers do not follow a rigorous methodology in their LCC analysis. It is also noted that LCC analysis are seldom accompanied by risk assessments and sensitivity analyses of key parameters that are likely to change with time (IISD 2016; OGC 2007).

An aspect that challenges procurers is the calculation of costs and time ratios related to interventions throughout the life cycle of infrastructure and building assets (e.g. maintenance, repair and replacement). Procurers needs dedicated databases that provide benchmarks, but data in the public domain is often too aggregated, incomplete or is presented in formats that make comparison and extrapolation difficult (IISD 2016).

3 Standardized Buildings Life-Cycle Economic Data

ISO 15686-5 and EN 16627 specifies a LCC calculation method and details the building life cycle information required. It also includes the rules for calculating the cash flows over the life cycle of buildings and defines a structure for the collection of economic data.

This structure intends to simplify the organization of the different costs involved in a building life cycle. It organizes the economic information into modules corresponding to the different life cycle stages. Table 1 summarize a standardized base for LCC calculations that can be applied to all scope of buildings costs (EN 15643-4 and EN 16627).

Public procurers can also use it to performing sensitivity and risk analyses of key parameters. The end of life stage (module C) is not considered in the scope of this study.

4 Case Study

4.1 Characterization

The case study covers a sample of 130 public school buildings (SB) constructed in Portugal between 1955 and 1988. The life cycle period of analysis starts with the land acquisition and ends in 2008. After 2008, the buildings were totally refurbished. This marks the beginning of a new life-cycle period that is out of the scope of this paper. Table 2 present some relevant parameters of the present case study.

The technical and functional characteristic of these 160 school buildings can be grouped in 3 types (A, B and D), which together represent 96% of the Portuguese school buildings. Type A use mixed construction technologies, based on resistant walls of ordinary masonry of plastered stone on which are attached the floor slabs and reinforced concrete stairs. The roof cover of Type A buildings uses wooden structures and coatings are predominantly wood in the floors of the classrooms and hydraulic mosaic in the corridors. On the walls is applied sand mortar and binder or stucco. The window frames are made of wood or prefabricated reinforced concrete elements with single glass. Types B and C present a structure of reinforced concrete with slabs of the same material and brick walls, plastered and painted. The coverings are flat or inclined coated with fibber cement elements. The frames are in wood (type B) or aluminium (type C) with a single glass. Type B normally adopted structural prefabricated elements.

Table 1. Before use stage (Module A) and Use stage (Module B) – standardization of building economic data

Cost category	Scope of costs
A0 – Land and associated fees/advice	Land acquisition; Professional fees; Taxes; Subsidies and incentives
A1 – Raw material supply	Supply of materials; Temporary works; Professional fees; Taxes; Subsidies and incentives
A2 – Transport (product)	Transport of products and material; Temporary works; Professional fees; Taxes; Subsidies and incentives
A3 – Manufacturing	Manufacturing of construction products; Temporary works; Professional fees; Taxes; Subsidies and incentives
A4 – Transport (construction process)	Transport of products, material, equipment and waste; Temporary works; Professional fees; Taxes; Subsidies and incentives
A5 – Construction installation process	Construction of asset/construction activities; Initial adaptation or refurbishment of asset; Accessibility of the site during construction; Landscaping, external works on the curtilage; Storage of products; Assembly, installation, commissioning and waste management; Provision of heating, cooling, ventilation, humidity control etc. during the construction process; Purchase or rental of the site; Temporary works (clear, prepare the site for construction and provide gas, electricity and water); Professional fees; Taxes; Subsidies and incentives
B1 – Use	Rent; Insurance; Cyclical regulatory costs; Utilities; Taxes
B2 – Maintenance	Costs related to all components and products used in maintenance activities; Cleaning; ground maintenance; Redecoration; Building-related facility management; Building-related insurance; Rent and lease; Third party income; Cyclical regulatory costs; End of lease and disposal inspections; Taxes; Subsidies and incentives
B3 – Repair	Repairs of minor components/small areas; Waste management; Taxes; Subsidies and incentives
B4 – Replacement	Replacement of minor components/small areas; Replacement of major systems and components; Revenue from sale of asset or elements; Waste management; Taxes; Subsidies and incentives
B5 – Refurbishment	Refurbishment or adaptation of asset in use; Waste management; Taxes; Subsidies and incentives
B6 – Operational energy use	Operational energy and utilities; Rent; Insurance; Taxes; Subsidies and incentives
B7 – Operational water use	Operational water and utilities; Rent; Insurance; Taxes; Subsidies and incentives

Table 2. Characterization of the sample of school buildings

Type	Construction year	Nr. of SB	Nr. of Students per SB	Land area (m^2) per SB	Constructed area (m^2) per SB
A	[1955–1964]	39	[500–800]	[13.000–27.000]	[7.000–13.000]
B	[1965–1979]	45	[400–950]	[15.000–43.000]	[5.000–12.000]
C	[1980–1988]	48	[500–1.000]	[13.000–52.000]	[5.000–11.000]

4.2 Collection of Historical Data

LCC data collection follows a well-established method described above (EN 16627). The LCC costs collected are related to modules A and B. It was not possible to discriminate costs related to module A1 to A5 due to low information granularity. A1 to A5 costs are thus presented as a group. In the present case study, nominal costs (expected price that will be paid when a cost is due to be paid, including estimated changes in price) was considered and the Discounted Factor (DF) is calculated with the Eq. 1:

$$DF(T) = \frac{1}{(1+r)^T} \tag{1}$$

Where r = annual real discount rate; and T = number of years between the reference date (start of the building life cycle) and the date of onset of the cost (2017 for the present case study). In this case study, r is variable according with the inflation/deflation in the Portuguese construction sector (Fig. 1).

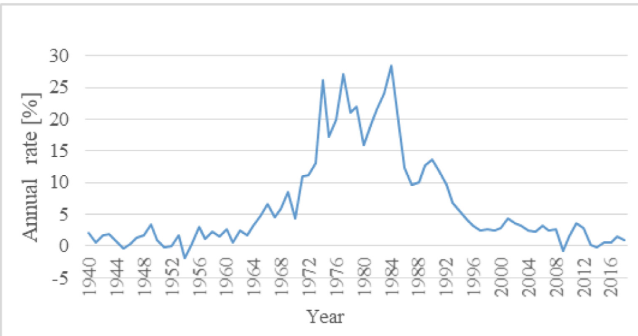


Fig. 1. Inflation/Deflation rate in the Portuguese construction sector

The collection of LCC nominal costs, presented in Fig. 2 for one SB, was performed for all the 130 SB of the present case study and a dynamic data base was established.

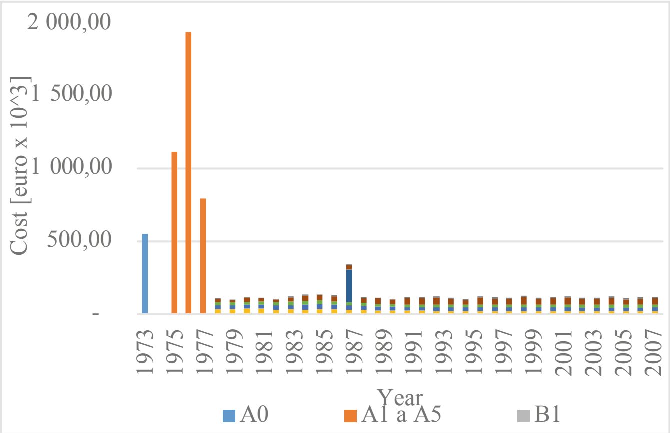


Fig. 2. Example of LCC nominal costs

5 Results

ISO 15686-5 describes several approaches for economic performance indicators. NPV (net present value, the sum of the discounted cash flows) per m^2 of the constructed asset was adopted in this case study. SB under the present case study have different life cycle periods. Therefore, in order to enable comparisons, the indicator AEV (annual equivalent value) is calculated according to the Eq. 2:

$$AEV = NPV * \frac{d(1+r)^T}{(1+r)^T - 1} \qquad (2)$$

The AEV indicator was calculated for each SB of the present case study and for different discount rates (the discount rate includes the inflation/deflation rate plus other components of the real discount rate). The k-means clustering method (Zalik 2008) was used to organize the data. The results of the average AEV for each cluster are shown in Table 3. A statistical analysis of the AEV is being undertaken and the results shall be presented in future publications.

After a statistical analysis of AEV to the entire case study the average of the relative cost ratios presented in Table 4 were calculated. According to EN 16627 and ISO 15686-5, it was considered the (i + 2)% discount rate. CAPEX (capital expenditure) includes A0-A5 and B5; OPEX (operational expenditure) includes B1-B4, B6 and B7; TOTEX (total expenditure) includes CAPEX and OPEX; Operation includes B1; and Maintenance includes B2, B3 and B4 costs. Variance of the average values of LCC ratios per Cluster (SB type) are showed in Fig. 3.

Table 3. Shows the variance of the average values of LCC ratios for each cluster

Cluster (SB Type)	Cost category	Discount rate			
		(i + 1)%	(i + 2)%	(i + 3)%	(i + 4)%
A	A0	11,55	19,39	32,23	54,76
	A1–A5	83,99	134,41	221,58	355,00
	B1	17,93	23,91	33,17	45,12
	B2	13,62	19,39	27,57	40,40
	B3	12,75	18,63	26,62	36,77
	B4	11,16	15,62	21,84	31,50
	B5	3,49	5,23	7,48	11,13
	B6	17,17	23,22	32,37	43,98
	B7	8,15	12,13	16,19	21,26
B	A0	11,91	20,20	33,93	56,45
	A1–A5	80,76	133,08	217,24	351,49
	B1	16,92	23,31	32,13	44,30
	B2	14,04	20,20	29,02	41,65
	B3	14,65	20,70	29,25	41,32
	B4	11,51	16,27	22,99	32,47
	B5	4,26	6,30	9,23	13,41
	B6	19,74	27,37	37,92	52,51
	B7	9,57	13,89	18,02	24,72
C	A0	12,39	20,40	34,61	57,01
	A1–A5	66,22	110,46	182,48	291,74
	B1	16,41	22,38	30,52	42,97
	B2	14,60	20,40	29,60	42,07
	B3	15,24	20,91	29,84	41,73
	B4	11,97	16,43	23,45	32,79
	B5	3,49	5,23	7,48	11,13
	B6	20,13	27,64	38,68	53,04
	B7	9,86	14,03	18,38	24,97

Table 4. Average value of LCC ratios

Ratio	Type A	Type B	Type C
R1 - OPEX/CAPEX	1/1,41	1/1,31	1/1,12
R2 - CAPEX/TOTEX	1/1,71	1/1,76	1/1,89
R3 - OPEX/TOTEX	1/2,41	1/2,31	1/2,12
R4 - CAPEX before use stage/CAPEX in use stage	1/0,03	1/0,04	1/0,04
R5 - Operation/Maintenance	1/2,24	1/2,45	1/2,58

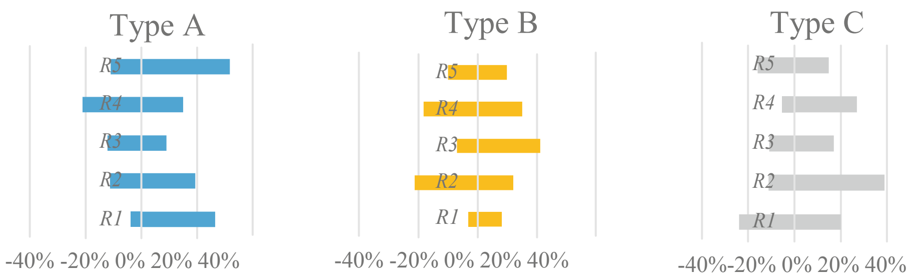


Fig. 3. Variance of the average values of LCC ratios

6 Conclusions

The AEC sector deals with several stakeholders and influences the different stages in the life cycle of building projects. The increasing demands for this sector, legal and regulatory, include the need to gather standardized information of LCC and the creation of databases. Such databases should be able to feed indicators and ratios that describe the economic performance of buildings. LCC-related information and benchmarking is a relevant component of sustainable public procurement practices. This information depends on the development of tools, communication platforms and databases with standardized information such as LCC indicators and ratios.

This paper addresses these concerns and presents a case study of 130 public school buildings constructed in Portugal since 1955, which were organized into 3 clusters. The average values of two economic life-cycle indicators are presented for each cluster: i) the average values of the AEV indicator; and ii) the average values of the cost ratios based on capital expenditure and operational costs. It was found relevant to considering the technical and functional characteristics of building portfolio clusters (building types A to C), as well as the operation and maintenance requirements in the long-term. A sensitivity analysis shows that a difference of 1% in the real discount rate represents an increase/decrease of 40% to 75% of the AEV indicator. This percentage becomes higher for longer periods of analysis.LCC relative ratios (R1 to R5) were also calculated for each type of SB. The technical and functional characteristics and the period of analysis also influence these ratios. The variance of results for the ratio average values can go up to 40% in Type C buildings, up to 35% in Type A and up to 20% in Type B. A detailed statistical analysis of the results shall be presented in future publications.

References

Arja, M., Sauce, G., Souyri, B.: External uncertainty factors and LCC: a case study. Build. Res. Inf. **37**(3), 325–334 (2009)

Cole, J., Sterner, E.: Reconciling theory and practice of life-cycle costing. Build. Res. Inf. **28**(5–6), 368–375 (2000)

Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC. Official Journal of the European Union, Brussels, Belgium

EN 15643-4:2012. Sustainability of construction works - Assessment of buildings -Part 4: Framework for the assessment of economic performance. European Committee for Standardization (CEN), Brussels, Belgium

EN 16627:2015. Sustainability of construction works. Assessment of economic performance of buildings – Calculation method. European Committee for Standardization (CEN), Brussels, Belgium

Flanagan, R., Norman, G.: Life Cycle Costing: Theory and Practice. BSP Professional Books, Oxford (1989)

Gluch, P., Baumann, H.: The life cycle costing (LCC) approach: a conceptual discussion of its usefulness for environmental decision-making. Build. Environ. **l39**, 571–580 (2014)

Goh, B.H., Sun, Y.: The development of life-cycle costing for buildings. Build. Res. Inf. **44**(3), 319–333 (2016)

IISD, International Institute for Sustainable Development: Life Cycle Costing: A Question of Value. Winnipeg, Manitoba, Canada (2016)

ISO 15686-5:2015, Building and constructed assets - Service-life planning - Part 5: Life-cycle costing. International Organization for Standardization, Geneve, Switzerland

Ive, G.: Re-examining the costs and value ratios of owning and occupy- ing buildings. Build. Res. Inf. **34**(3), 230–245 (2006)

Jrade, A., Alkass, S.: Computer-integrates system for estimating the costs of building projects". J. Archit. Eng. **13**(4), 05–222 (2007)

Langdon, D.: Life cycle costing (LCC) as a contribution to sustainable construction: a common methodology. In: Davis Langdon Management Consulting, London, UK (2007)

Ludvig, K., Gluch, P., Lindahl, G.: Life cycle costing in construction projects - a case study of a municipal construction client organization. In: 3rd International World of Construction Project Management Conference 2010, Coventry UK, pp. 1–8 (2010)

Ma, Z., Liu, Z., Wei, Z.: Formalized representation of specifications for construction cost estimation by using ontology. Comput. -Aided Civ. Infrastr. Eng. **31**(1), 4–17 (2016)

Meckler, G.: Building performance with minimal life cycle cost. Constr. Specif. **30**(5), 36–58 (1977)

OGC, Office of Government Commerce: Whole Life Costing and Cost Management Achieving Excellence in Construction Procurement Guide. United Kingdom (2007)

Zalik, K.: An efficient k-means clustering algorithm. Pattern Recognit. Lett. **29**(9), 1385–1391 (2008)

Using Indicators to Deal with Uncertainty in the Capital Renewals Planning of an Industrial Water Supply System: Testing the Infrastructure Value Index

Joao Vieira[1], Nuno Almeida[2(✉)], Jaime Gabriel Silva[3], Cristina Castro[4], Branca Lima[4], and Manuela Trindade[5]

[1] IST, University of Lisbon, Lisbon, Portugal
joaopcvieira@outlook.com
[2] IST, CERIS, University of Lisbon, Lisbon, Portugal
nunomarquesalmeida@tecnico.ulisboa.pt
[3] Instituto Superior de Engenharia do Porto, Águas de Santo André, Porto, Portugal
g.silva@adp.pt
[4] Águas de Santo André, S.A., Santiago do Cacém, Portugal
{c.castro,b.lima}@adp.pt
[5] Infraestruturas de Portugal, CERIS, Lisbon, Portugal
manuela.trindade@tecnico.ulisboa.pt

Abstract. Uncertainties are a key issue when planning for long term capital investments in asset intensive organizations. Amongst the sources of uncertainty are the quality and accuracy of readily available data and the assumptions underlying the various estimations involved in capital renewal plans, namely those related to re-placement costs and useful lives of the asset base. This paper addresses the effect of such uncertainties in the capital investment strategies for the industrial water supply system of the Industrial and Logistics Zone of Sines (ILZS), installed in the Portuguese seacoast around 40 years ago. The discussion includes deterministic and stochastic approaches to an asset management KPI that is presently being tested in the Portuguese water sector - the Infrastructure Value Index (IVI).

1 Introduction

The increasing challenges related with aging water supply infrastructures are widely covered by the literature. For example, Mirza (2007) addresses Canadian municipal infrastructures installed between 1950 and 1970 that are now reaching the end of their design life.

The ASCE reports a backlog of insufficient expenditure, a deteriorating infrastructure and the need for higher renewal rates, namely in the water sector (ASCE 2009). Selvakumar et al. (2015) mention cases of massive water losses by leaking pipes. In Brazil, the average water losses rate exceeds 36% and worldwide this rate ranges from 30% to 40%. In Europe, the volume of uncounted water is between 9% and 30% (Vilanova et al. 2015).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 226–237, 2020.
https://doi.org/10.1007/978-3-030-48021-9_26

The deferral of critical investment and the subsequent decay of the infrastructures' condition (Francisque et al. 2016) compromises the levels of service. Capital renewal investment must increase considerably in the forthcoming decades (Rokstad and Ugarelli 2015). It is therefore crucial for infrastructure asset managers to adopt a systematic approach to optimize and balance infrastructure performance (Arthur et al. 2014), risk and cost, both in the short and long-term (Rokstad and Ugarelli 2015).

Various tools have been developed and implemented in the water sector with this aim, most of which are applicable to water distribution networks, such as life-cycle cost analysis and risk analysis (Francisque et al. 2016), multi-criteria decision analysis (Lienert et al. 2014; Salehi et al. 2018) and decision trees (Winkler et al. 2018). Sewer drainage and water distribution networks represent, inside the water sector, the major focus of the existing failure predicting studies and renewal planning methodologies (Arsénio et al. 2015; El-Abbasy et al. 2016; Shahata and Zayed 2012; Nafi and Tlili 2015).

To handle complex asset management optimisation challenges, there is a trend to use evolutionary optimisation techniques such as genetic algorithms. These algorithms have been used in life cycle analysis and renewal planning for different types of infrastructures, including water and sewer networks (Rashedi and Hegazy 2015).

A rising number of software packages have been developed to support renewal strategies (Rokstad and Ugarelli 2015). These packages usually consider standard cost models, reliability performance metrics and risks of failure. WiLCO, CARE-W, I-WARP and AWARE-P are some of initiatives designed to support the asset management activities in the water sector. However, the use of these tools is not as widespread as one could expect. These tools have been sometimes criticized for not considering that water systems can be renewed in blocks or groups due to economical and practical reasons (Rokstad and Ugarelli 2015).

There are also specific contributions to the application of risk-based approaches to water infrastructures. Risk modelling techniques, as mentioned in the literature, vary in scope and complexity (Shahata and Zayed 2015; Korving et al. 2009). Kleiner et al. (2004) characterised consequences of failure combined with the possibility of failure. Most of these studies have focused on the probability of failure with little focus on the consequences of failure (Shahata and Zayed 2015). Fares and Zayed (2009) developed a risk model that assesses the risk of failure associated to each pipeline in the network. This model considers environmental, physical, operational, and post-failure factors. Rogers (2006) developed a model to evaluate water main failure risk based on a weighted average method to calculate the probability of failure. The model developed by Rogers (2006) considers the consequence of failure through "what-if" investment scenarios.

According to Lienert et al. (2014), there is rising interest for the use of indicators to compare different infrastructure intervention options, usually combined with life cycle analysis. This paper discusses this indicator-based approach combined with life cycle thinking. It presents the developments of previous study (Silva et al. 2017) about an indicator that is being tested by the Portuguese regulator of the water sector, namely the *Infrastructure Value Index* (IVI). The scope of the pilot-study is the industrial water supply system of the ILZS.

2 Infrastructure Value Index

According to Alegre et al. (2014), the IVI provides, for a given moment in time, the ratio between the current value (the fair value) of the infrastructure and its replacement cost (Eq. 1). IVI varies between 0 and 1 and is intended to translate the aging and the intervention needs of a given infrastructure and its assets. IVI values from 0,40 to 0,60 indicate the presence of a stabilized infrastructure (Amaral et al. 2016; Alegre et al. 2014). Higher values indicate lower needs of investment in the short term and vice-versa. The IVI was originally designed as a support tool for long term investment planning.

$$IVI(t) = \frac{\sum_{i=1}^{N}\left(rv_{i,t} * \frac{rul_{i,t}}{eul_i}\right)}{\sum_{i=1}^{N} rv_{i,t}} \quad (1)$$

Where: t: reference time; $IVI(t)$: Infrastructure Value Index at time t; N: Total number of assets; $rv_{i,t}$: replacement value of asset i at time t; $rul_{i,t}$: residual useful life of asset i at time t; eul_i: expected useful life of asset i.

Despite the apparent simplicity of the calculation method, the interpretation of results must be accompanied with judgment and complementary information. Alegre et al. (2014) and Amaral et al. (2016) draw attention to the fact that IVI parameters can be determined in different ways. It is possible to use an asset-oriented or a service-oriented approach to the IVI, depending if the useful lives and replacement costs of each asset are the only parameters to be considered, or if also the performance of functional units, as well as cost and risk, are to be considered. Moreover, there are also complexities related to determining useful lives of interrelated assets and components, given their different technological specificities, physical contexts or functional environments.

Despite the use of the term ‘value’ in its designation, the IVI is directly related to capital expenditure (CAPEX), leaving aside operating and maintenance costs (OPEX) and other aspects, both financial and non-financial, that are usually considered when using the concept of ‘value’ in a broader sense. Future developments of the IVI should also include specific operational contexts and uncertainties, which are out of the scope of the present study.

3 Method

The authors propose a revised method to calculate the IVI in a way that can assist long-term investment planning. When determining the actual value of the IVI (2017 in the pilot-case study presented in this paper), the steps of this revised method are:

Step 1: Collect asset-related data, including installation dates and costs, location, function and physical parameters, condition, risks, etc.;
Step 2: Establish a manageable breakdown structure of the asset base according to readily available data;

Step 3: Determine the theoretical useful lives of the asset base (ideally taking into consideration their actual physical condition and degradation trends);
Step 4: Determine the actual replacement costs at a level of the breakdown structure as low as possible or practical (e.g. asset component, asset, asset system);
Step 5: Calculate the IVI with Eq. 1 using a stochastic approach to the inputs (expected useful lives) and the outputs (IVI and current value) of the calculation process;
Step 6: Perform a long-term analysis of the IVI. The calculations of the IVI should differentiate the effect of renewals, replacements and expansion expenditures. The authors propose that the parameters used in the calculations be adjusted according to Eqs. 2 to 5.

$$rv_m^t = \beta * rv_m^{t-1} + \alpha_{rep} * CAPEX_{rep}^t + \alpha_{rep} * CAPEX_{rep}^t \tag{2}$$

$$cv_m^t = cv_m^{t-1} + \alpha_{ren} * CAPEX_{ren}^t + \alpha_{rep} * CAPEX_{rep}^t + \alpha_{exp} * CAPEX_{exp}^t \tag{3}$$

$$rul_m^t = eul_m * \left[\frac{cv_m^{t-1} + \alpha_{ren} * CAPEX_{ren}^t + \alpha_{rep} * CAPEX_{rep}^t + \alpha_{exp} * CAPEX_{exp}^t}{\beta * rv_m^{t-1} + \alpha_{rep} * CAPEX_{rep}^t + \alpha_{exp} * CAPEX_{exp}^t} \right] - 1 \tag{4}$$

$$IVI_m^t = \frac{cv_m^{t-1} + \alpha_{ren} * CAPEX_{ren}^t + \alpha_{rep} * CAPEX_{rep}^t + \alpha_{exp} * CAPEX_{exp}^t}{\beta * rv_m^{t-1} + \alpha_{rep} * CAPEX_{rep}^t + \alpha_{exp} * CAPEX_{exp}^t} - \frac{1}{eul_m} \tag{5}$$

Where: rv_m^t: replacement value of asset family *m* at year *t;* rv_m^{t-1}: replacement value of asset family *m* at year *t*−*1*; cv_m^t: current value of asset family *m* at year *t;* cv_m^{t-1}: current value of asset family *m* at year *t*−*1*; eul_m: expected useful life of asset family *m;* IVI_m^t: Infrastructure Value Index of asset family *m* at year *t; CAPEX*: *capital expenditure*; α*:* reflects the impact of a given intervention on the useful life*;* β*:* translates the disposal of existing assets. The different coefficients α and β and different *CAPEX* values for renewals (*ren*), replacements (*rep*) and expansions (*exp*) of the asset base relate as follows:

$$\begin{cases} \beta = 1, CAPEX_{ren}^t = 0 \\ \beta = 1, CAPEX_{ren}^t \neq 0 \end{cases}; \begin{cases} \alpha_{ren} = 0, CAPEX_{ren}^t = 0 \\ \alpha_{ren} > 0, CAPEX_{ren}^t \neq 0 \end{cases}$$

$$\begin{cases} \beta = 1, CAPEX_{rep}^t = 0 \\ 0 < \beta < 1, CAPEX_{rep}^t \neq 0 \end{cases}; \begin{cases} \alpha_{rep} = 0, CAPEX_{rep}^t = 0 \\ \alpha_{rep} = 1, CAPEX_{rep}^t \neq 0 \end{cases}$$

$$\begin{cases} \beta = 1, CAPEX_{exp}^t = 0 \\ \beta = 1, CAPEX_{exp}^t \neq 0 \end{cases}; \begin{cases} \alpha_{exp} = 0, CAPEX_{exp}^t = 0 \\ \alpha_{exp} = 1, CAPEX_{exp}^t \neq 0 \end{cases}$$

4 Pilot Case Study

The pilot case study for testing the proposed IVI calculation method is the industrial water supply system of *Águas de Santo André* (AdSA). AdSA is a Portuguese water utility company located near the Port of Sines and the ILZS. The distribution network and the wastewater infrastructures are excluded from the pilot case study. The industrial water services provide 53% of the total revenues generated by AdSA (around $10*10^6$ €) and comprises asset management challenges related to service levels, supply penalties, lack of redundancy, demand uncertainty, among others.

The industrial water supply system is constituted by the following infrastructures (by water flux order): one intake and pumping station, located in Sado River, with 4 pumps of 1,6 m^3/s and 2 pumps of 0,5 m^3/s of capacity; one aerial pipeline, with 3 km of approximate extension; one water channel, with trapezoidal shape and 22 km of extension; one tunnel, with an extension of 13 km; one embankment dam with a useful storage capacity of $25*10^6$ m^3 (it functions as a natural reservoir to the water supply process and is out of AdSA's current stewardship) and a pumping station with 1,5 m^3/s of capacity; one pipeline with 1 km of extension which transports the pumped water to the water treatment plant (WTP); one WTP, with a maximum treatment capacity of 0,8 m^3/s; one gravity flow pipeline, which transports the treated water to the distribution reservoir throughout a 10 km course; one distribution reservoir, located in Monte Chãos, with a 50.000 m^3 capacity of storage.

AdSA has, since 2001, the stewardship of the industrial water supply system originally constructed in the 1980's. In 2002, an acquisition value was assigned to each asset of that system. In this paper, for simplification purposes, the discounted acquisition values of the assets are used as replacement costs (*rv*), although some authors recommend taking the unit costs (according to the most recent technology of equivalent assets) and the corresponding quantities to calculate asset replacement cost more rigorously. This second option requires a deeper knowledge about the assets and involves greater effort to calculate (and regularly update) the replacement costs based on current prices and available technologies. The simplification adopted by the authors may involve lower precision in determining the replacement values but it is more practical and efficient when updating these estimates on a yearly basis.

On the other hand, a non-deterministic approach to useful lives was undertaken to determine the current value (*cv*). Three values for useful lives were considered: a minimum, a maximum and a mode, which tends to be viewed as the mean value when collecting estimates from experts (Vose 2000). The Monte Carlo Simulation (MCS) was used with the support of risk analysis software - *@Risk*.

Table 1 presents an extract of the IVI calculation inputs. Values for useful lives are based in existing literature.

Table 1. Extract of data inputs for IVI calculation (t = 2017)

Infr.	Asset family	Asset type		Year	Replacement value	Useful life		
						Min	Mode	Max
Total (Industrial Water Supply System)				–	**74 721 980 €**	–	–	–
Ermidas intake				–	4 204 223 €	–	–	–
	Civil works	Available to operation	Buildings	1980	761 815 €	40	60	100
			Reservoirs	1980	704 161 €	40	60	100
			Transformation posts	1980	376 526 €	40	60	100
		Pipes inside facilities	Pipeline	2009	1 519 €	40	60	80
		Pipes outside facilities	Pipeline	1980	1 105 775 €	40	50	60
	Equipment	Hydraulic circuits	Floodgates	1980	26 267 €	40	60	75
			Shut-off valves	1980	9 324 €	15	25	40
				2006	46 130 €	15	25	40
		…	…	…	…	…	…	…
	Electrical installations	…	…	…	…	…	…	…
(Remaining infrastructures)				…	…	…	…	…

5 Results and Analysis

5.1 IVI (2017): Stochastic Approach

The outputs of the MCS for the industrial water supply system IVI are presented in Table 2. Because the channel (12,8 M€, corresponding to 17% of the total value) and the tunnel (32,9 M€, 44% of the total value) are one-asset infrastructures and have a very significant replacement cost, the results are presented with and without considering these items in the calculation.

The channel and tunnel have a high impact on the global result of the IVI. These infrastructures have a mean IVI of 0,53 and 0,64, respectively, contrasting with a mean value of 0,37 to the remaining infrastructures. The results of the IVI suggest a good condition for the channel, the tunnel and the Monte Chãos Reservoir, all with an IVI above 0,40. But there are several aged infrastructures with needs of rehabilitation, namely those with a mean IVI value below 0,40.

5.2 IVI (2018–2038): Deterministic Approach

5.2.1 Ermidas Intake

As an example of a detailed analysis of a part of the industrial water supply system, Fig. 1 presents the capital investments scheduled until 2038 for the Ermidas intake. Figure 2 presents the results of the IVI considering deterministic useful lives (mean values).

Table 2. Stochastic outputs of the IVI calculation (t = 2017)

Infrastructure	IVI (2017)				
	Min	5%	Mean	95%	Max
Ermidas intake	0,12	0,20	0,25	0,30	0,35
Ermidas – Channel pipe	0,09	0,22	0,37	0,48	0,53
Channel	0,39	0,45	0,53	0,60	0,63
Tunnel	0,54	0,57	0,64	0,71	0,75
Morgavel Dam	0,07	0,12	0,15	0,18	0,20
Morgavel Dam – WTP pipe	0,08	0,16	0,26	0,34	0,38
WTP	0,27	0,34	0,39	0,42	0,45
WTP – Monte Chãos pipe	0,12	0,27	0,37	0,46	0,52
Monte Chãos reservoir	0,27	0,36	0,48	0,57	0,63
TOTAL - With channel and tunnel	**0,45**	**0,48**	**0,52**	**0,55**	**0,58**
TOTAL - Without channel and tunnel	**0,29**	**0,33**	**0,37**	**0,39**	**0,43**

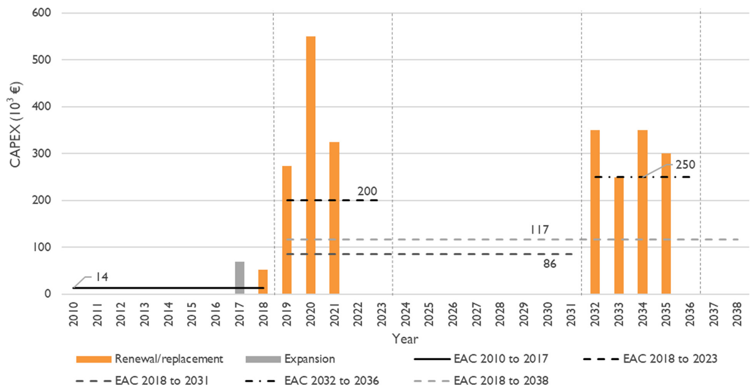


Fig. 1. CAPEX investment for Ermidas intake

The variations of IVI values depicted in Fig. 2 are related to the interventions presented in Fig. 1. The CAPEX interventions programmed until 2023 follow a risk-based approach. The planning of subsequent interventions is based on historical data and the estimated useful lives of each asset family. The 20-year period was established so that the analysis could cover, at least, the first 5 years after the current concession period (2032–2037).

The low values of IVI of Ermidas intake in 2017 are the outcome of very low CAPEX investment in the past. The programmed investments until 2031 (last year of the concession period) do not ensure a global IVI within the desired range of 0,40 to 0,60, with a significant deferring of investment in civil works until 2033. The low IVI for civil works during the concession period (below 0,30 in 2024) is a risk-based tradeoff for this infrastructure as a whole, given the present budget restrictions and higher priority of CAPEX investment in equipment and electrical installations.

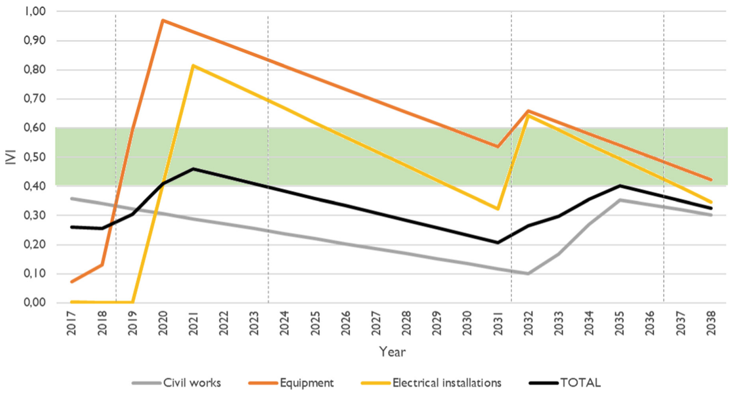


Fig. 2. IVI of Ermidas intake, by asset family

The estimated annual investment to maintain a steady value of the IVI is around 2,1–2,6% of the replacement cost of Ermidas intake. This estimate is calculated considering the levels of CAPEX investment for given periods (e.g. the periods spanning from 2020 to 2024 and 2024 to 2035, where the IVI is 0,40) and an average renewal ratio based on the weighted average useful life of Ermidas intake (48 years).

5.2.2 Industrial Water Supply System (Global)

Figures 3 and 4 illustrate the total CAPEX investments of this case study and the IVI evolution until 2038 for each infrastructure of the pilot case study, respectively.

Some infrastructures present wider variation of IVI values (like Morgavel Dam and Morgavel Dam-WTP pipe) than others throughout the 20-year period of analysis.

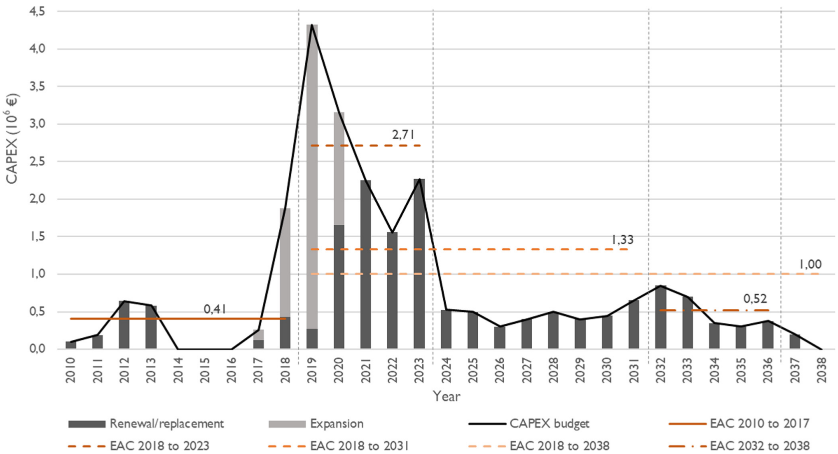


Fig. 3. CAPEX investment for the industrial water supply system

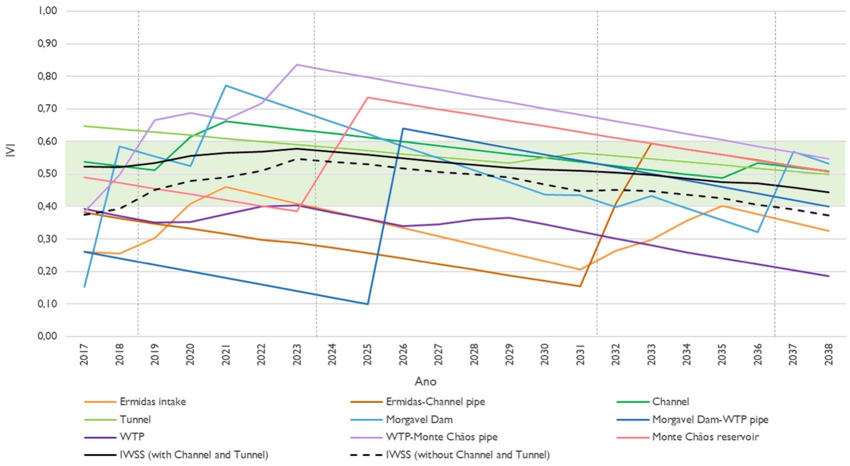


Fig. 4. IVI of HLIWSS, by infrastructure

These variations are explained by the specificities of each infrastructure and their assets, with different useful lives and replacement values, and their own renewal and upgrade strategies. The comparison of the IVI of each infrastructure must be carefully done. The analysis of the coefficients α and β must take part along with the interpretation of CAPEX investments and their impacts on the IVI.

There is a decrease of the initial gap between the values of the IVI with the channel and the tunnel and those without considering them. This is explained by the effect of the investment levels until 2029, on one hand, but also by the low decrease rates of the channel and tunnels' IVI values (due to their higher expected useful lives as compared to the other infrastructures).

5.3 Discussion

Generally, the application of the proposed methodology enhanced the interchanges and the engagement of different functions in the asset management activities of AdSA. Namely, the IVI promoted alignment of engineering and operations with finance due to the need of valuating the assets and establish useful lives according to given service levels. However, improvements to the revised method proposed by the authors are needed, namely regarding the accuracy of the inputs and the validation of the coefficients α and β used in the calculations.

The stochastic approach undertaken for the calculations of the IVI in 2017 should be extended to the estimates of the whole 20-year period. This would allow a better understanding of the long-term effects of uncertainty inherent to possible deviations of asset condition from the average scenario. This would also allow decision markers to deal with uncertainty considering their risk attitude towards the variations of asset condition and its implications on performance and the reliability of the system.

Furthermore, the declining of IVI could be calibrated against existing deterioration models for different types of assets (Ellingwood and Lee 2015). IVI calculations were found to provide a valuable contribute to maintain a record of investments on assets and to establish inspection priorities when it is not feasible to track the deterioration history of every asset over the years.

The IVI is very sensitive to the concept of useful life adopted, which can vary, at an operational level, from expert to expert and depending on the asset granularity being considered, and at a strategic level, from organization to organization. The values for the expected useful life of each asset family (eul_m) should be periodically revised according to the reality of each organization, its current physical condition and the record of service failures (due to infrastructural issues).

The concept of replacement value in the original equation of IVI lead to a discussion about the rationale underlying the concept of asset replacement. Certain asset families, when they come to a lower level of condition, do not exactly require a full replacement (e.g. a certain share of existing assets within a given asset family may be renewed while others remain unchanged), but the remaining useful lives of the asset family are extended to a certain degree with an accompanying increase of the IVI. However, when the asset family no longer fulfil the required levels of service, the full replacement of these assets may become a valid alternative and this implies a value of IVI = 1,00 for that asset family. For this reason, this paper adopts an alternative calculation method, which covers these and other conceptual particularities.

6 Conclusions

This paper used the industrial water supply system of a Portuguese utility to study both deterministic and stochastic approaches to an asset management KPI that is presently being tested in the Portuguese water sector - the Infrastructure Value Index (IVI).

The IVI was found to enable an indicator-based methodology to plan capital needs in the long term and to inform finance, engineering and operations of the consequences of the decisions made in this regard.

A revised calculation method was proposed to reduce the level of effort and resources needed to calculate the IVI. The proposed method seems to be easier than the original proposition by the regulator, but the interpretation of results must be accompanied with judgment and complementary information about the asset systems.

The right level of information granularity is key to enable the determination of the useful lives and the replacement value of assets, and thus to enable straightforward calculations, not only in terms of its effort and resources involved, but also in the quality of the resulting outputs. In any case, the input data and the assumptions made to calculate the IVI of a given asset system/infrastructure should always be clearly reported and analysed together with the results.

References

Arthur, D., Hodkiewicz, M., Schoenmaker, R., Muruvan, S.: Asset planning performance measurement framework. In: CEED Seminar Proceedings 2014, Crawley, Australia (2014)

Alegre, H., Vitorino, D., Coelho, S.: Infrastructure value index: a powerful modelling tool for combined long-term planning of linear and vertical assets. Procedia Eng. **89**, 1428–1436 (2014)

Amaral, R., Alegre, H., Matos, J.S.: A service-oriented approach to assessing the infrastructure value index. Water Sci. Technol. **74**, 2 (2016)

Arsénio, A., Dheenathayalan, P., Hanssen, R., Vreeburg, J., Rietveld, L.: Pipe failure predictions in drinking water systems using satellite observations. Struct. Infrastruct. Eng. **11**(8), 1102–1111 (2015). https://doi.org/10.1080/15732479.2014.938660

ASCE: Report card for America's infrastructure. American Society of Civil Engineers (2009)

El-Abbasy, M., Chanati, H., Mosleh, F., Senouci, A., Zayed, T., Al-Derham, H.: Integrated performance assessment model for water distribution networks. Struct. Infrastruct. Eng. **12**(11), 1505–1524 (2016). https://doi.org/10.1080/15732479.2016.1144620

Ellingwood, B., Lee, J.: Life cycle performance goals for civil infrastructure: intergenerational risk-informed decisions. Struct. Infrastruct. Eng. **12**(7), 822–829 (2015). https://doi.org/10.1080/15732479.2015.1064966

Fares, H., Zayed, T.: Risk assessment for water mains using fuzzy approach. In: Proceedings of 2009 Construction Research Congress Building a Sustainable Future, ASCE, Reston, VA, pp. 1125–1134 (2009)

Francisque, A., Tesfamariam, S., Kabir, G., Haider, H., Reeder, A., Sadiq, R.: Water mains renewal planning framework for small to medium sized water utilities: a life cycle cost analysis approach. Urban Water J. (2016). https://doi.org/10.1080/1573062X.2016.1223321

Kleiner, Y., Sadiq, R., Rajani, B.: Modeling failure risk inburied pipes using fuzzy markov deterioration process. In: Proceedings of Pipeline Engineering and Construction: What's on the Horizon? ASCE, Reston, VA, pp. 1–12 (2004)

Korving, H., Van Noortwijk, J., Van Gelder, P., Clemens, F.: Risk-based design of sewer system rehabilitation. Struct. Infrastruct. Eng. **5**(3), 215–227 (2009). https://doi.org/10.1080/15732470601114299

Lienert, J., Scholten, L., Egger, C., Maurer, M.: Structured decision-making for sustainable water infrastructure planning and four future scenarios. EURO J. Decis. Process (2014). https://doi.org/10.1007/s40070-014-0030-0

Mirza, S.: Danger Ahead: The Coming Collapse of Canada's Municipal-Infrastructure. Federation of Canadian Municipalities, Ottawa (2007)

Nafi, A., Tlili, Y.: Functional and residual capital values as criteria for water pipe renewal. Struct. Infrastruct. Eng. **11**(2), 194–209 (2015). https://doi.org/10.1080/15732479.2013.862728

Rashedi, R., Hegazy, T.: Capital renewal optimisation for large-scale infrastructure networks: genetic algorithms versus advanced mathematical tools. Struct. Infrastruct. Eng.: Maint. Manag. Life-Cycle Des. Perform. **11**(3), 253–262 (2015). https://doi.org/10.1080/15732479.2013.866968

Rogers, P.: Failure assessment model to prioritize pipe replacement in water utility asset management. Ph.D. dissertation, Colorado State Univ., CO (2006)

Rokstad, M., Ugarelli, R.: Minimising the total cost of renewal and risk of water infrastructure assets by grouping renewal interventions. Reliabil. Eng. Syst. Saf. **142**, 148–160 (2015)

Salehi, S., Ghazizadeh, M., Tabesh, M.: A comprehensive criteria-based multi-attribute decision-making model for rehabilitation of water distribution systems. Struct. Infrastruct. Eng. **14**(6), 743–765 (2018). https://doi.org/10.1080/15732479.2017.1359633

Selvakumar, A., Matthews, J., Condit, W., Sterling, R.: Innovative research program on the renewal of aging water infrastructure systems. J. Water Supp.: Res. Technol.-AQUA **64**(2), 117–129 (2015). https://doi.org/10.2166/aqua.2014.103

Shahata, K., Zayed, T.: Data acquisition and analysis for water main rehabilitation techniques. Struct. Infrastruct. Eng. **8**(11), 1054–1066 (2012). https://doi.org/10.1080/15732479.2010.502179

Silva, J.G., Pitta, J., Pinto, A., Amaral, R., Ghira, L., Castro, C., Lima, B., Brôco, N.: An implementation of asset management in an industrial and urban environmental utility, using risk based investment prioritization and a valuation index. Infrastructure Asset Management & Utility Bankability, Livorno (2017)

Vilanova, M., Filho, P., Balestieri, J.: Performance measurement and indicators for water supply management: review and international cases. Renew. Sustain. Energy Rev. **43**, 1–12 (2015)

Vose, D.: Risk Analysis – A Quantitative Guide, 2nd edn. Wiley, Hoboken (2000)

Winkler, D., Haltmeier, M., Kleidorfer, M., Rauch, W., Tscheikner-Gratl, F.: Pipe failure modelling for water distribution networks using boosted decision trees. Struct. Infrastruct. Eng. (2018). https://doi.org/10.1080/15732479.2018.1443145

Technico-Economic Modelling of Maintenance Cost for Hydroelectric Turbine Runners

Thibaud Lamothe[1(✉)], Michel Blain[2], Martin Gagnon[2], Jérôme Lonchampt[3], and Antoine Tahan[1]

[1] Ecole de Technologie Supérieure, Montreal, Canada
thibaud.lamothe.1@etsmtl.net, antoine.tahan@etsmtl.ca
[2] IREQ, Varennes, Canada
{blain.michel,gagnon.martin}@ireq.ca
[3] EDF R&D, Chatou, France
jerome.lonchampt@edf.fr

Abstract. Large utilities need to optimize investments to maintain their assets. For a utility like Hydro-Québec (37 GW of installed power) an important part of those investments is used to maintain their hydroelectric facilities. To minimize the maintenance costs, technico-economic model are essentials. They allow to propagate the uncertainties associated with the degradation processes for a given component. Therefore, we developed two technico-economic models for Francis hydroelectric turbine runners: one for crack propagation and one for cavitation. Since these are the main degradation mechanisms leading to failure of Francis runners, they enable us to study the effect of maintenance strategies on the maintenance cost. The models have been created using VME, an asset management software developed by EDF R&D (Électricité de France). VME uses Monte-Carlo simulations to generate stochastic failure dates and obtains probabilistic indicators of the net present value of a given management strategy. Using data based on a Hydro-Québec (Québec, Canada) facility, we illustrate the importance of the proper assessment of current and expected long-term reliability on maintenance cost.

1 Introduction

Turbines are mostly designed based on their efficiency. Most of the operation and maintenance costs are often neglected. Our objective using technico-analysis is to get more information about the costs associated to maintenance and long-term operation. The assessment of an asset management strategy needs to be done by accounting for costs across the assets life cycle. Given a current reference strategy, other strategies can be proposed and then compared using financial indicator. The cost of a given strategy depends on the possible failures of the asset and the dates of the events. This introduces uncertainty according to probabilistic reliability methods. From these, one can compute risk indicators for a given asset management strategy to support efficient and robust decisions.

The repair of a turbine runner blade can cost a significant amount of money in terms of production losses. A common repair strategy is presented in Fig. 1. This strategy can

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 238–245, 2020.
https://doi.org/10.1007/978-3-030-48021-9_27

be described as follows: first, a time interval between inspections. Then each time this interval is reached, an inspection event occurs. An inspection can mean weeks of turbine downtime. The magnitudes of such inspection intervals are in number of years for turbine runner. At every inspection, if a crack is detected the turbine runner is automatically repaired, except if the age of the turbine is larger than a given number of years. We have used fifty years in this case. At that time, the facility is refurbished with a new runner. This is the strategy used in this study.

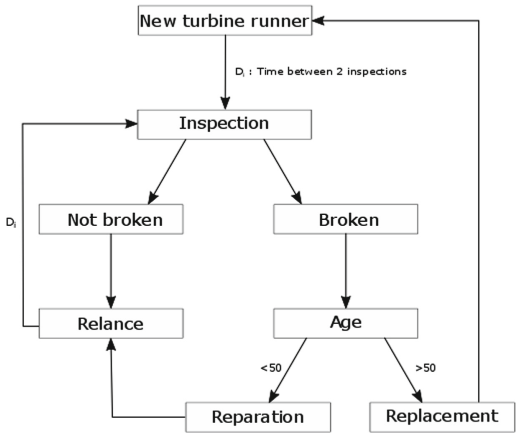


Fig. 1. Current repairing strategy

VME, an asset management tool, was used to model the strategy. With VME, the first step is to build a graphic model with blocs such as ‘events’, ‘maintenance tasks’ or ‘resources’ of the chosen management strategy. Each bloc is filled with technical data - such as degradation laws and comportment after issues – and economic information. Then, the model is used for Monte-Carlo simulations which link the events and their probability to economic consequences. At the end, technical (like average number of failures or resources available…) and economic (like costs or value of production losses…) indicators are provided. Using these, the analyst can then give information to administrators help them during decision making. One of the main indicators used to evaluate the profitability of a strategy is the Net Present Value (NPV). The NPV is a classic indicator used in finance that sums the cash-flows (Fig. 2), both positives ones (profits) and negative ones (investments) using the proper discount rate. A basic study, at Hydo-Quebec scale can easily results in millions of dollars benefits and this is why such calculations are so meaningful.

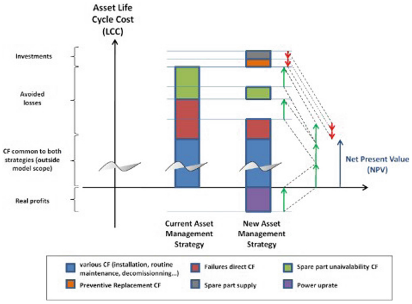


Fig. 2. NPV representation in asset management

The paper is structured as follows. First, we will have a look at the two-degradation process occurring in turbines runner: crack propagation and cavitation. Next, the study cases are presented to discuss the results obtained. Finally, we discuss the sensitivity of the assessment to uncertainties.

2 Crack Propagation

Crack propagation consists in the propagation of an initial defect to a critical value due to stress cycles. In our case, the fatigue process is linked to two kinds of stresses. The first is related to the rotation of the turbine runner and hydraulic instabilities. These are considered part of high cycle fatigue (due to their high frequencies) while the second is linked to start-and-stop cycles that are part of low cycle fatigue (Gagnon et al. 2012). Both have an impact on degradation process. The limit state used to determine when a stress cycle contributes to propagation is defined according to (Kitagawa 1976). This limit state is a function of stress amplitude and defects size a, as shown in Fig. 3. Then First Order Reliability Method (FORM) is used to obtain the probability of detecting a crack (Gagnon et al. 2013). Once a crack is detected, it is automatically repaired due to the relative long inspection intervals.

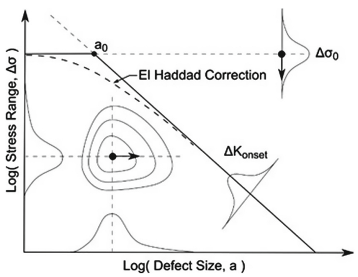


Fig. 3. Kitagawa diagram and propagation threshold

2.1 Degradation Types

The degradation is composed with three components (Fig. 4). At the beginning there is a high probability of failure because of the lack of knowledge due to inspection methods which cannot confirm the absence of defects (dead on arrival). In this case, we cannot be sure that the turbine will survive until the first inspection. After the first inspection, if the turbine is not in a failed state, only at that time, we have the knowledge that the crack probability is null. In parallel there is also a degradation component which is a function of time. Finally, there is an uncertainty component which was not included in our model.

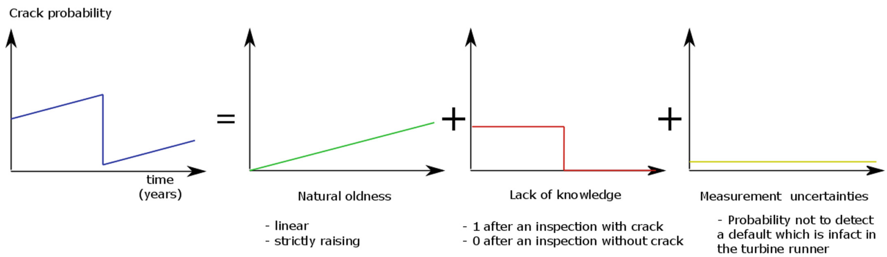


Fig. 4. Components of degradation

Because we consider that repaired turbines have worst mechanical characteristics, an increased rate due to aging is expected. Furthermore, in our model, after a repair, the lack of knowledge probability is also expected to be higher. In fact, the runner blade reliability has three states: the initial state, the repaired state and zero after inspections without failure. The combination of these values leads to the results in Fig. 5.

While in Fig. 5 left, we show results from typical values, we have used significantly higher degradation values in Fig. 5 right to highlights the correlation between the two phenomena. If natural degradation is high the percentage of new runners rises and then the proportion of death on arrival also rises. Then, when many runners are replaced, the contribution of the natural degradation decreases due to the increased percentage of new turbines.

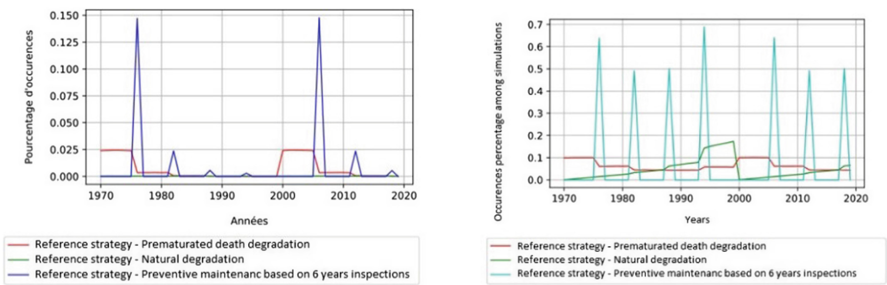


Fig. 5. Technical results from crack propagation model (200 000 iterations)

2.2 Turbine Runner Economic Model

The turbine runner modelled in this study has 13 blades. Also, we consider the runner in failed state if at least one blades has failed. To model properly the reliability, the 13 blades are modelled in series rather than the runner as a single unit. As shown in Fig. 6, since each blade is independent, and it has its own influence on the turbine this allows a better precision. This should lead to more realistic behaviour and results (Fig. 7).

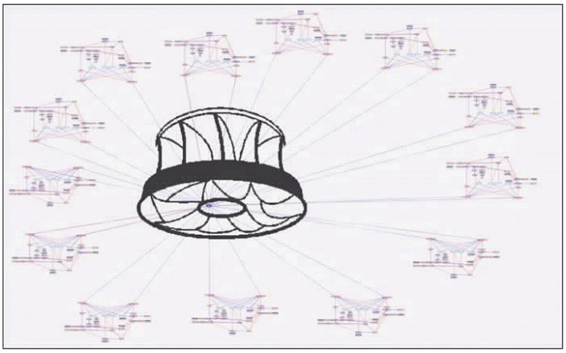


Fig. 6. View of the VME crack model

In the case shown Fig. 7, the reference strategy results are linked to corrective maintenance and penalties of non-inspection. The evaluated strategy corresponds to preventive maintenance and inspection costs. With time the observed steps in the evolution of cost are smaller and smaller due to actualisation. Here the evaluated strategy is less expensive which explains a positive NPV.

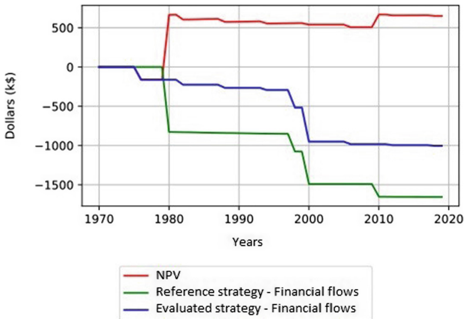


Fig. 7. Evolution of the costs between two strategies given the model

2.3 Verification

To validate these results, we used the Kaplan-Meyer method. First, we defined that a single blade was in the repaired state, while keeping the other in the initial state. By increasing the number of repaired blades in turbines, we verified that the degradation increased to ensure the proper behaviour of the model. Notice that an issue related to VME is that we cannot account for the time between a failure (crack) and an inspection. Since this time interval is correlated with the repair costs this might induce some unrealistic behaviour in our study.

3 Cavitation Phenomenon

Cavitation is a completely different type of degradation. Cavitation occurs due to water's change of phase. While usually water gets to boiling point with a temperature increase, in cavitation changes to gas with a pressure decrease for a constant temperature (Fig. 8).

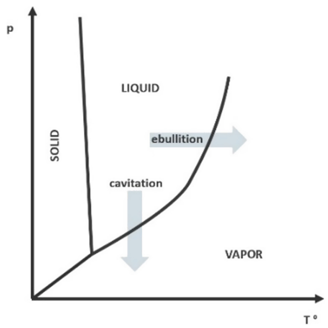


Fig. 8. Water state diagram

In turbines, due to pressure changes small vapour bubbles appear. When these bubbles explode on the blade, they erode the surface of the turbine runner blades. First mechanical stresses appear leading to a hardening followed by a tearing of material. These two stages of degradation were modelled as shown in Fig. 9.

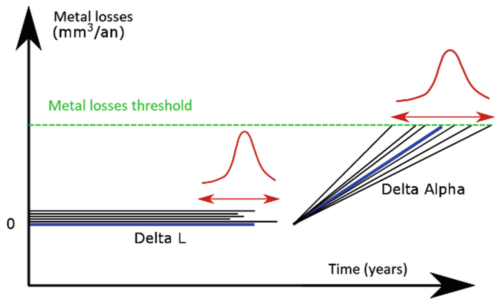


Fig. 9. Schematic representation of the cavitation phenomenon

Both stages have uncertainties (Fig. 9): the first one is the delay before erosion starts and the second one is the speed rate of the erosion. The loss of material will change from an inspection to another one as a function of those uncertain parameters. A variable threshold is usually set by cavitation experts in function of the use case to decide if a turbine runner need repair.

Figure 10 shows the result obtained for such model. The proportion of turbines needing repair changes in a periodic manner depending on the erosion rate. We noticed in this cased that the results converge after around 200 000 Monte Carlo simulations. The results taken from our first use case are very encouraging and the use of such a tool (VME) might be scalable to much bigger situations.

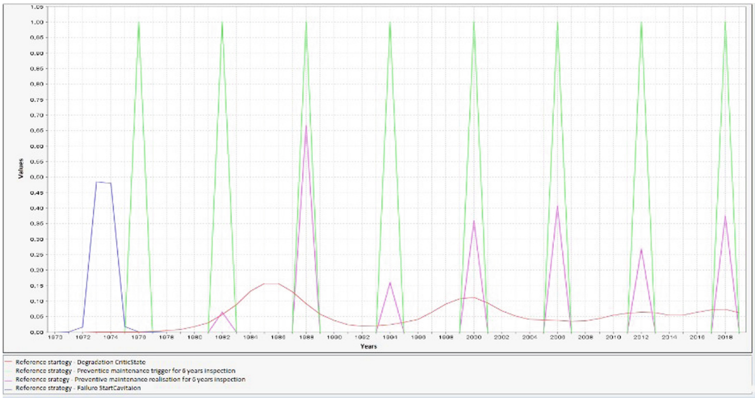


Fig. 10. Technical results from the cavitation model (200 000 simulations)

4 Conclusion

To get a better idea of turbines life cycle, we were able to build technico-economic model for fatigue and cavitation analysis of hydroelectric turbine runner. We made some simplifications and assumptions so that the model obtained enable us to get economic indicators based on technical knowledge. The VME software proved to be able to account for most of our degradation process even if we noticed that it cannot properly account for the time between failure and inspection. Our results show that by modifying maintenance strategies, the model helps to understand the impacts of maintenance's strategies and chose the most financially advantageous. By combining the two degradation models, future studies will be able to model and understand the financial value of different strategies using more realistic degradation process. Later on this will be converted into considerable profits for the company.

Acknowledgement. This report would not be possible without contribution from EDF R&D team and the free utilisation of VME they provide us for our modelization and case study. Financial support by MITACS is also gratefully acknowledged.

References

Bourdon, P.: Détection vibratoire de l'érosion de cavitation des turbines Francis. Thesis of Ph.D., Ecole polythechnique fédérale de Lausanne (2000)

Gagnon, M., Tahan, A., Bocher, P., Thibault, D.: On the fatigue reliability of hydroelectric Francis runners presented at 5th Fatigue Conference, Fatigue Design (2013)

Gagnon, M., Tahan, A., Bocher, P., Thibault, D.: A probabilistic model for the onset of High Cycle Fatigue (HCF) crack propagation: application to hydroelectric turbine runner. Int. J. Fatigue **47**, 300–307 (2012)

Kitagawa, H., Takahashi, S.: Applicability of fracture mechanics to very small crack. In: ASM Proceedings of 2nd International Conference. Mechanical Behaviour of Materials, Metalspark, Ohio, pp. 627–631 (1976)

Lafleur, F.: Vibratory detection system of cavitation erosion: historic and algorithm validation. In: Proceedings of the 8th International Symposium on Cavitation, Abstract N°150 (2012)

Lonchampt, J.: VME a tool for risk informed engineering asset management (2017)

Part III: Asset Data Management and Data Science

Introducing Lean into Maintenance Data Management: A Decision Making Approach

Salla Marttonen-Arola[1](✉), David Baglee[1], Sini-Kaisu Kinnunen[2], and Maria Holgado[3]

[1] University of Sunderland, Sunderland, UK
salla.marttonenarola@gmail.com,
david.baglee@sunderland.ac.uk
[2] Lappeenranta University of Technology, Lappeenranta, Finland
sini-kaisu.kinnunen@lut.fi
[3] University of Cambridge, Cambridge, UK
M.Holgado@sussex.ac.uk

Abstract. The amount of different data available for maintenance decision makers is extensive. However, in practice most companies are not exploiting the data in the best possible way but are wasting their resources in sub-optimal data management processes. This paper reviews the current literature on exploiting data in maintenance decision making and analyses how the principles of lean management could contribute to the issue of wasted resources. The objective of the paper is to create a literature-based framework, which will be used as the starting point for empirical research and value modelling in the later stages of the study. The presented framework highlights the role of data in maintenance management decision-making situations, and suggests how the principles of lean management could be adopted in the process of managing maintenance data to improve its value and resource efficiency. The results of this paper will contribute to future research which will include modelling and optimizing the use of data in maintenance decision making.

Keywords: Maintenance data · Decision making · Lean · Literature review · Framework

1 Introduction

The amount of data available to maintenance decision makers is growing exponentially due to technological developments such as the rapid increase in the amount of sensors, use of cloud-based computing, eMaintenance, and systems and assets connected through, for example the Internet of Things (Baglee et al. 2015; Candell et al. 2009). The data provides vast possibilities for smart, autonomous assets and predictive maintenance (Bumblauskas et al. 2017; Crespo Marquez 2007). However in practice issues related to the amount, quality, integration, and exploitation of maintenance data challenge the decision makers (see Kinnunen et al. 2016; Ylä-Kujala et al. 2016). As a result, many companies are utilising ineffectively their resources in the data management process and not succeeding in harvesting the value of the data. This paper reviews the current literature on exploiting data in maintenance decision making and analyses

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 249–257, 2020.
https://doi.org/10.1007/978-3-030-48021-9_28

how the principles of lean management could contribute to the issue of under utilised resources. The objective is to create a literature-based framework describing the role of data in maintenance decision-making, and conceptualising the adoption of lean principles in the maintenance data management process. The paper will contribute to future empirical research which will include modelling and optimizing the use of data in maintenance. The structure of the paper is as follows. Section two justifies the research design, in section three the literature on the use of data in maintenance decision making is addressed, followed by literature of lean management in the context of maintenance data. Section four will summarize the literature in the form of a framework, and in section five the paper finishes with a conclusion.

2 Research Design

This paper contributes to the groundwork of LeaD4Value (2017–2019), an international research project which aims to demonstrate how the business value of maintenance can be maximised through adopting lean data-based decision making. The role of this literature review is to examine, summarize and integrate the knowledge and existing methods currently available in the literature as a starting point for the industry-academia collaborative research in LeaD4Value. The literature and theoretical frameworks described in this paper will act as the foundations on which the empirical research and normative managerial tools, to be created in the project, will be developed. The next stage in the project will include case studies in collaboration with UK manufacturing companies. The current state of maintenance data management in the companies, as well as potential for improvements, will be analysed. Cost modelling will be used to define the value of the data to support the maintenance decision makers. In the final phase of the project a process model and a performance measurement system will be designed to help companies in improving their data management processes.

3 Lean Data-Based Decision Making in Maintenance

3.1 Exploiting Data in Maintenance Decisions

The existing literature acknowledges several important aspects regarding exploiting data for maintenance decisions. Figure 1 below emphasizes these aspects. Firstly, the twofold and iterative flow of information is required to: 1) support maintenance decision making through various data-based models and predictions, and 2) gather history data to improve the reliability and validity of the decision support constructs (Takata et al. 1999). As described by Sharma et al. (2011), numerous quantitative and qualitative models have been introduced to support the decision making in specific maintenance situations. However in this literature review a more generic perspective is adopted to increase the understanding of the role of data in maintenance decisions.

As depicted in the figure, the data is likely to be multifaceted, including technical and economical perspectives. Both quantitative and qualitative items as well as continuous and discrete parameters are usually included (Kortelainen et al. 2015). The

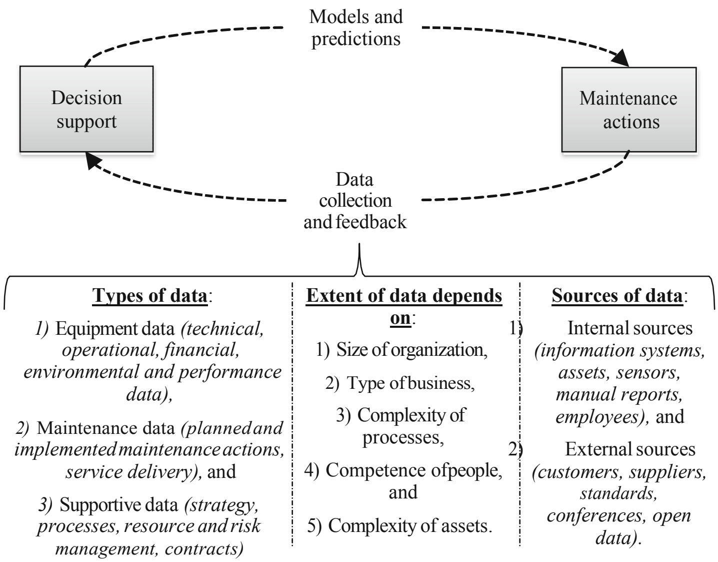


Fig. 1. Data as the foundation of maintenance actions (see Bahga and Madisetti 2012; BS EN 13306 2010; BS EN ISO 14224 2006; BS EN ISO 9001 2015; BS ISO 55000 2014; BS ISO 55002 2014; Kortelainen et al. 2015; Murthy et al. 2015; Takata et al. 1999; Wang et al. 2017)

extent to which data should be collected is highly case specific. The recent technological developments have pushed maintenance decision making to-wards real-time decisions. This means that the time available for creating solutions for maintenance optimization problems is decreasing, but so is the time required for data collection (Kinnunen et al. 2016). The importance of having a systematic, value-based plan for exploiting data in maintenance decision making is highlighted.

3.2 Lean Maintenance Management

The Toyota Company of Japan first introduced lean production in the 1950s (Gupta et al. 2016; Huang et al. 2012). Gupta et al. (2016, p. 1026) reviewed lean definitions in literature and summarized their main content into the following:

Lean is "an integrated multi-dimensional approach encompassing wide variety of management practices based on the philosophy of eliminating waste through continuous improvement".

The elimination of waste is an important aspect of lean production, and both the development actions and the performance assessment of lean should be based on this criterion (see e.g. Pakdil and Leonard 2014). The literature presents several main types of waste present in the production process. In addition to manufacturing processes, lean has also been successfully applied to several service production processes (Andersson et al. 2015; Jylhä and Junnila 2013). Lean maintenance can be seen as one of the prerequisites for lean production (Mostafa et al. 2015); without value-based, optimized maintenance services production is more likely to suffer from e.g. unplanned or excessive stoppages. Compared to 'traditional' production systems, lean production tends to require e.g. more planned maintenance (as opposed to unplanned maintenance), more training for maintenance personnel, less inventories, and less downtime and waiting time in relation to production time (Moayed and Shell 2009). The differences and similarities between Total Productive Maintenance (TPM) and lean maintenance remain somewhat vague in the literature. Some authors use the two as synonyms, and TPM can certainly be considered an integral part of lean maintenance, aiming to prevent losses, accidents, defects and breakdowns (see e.g. Andersson et al. 2015; Rolfsen and Langeland 2012). On the other hand TPM emphasizes some specific approaches like operator maintenance, whereas lean maintenance could also be achieved with other techniques.

Studies addressing the lean principles in maintenance contexts are still scarce. Mostafa et al. (2015) conclude that the same main principles can be applied in lean production and lean maintenance. However, the meaning of the key terminology (concepts like "customer", "product", "value", and "waste") is different. The main objective of lean is to maximise the value of the product to the customer through eliminating wastes (see e.g. Andersson et al. 2015; Lacerda et al. 2016).

Table 1 demonstrates how the key terms are typically understood in lean production and lean maintenance processes.

Table 1. The basic concepts of lean on the level of production and maintenance (Andersson et al. 2015; Huang et al. 2012; Mostafa et al. 2015).

Key term	Lean production	Lean maintenance
Customer	The customer to whom the product or service is delivered	The asset
Product	Product or service	Maintenance service
value	Based on the customer's needs	E.g. improved asset availability
Waste	1) Overproduction, 2) Overstock, 3) Unnecessary material transportation, 4) Unnecessary movement, 5) Waiting for next working procedure, 6) Incorrect processing, 7) Unqualified products, 8) Unused employee creativity	1) Too much maintenance, 2) Waiting for maintenance resources, 3) Centralised maintenance (excess transportation), 4) Non-standard maintenance, 5) Excessive stock, 6) Double handling, 7) Poor maintenance, 8) Under-utilisation of maintenance crew

To maximise the value of the data needed in maintenance, lean principles are introduced to the maintenance data management process. However, to identify and eliminate the wastes in the maintenance data management process, we need to introduce one more concept: lean maintenance data management, which can be seen as a prerequisite for lean maintenance (and finally for lean production). This will be addressed in the framework presented in the next section.

4 Lean Maintenance Data Management Framework

Only a few previous discussions of applying lean principles to data management exist (see e.g. Keltanen 2013; Obeysekare et al. 2016). Figure 2 presents our framework of adapting lean principles to maintenance data management. In this case, the customer is the maintenance decision maker or a stakeholder in the maintenance process, and the product is information valuable in the decision making situations. To define what constitutes value for the maintenance decision makers, the view of Bucherer and Uckelmann (2011) on value proposition in information services is adopted; the value of the information depends on the information items provided as well as on the amount,

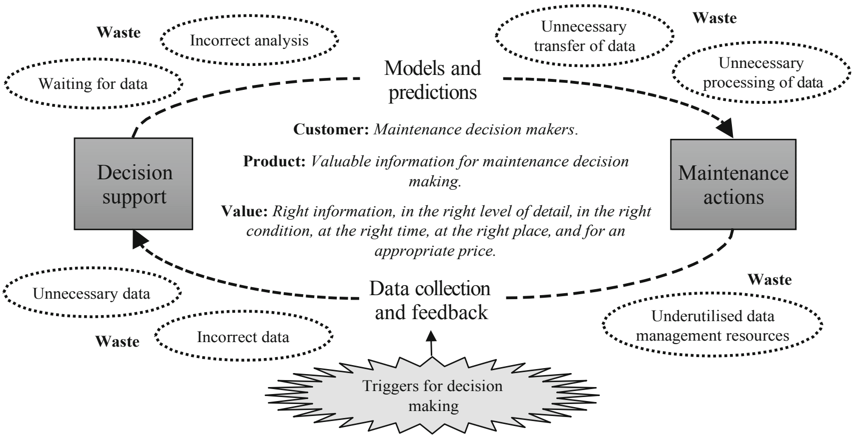


Fig. 2. A literature-based framework on lean maintenance data management.

quality, format, time, place, and price or cost of the information.

Following the logic of categorizing waste in lean maintenance and lean production, the main types of waste in maintenance data management process can be seen to include incorrect or unnecessary data, waiting for data, incorrect analysis, unnecessary data transfer and processing, as well as resource underutilization. Table 2 presents a brief description of each of these waste types. These wastes should be minimized to maximize the value, while reacting to the various triggers inducing decision-making needs.

The presented framework can be applied in mapping the data flows to analyze how much waste there is in maintenance data management processes. The nextstep is cost modelling to present the value loss in monetary terms. This would enable assessing and comparing the value of different solutions to decrease or eliminate the identified waste.

Table 2. The main types of waste in maintenance data management.

Waste type	Description
Waiting for data	The decision maker has to wait/look for data item(s) which are necessary for the decision
Unnecessary data	The decision maker has an overload of data and has to use additional time and resources to focus on what is needed for the decision making situation
Incorrect data	The decision maker has incorrect data which potentially leads to incorrect conclusions in decision making
Incorrect analysis	Errors in data analysis create incorrect conclusions in decision making
Unnecessary transfer of data	The data are transferred between people, systems or organizations without creating additional value to anyone
Unnecessary processing of data	The data are processed without creating additional value to anyone, e.g. creating summaries and reports that no-one reads
Underutilized data management resources	Unused data management resources (e.g. personnel or IT systems) create additional fixed costs

This can be described with a simple example. Let us assume that a company wishes to increase the value of data related to deciding whether to repair a broken down asset immediately or later on. The trigger for decision making in this particular example would be the asset breakdown. After mapping the data management process in the decision making situation, the company identifies two main sources of waste: 1) incorrect data related to the spare parts available in the inventory, and 2) waiting for production managers to share their latest data related to the production plans for the next days. When the decision maker thinks the required spare parts are available even though they are not (or vice versa), the risk of making a suboptimal decision increases. This may result in a lot of additional asset downtime if maintenance engineers have to wait for the spare parts. As for the production plans, the maintenance decision maker needs to be aware of them to assess the criticality of the broken down asset in the short term as it significantly affects the optimal timing of the repair. To quantify the two identified wastes, the company then needs to evaluate how much additional waiting, downtime etc. they cause e.g. in a year. Through applying a cost model the company could then find out how much money could be saved by eliminating the wastes. Presenting the waste in monetary terms would enable investment appraisal of various potential solutions. For example the company could assess the additional value of RFID tagging their critical spare parts to increase the reliability of the spare part data, or uploading their production plans into a real-time system which could be accessed by maintenance managers with their mobile phones anytime.

5 Conclusions

This paper reviews the use of data to support maintenance decision making, and complements this through introducing the concept of lean maintenance data management. The theoretical contribution of the paper extends the discussion of lean principles to cover data management applications, as opposed to previous lean management research, which has focused primarily on the optimal management of production processes, or maintenance processes. The framework also supports integrating data from both technical and economical perspectives. Existing research regarding the value of maintenance data management is scarce. Research which highlights models addressing the value of data or information is limited and almost none focus on maintenance decision making contexts. There are unsolved challenges in, for instance, defining the value of maintenance data. Managerial contributions of this research represent a potential advance on the way how maintenance decision makers handle data related processes. The final goal is to allow them to exploit data in the best possible way and reduce wasted resources, thus, improving the overall maintenance data management process. It should be noted that the maintenance decision maker can either be a maintenance manager at the company owning the asset, or a service provider involved in the asset maintenance, looking for ways to increase customer value through improving their service offering. The main limitation of the paper is that the framework will require empirical validation. This task will be examined in detail during the next stages of the research project.

Acknowledgements.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska Curie grant agreement No 751622.

References

Andersson, R., Manfredsson, P., Lantz, B.: Total productive maintenance in support processes: an enabler for operation excellence. Total Qual. Manag. **26**(10), 1042–1055 (2015)

Baglee, D., Marttonen, S., Galar, D.: The need for big data collection and analyses to support the development of an advanced maintenance strategy. In: The Proceedings of the 11th International Conference on Data Mining, Las Vegas, 27–30 July, pp. 3–9 (2015)

Bahga, A., Madisetti, V.K.: Analyzing massive machine maintenance data in a computing cloud. IEEE Trans. Parallel Distrib. Syst. **23**(10), 1831–1843 (2012)

BS EN 13306 Std.: Maintenance. Maintenance terminology, BSI Stand- ards Ltd., ISBN 978-0-580-64184-8 (2010)

BS EN ISO 9001 Std.: Quality management systems. Requirements, BSI Standards Ltd., ISBN 978-0-580-91816-2 (2015)

BS EN ISO 14224 Std.: Petroleum, petrochemical and natural gas industries – Collection and exchange of reliability and maintenance data for equipment, BSI Standards Ltd., ISBN 978-0-580-50138-8 (2006)

BS ISO 55000 Std.: Asset management. Overview, principles and terminology, BSI Standards Ltd., ISBN 978-0-580-86467-4 (2014)

BS ISO 55002 Std.: Asset management. Management systems – Guidelines for the application of ISO 55001, BSI Standards Ltd., ISBN 978-0-580-86468-1 (2014)

Bucherer, E., Uckelmann, D.: Business models for the internet of things. In: Uckelmann, D., Harrison, M., Michahelles, F. (eds.) Architecting the Internet of Things, 352 p., e-ISBN 978-3-642-19157-2. Springer (2011)

Bumblauskas, D., Gemmill, D., Igou, A., Anzengruber, J.: Smart maintenance decision support systems (SMDSS) based on corporate big data analytics. Expert Syst. Appl. **90**(C), 303–317 (2017)

Candell, O., Karim, R., Söderholm, P.: eMaintenance – Information logistics for maintenance support. Robot. Comput.–Integr. Manuf. **25**(6), 937–944 (2009)

Crespo Márquez, A.: The maintenance management framework. In: Models and Methods for Complex Systems Maintenance, Springer series in reliability engineering, ISBN 978-1-84628-820-3 (2007)

Gupta, S., Sharma, M., Sunder, V.M.: Lean services: a systematic re- view. Int. J. Prod. Perform. Manag. **65**(8), 1025–1056 (2016)

Huang, J., Bian, Y., Cai, W.: Weapon equipment lean maintenance strategy research. In: The International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE), Chengdu, China, pp. 1217–1221, ISBN 978-1-4673-0788-8 (2012)

Jylhä, T., Junnila, S.: Learning from lean management – going beyond input-output thinking. Facilities **31**(11/12), 454–467 (2013)

Keltanen, M.: Why 'lean data' beats big data (2013). https://www.theguardian.com/media-network/media-network-blog/2013/apr/16/big-data-lean-strategy-business. Accessed 10 Nov 2017

Kinnunen, S.K., Marttonen-Arola, S., Ylä-Kujala, A., Kärri, T., Ahonen, T., Valkokari, P., Baglee, D.: Decision making situations define data requirements in fleet asset management. In the proceedings of the 10th World Congress in Engineering Asset Management (WCEAM 2015). Lecture Notes in Mechanical Engineering 2195-4356, pp. 357–364, ISBN 978-3-319-27062-3. Springer (2016)

Kortelainen, H., Kunttu, S., Valkokari, P., Ahonen, T., Kinnunen, S.-K., Ali-Marttila, M., Herala, A., Marttonen-Arola, S., Kärri, T.: D2BK data to business knowledge model – Data sources and decision making needs. FIMECC S4Fleet P3 SP1 Fleet information network and decision making, 29 p. (2015)

Lacerda, A.P., Xambre, A.R., Alvelos, H.M.: Applying value stream mapping to eliminate waste: a case study of an original equipment manufacturer for the automotive industry. Int. J. Prod. Res. **54**(6), 1708–1720 (2016)

Moayed, F.A., Shell, R.L.: Comparison and evaluation of maintenance operations in lean versus non-lean production systems. J. Qual. Maint. Eng. **15**(3), 285–296 (2009)

Mostafa, S., Lee, S.-H., Dumrak, J., Chileshe, N., Soltan, H.: Lean thinking for a maintenance process. Prod. Manuf. Res. **3**(1), 236–272 (2015)

Murthy, D.N.P., Karim, M.R., Ahmadi, A.: Data management in maintenance outsourcing. Reliab. Eng. Syst. Saf. **142**, 100–110 (2015)

Obeysekare, E., Marucci, A., Mehta, K.: Developing a lean data management system for an emerging social enterprise. In: IEEE Global Humanitarian Technology Conference, Seattle, WA, USA, 13–16 October, pp. 54–61, ISBN 978-1-5090-2432-2 (2016)

Pakdil, F., Leonard, K.M.: Criteria for a lean organization: development of a lean assessment tool. Int. J. Prod. Res. **52**(15), 4587–4607 (2014)

Rolfsen, M., Langeland, C.: Successful maintenance practice through team autonomy. Empl. Relat. **34**(3), 306–321 (2012)

Sharma, A., Yadava, G.S., Deshmukh, S.G.: A literature review and future perspectives on maintenance optimization. J. Qual. Maint. Eng. **17**(1), 5–25 (2011)

Takata, S., Inoue, Y., Kohda, T., Hiraoka, H., Asama, H.: Maintenance data management system. Ann. CIRP **48**(1), 389–392 (1999)

Wang, L., Qian, Y., Li, Y., Liu, Y.: Research on CBM information system architecture based on multi-dimensional operation and maintenance data. In: IEEE International Conference on Prognostics and Health Management, 19–21 June, Allen, TX, USA, ISBN 978-1-5090-0382-2 (2017)

Ylä-Kujala, A., Marttonen, S., Kärri, T., Sinkkonen, T.: Inter organisational asset management: linking an operational and a strategic view. Int. J. Process Manag. Benchmark. **6**(3), 366–385 (2016)

From Data to Asset Health Management in the Context of Hydropower Generation: A Holistic Concept

Olivier Blancke[1(✉)], Dragan Komljenovic[2], Antoine Tahan[1], Normand Amyot[2], Claude Hudon[2], Jean-François Boudreau[2], and Mélanie Lévesque[2]

[1] École de Technologie Supérieure, Montréal, QC, Canada
olivier.blancke.1@ens.etsmtl.ca,
antoine.tahan@etsmtl.ca

[2] Institut de Recherche D'Hydro-Québec (IREQ), Varennes, QC, Canada
komljenovic.dragan@ireq.ca

Abstract. Asset management is an enabling discipline that is defined by the International Standard ISO 55000 as the coordinated activities of an organization to realize value from assets. Prognostic and health management (PHM), component and system reliability, risk assessment, asset management planning and organizational structure are all part of the asset health management process. Those activities are themselves specific disciplines which have their own standards, goals, expert communities and visions. However links between them as a whole process still have to be clearly understood and defined. The aim of this paper is to propose a holistic asset health management model depicting interrelations between activities. The well-known Data-Information-Knowledge-Wisdom (DIKW) hierarchy is applied to the asset management process to shed light on interactions between key activities and their contributions to the asset management value chain. To illustrate the concept, an example is shown in the context of hydropower generation.

1 Introduction

Over the last decade, modern companies have attempted to establish holistic system approaches for Asset Management (AsM) to increase their value creation by taking advantage of the latest technological advances. In this context, experience gained and significant progress in the field of Asset Management has led to the establishment of an international standard in AsM, the ISO 55000 standard (ISO 2014).

Asset management is defined by the ISO Standard as a set of *coordinated activities of an organization to realize the value of assets* (ISO 2014). These coordinated activities are themselves specific disciplines and have been listed through a conceptual model proposed by the Institute of Asset Management (IAM) (IAM 2015). It breaks down these activities into 6 subject groups covering a total of 39 asset management related subjects. These various activities are covered by different standards, goals, expert communities and visions. They have their own models that interact and generate

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 258–266, 2020.
https://doi.org/10.1007/978-3-030-48021-9_29

a flow from raw data to knowledge and then to appropriate actions. Their interactions and interdependencies are complex and characterized by significant intrinsic uncertainties (Komljenovic et al. 2017). This whole process has already been considered and analyzed by various authors as a Complex adaptive System (CAS) (Lacroix and Stevenin 2016; Komljenovic et al. 2017; Katina and Keating 2015).

In this context, organizations must have a good understanding of the flow of data, information and knowledge leading to the decision-making process, the role and limits of their models and the influence of uncertainties. Komljenovic et al. have developed a holistic Risk-Informed Decision Making (RIDM) approach for AsM (Komljenovic et al. 2017). The proposed model is composed of sub-models that govern the establishment of the decision process in AsM at a macro-level. The RIDM model identifies the nature and strength of interrelations between sub-models and their relative underlying uncertainties. Otherwise, Aven has proposed a conceptual framework for linking risk to the elements of the DIKW hierarchy (Aven 2013). However, it seems that, there is no conceptual model that describes the flow of information from raw data to asset health management involving different disciplines and models such as Prognostic and Health Management (PHM), component and system reliability, risk assessment and asset management planning. Links between those activities and their specific standards as a whole process still have to be clearly understood and defined.

The aim of this paper is to propose a holistic asset health management model depicting interrelations between activities and impacts on expected results by using the DIKW hierarchy. A case study is proposed in the context of hydropower generation. It illustrates the applicability of the model while identifying the main challenges and future research tracks.

2 DIKW Hierarchy

DIKW hierarchy is often used to analyze knowledge in different types of applications. The DIKW elements enable to structure and break down the learning process into data, information and knowledge towards informed decision-making in order to take appropriate actions. Different extensions of the DIKW hierarchy exist in the literature. In this paper, the approach proposed by Aven to break down the risk assessment process (Aven 2013) has been applied. Brief definitions and metaphors of the DIKW elements are proposed below. For more details regarding the DIKW concept refer to Aven (Aven 2013) and (Wallace 2007; Rowley 2007).

- Data *(Know-nothing): Symbolic representation of observable properties of the world.* (Rowley 2007)
- Information *(Know-what): Relevant, or usable, or significant, or meaningful, or processed data.* (Rowley 2007)
- Knowledge *(Know-how): Fluid mix of framed experience, values, contextual information, expert insight and grounded intuition that provides an environment and framework for evaluating and incorporating new experiences and information.* (Wallace 2007).

- Wisdom *(Know-why):* Capacity to put into action the most appropriate behaviour, taking into account what is known (knowledge) and what is most beneficial (ethical and social consideration). (Rowley 2007).

3 Conceptual Model of Asset Health Management Though DIKW Hierarchy

3.1 General Considerations

The development of the conceptual model takes into account different research works, technical reports and standards in order to get a global picture of the asset health management process. The conceptual AsM model (IAM 2015), international standard ISO 55000 (ISO 2014), NASA reports on RIDM and Probabilistic Risk Assessment (NASA 2011, NASA 2010), DIKW hierarchy of risk (Aven 2013), the AsM holistic decision-making approach (Komljenovic et al. 2017) and the book on prognostics (Goebel et al. 2017) have been used for developing the overall concept.

3.2 Conceptual Model of Asset Health Management

Conceptual model of asset health management though DIKW hierarchy is presented in Fig. 1. The hierarchy is composed of 6 elements: Data, Information, Expert's Knowledge & Wisdom and Decision-Maker's Knowledge & Wisdom. This concept intends to better illustrate the role of experts and decision makers (DM) in the overall decision-making process and to highlight their interdependence. To link the conceptual model with that proposed by the IAM (IAM 2015), colours of the elements in Fig. 1 correspond with those representing the different groups of subjects in the IAM conceptual model. The whole AsM process shows the flow of information and knowledge from raw data to decision-making and highlights the significant role of the various models involved and uncertainties they may introduce.

3.2.1 Asset Cost-of-Failure Estimation Process

Asset Cost-of-Failure estimation should be a starting point enabling to identify critical failure modes from a cost perspective. Table 1 presents its DIKW hierarchy.

3.2.2 Asset Failure Probability Estimation Process

According to the available data and criticality of failure modes, the estimation of their probability of occurrence can involve different types of models: Statistical reliability models, Bayesian models and PHM models. System reliability models ensure that all model's results will be aggregated to the asset system level where the risk will be assessed. Table 2 presents The DIKW hierarchy of asset probability of failure estimation.

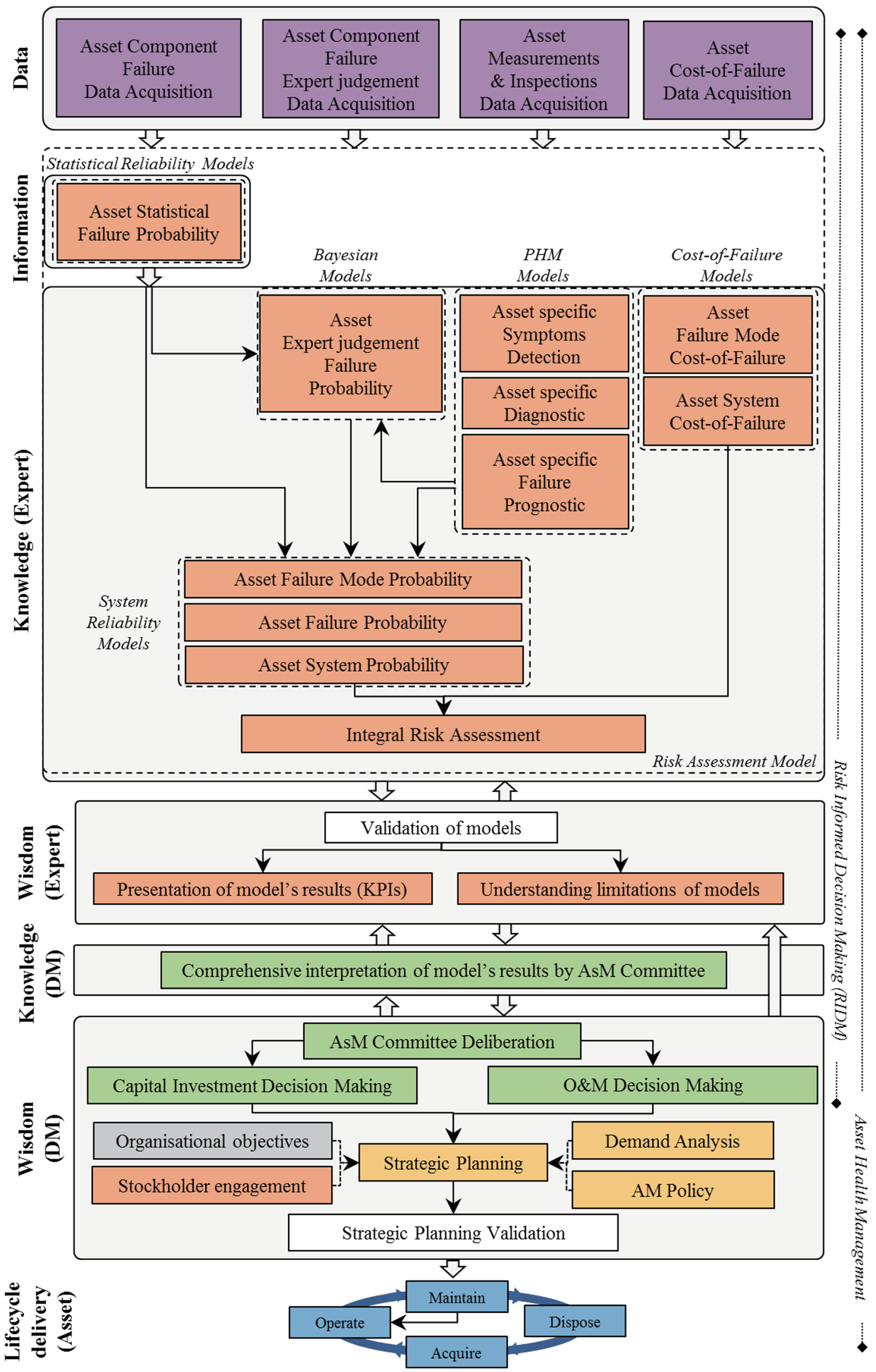


Fig. 1. Conceptual model of Asset Health Management through the DIKW Hierarchy

Table 1. DIKW hierarchy of the Asset Cost-of-Failure estimation process

	Cost-of-Failure Analysis
Data	Historical data of asset failure cost (labour, administrative, material, special services, time to repair, loss of production)
	Asset valuation and production rate
Knowledge (Expert)	- Having a good understanding of the impact of an asset system failure on the operational and organizational objectives
	- Construction of cost analysis models to estimate possible costs of planned and unexpected asset failure modes taking into account maintenance costs (labour, administrative, material, special services), loss of production costs and other potential cost. (EPRI 1999)
	- Understand cost uncertainties and their potential variability related to changes in operational, organizational, regulatory, market and others constraints. (Komljenovic et al. 2017)

Table 2. DIKW hierarchy of asset probability of failure estimation process

	Statistical approach	Bayesian approach	PHM approaches
Data	Asset usage & experiment failure data	Asset usage & experiment failure data, expert judgment, sensors data, inspection data	Asset usage & experiment failure data, expert judgment, sensors data, inspection data
Information	- Consolidate and interpret data - Estimate parameters of the population-based statistics model.	- Consolidate and interpret data	- Consolidate and interpret data
Knowledge (Expert)	- Understand what the model is able to do and what its limitations are (Aven 2013; Apostolakis 2004)	- Having a good understanding of existing failure modes, failures mechanisms and potential stochastic aspect of their propagation	- Having a good understanding of existing failure modes, failures mechanisms and potential stochastic aspect of their state estimation, future loading and propagation
		- How to perform an elicitation process combining expert judgment and hard data minimizing cognitive bias. (Aven 2013)	- Construction of fault detection, while state estimation and propagation models (prognosis algorithms). (Goebel et al. 2017)

(*continued*)

Table 2. (*continued*)

	Statistical approach	Bayesian approach	PHM approaches
		- Understand and characterize epistemic uncertainties	- Understand the resulting quantities their relative uncertainties
Knowledge (Expert)	System Reliability		
	- Having a good understanding of the asset system design and all significant component failure modes combinations that lead to the asset system failure		
	- Construction of Fault Tree (or other) models integrating the probability of failure model results of different entities to estimate a probability of failure of the asset system level where the risk is assessed		
	- Understand the logical models and conditional probabilities used to ensure that the model is inline with system design. (NASA 2011)		

3.2.3 Integral Risk Assessment & Model Validation Process

Integral Risk Assessment aims at integrating all significant risk factors including results of Sects. 3.2.1 and 3.2.2. It is a complex task that requires people from different disciplines to work together to achieve a common goal. The DIKW hierarchy of Integral Risk Assessment and validation process is presented in Table 3.

Table 3. DIKW hierarchy of the Integral Risk Assessment and validation process

Knowledge (Expert)	Integral Risk Assessment
	- Having a good understanding of the context and objectives of the organization, the overall risk management process, the asset system under study
	- How to conduct an integral risk assessment incorporating asset system failure model's result and other identified risk contributors (e.g. safety, environment, loss of reputation) (ISO 2009)
	- Understand produced risk metrics, interdependencies between various types of risks, their related uncertainties and the Strength-of-Knowledge of underlying knowledge supporting them. (Aven 2016)
Wisdom (Expert)	Validation of model
	- Being able to measure the performance of the models and algorithms (e.g. correctness, timeliness, confidence) based on other historical data. (Goebel et al. 2017)
	- Being able to assess the overall trustworthiness of the risk assessment using Zeng's framework. (Zeng and Zio 2017, Apostolakis 2004)
	Presentation of model's results and their limitations
	- Being able to present: results of risk assessment in a decision-making context, which sources of uncertainty are critical to those results, model's limitations and underlying Strength-of- Knowledge supporting it. (Aven 2013; NASA 2011; Aven 2016; Apostolakis 2004)

3.2.4 Decision-Making and Strategic Planning Process

This part concerns decision makers and highlight knowledge and the decision-making role they need to ensure. The DIKW hierarchy of Decision-Making and Strategic Planning process is presented in the Table 4.

Table 4. DIKW AM Committee Deliberation and strategic planning process

Knowledge (DM)	Decision Maker, Asset Management Committee: Comprehensive interpretation of models' results
	- Having a good understanding of the organization's context and its objectives, the overall risk assessment and analysis performed, the strength-of-knowledge supporting analysis, uncertainties, potential cognitive and motivational bias in the decision-making process
	- Understand the overall picture and key interdependencies between influence factors
	- Understand the impact of balancing the organization's long-term and short-term strategies on business while keeping the focus on organizational objectives regarding asset management policy, strategy and objectives. (NASA 2010; Komljenovic et al. 2017)
Wisdom (DM)	Decision Maker, Asset Management Committee Deliberation
	- Being able to adequately proportionate the importance, weight and limits of quantitative inputs, balance short-term and long-term strategies, deliberate pros and cons of each alternative, select alternatives and documents decision rationale. (NASA 2010; Komljenovic et al. 2017; Apostolakis 2004)
	Strategic Planning
	- Being able to grasp and consider asset management strategies, insights from deliberations, broader asset portfolio context, operational resource constraints, organisational objectives, demand analysis, asset management policy, risks and stockholder engagement and other factors leading to a detailed asset management plan. (IAM 2015)

4 Case Study – Hydropower Generation Context

This Section illustrates the applicability of the proposed Asset Management concept in the context of Hydropower generation at a major electric power utility. Companies operating hydroelectric power plants have a long-term vision of moving towards this type of conceptual model. However, the implementation of the various related processes remains a challenging long organizational and operational transformation for large organizations. The case of the Hydro-Québec hydro generators portfolio is proposed here to illustrate this transformation. Hydro-Québec is one of the most important electric utilities in North America and operates more than three hundred hydro generators.

In this context, extensive research was conducted from the 1970s to the end of the 1990 to develop sensors and measurement tools to acquire raw data from the hydro generators stators subcomponents (D and I phases of DIKW). The interpretation of these data has then necessitated an important development of knowledge through different scientific fields in order to translate the raw data into degradation assessments for each critical subcomponent of the stator (K phase). In the first half of the 2000's the raw data provided by measurement devices and sensors was translated in individual health indexes. Then, these health indexes were aggregated to give a single health index for the whole stator (W – experts). In 2008 a web application was released and implemented in all power plants. This enabled the ranking of all hydro generators with respect to their stator health index. It also allowed decision makers to take into account the state of health of their asset portfolio in their decision making process and thus to improve the prioritization of their short-term and long-term investments (K, W phases: decision makers). With the growing amount of diagnostic data available, data analytics capabilities and the expert knowledge acquired on asset failure mechanisms, the companies have realized the importance of research program development in predictive maintenance in order to take advantage of those advances. As an example, for hydroelectric generating units, prognostic models are under development for the generator and the turbine. Further research will be needed to integrate those model results into an integral risk assessment as described in the proposed conceptual models. The model helps guiding the development of various activities to achieve a common goal. As the organization develops and implements new technologies, decision-making processes evolve to take into account this new information. However, the transformation of organizations towards the proposed conceptual model remains a long process through which many challenges remain.

5 Conclusion

This paper presents a holistic concept of asset health management that highlights the flow of information and knowledge from the various disciplines involved. In the literature these different disciplines are often described independently through different standards. However, this study integrates them into a whole process and illustrates their interrelationships and interdependencies using the DIKW hierarchy. The role of experts and decision makers is also highlighted. The case study carried out on the hydroelectric context shows that such a conceptual model remains a long-term vision for the organizations and involves many operational and organizational challenges in its implementation. However, as new technologies are implemented, organizations are gradually adapting their decision-making process to take into account this new and relevant information. It shows that the whole process and its implementation can be considered as a Complex Adaptive System (CAS).

References

Apostolakis, G.E.: How useful is quantitative risk assessment? Risk Anal. **24**, 515–520 (2004)

Aven, T.: A conceptual framework for linking risk and the elements of the data– information–knowledge–wisdom (DIKW) hierarchy. Reliab. Eng. Syst. Saf. **111**, 30–36 (2013)

Aven, T.: Risk assessment and risk management: review of recent advances on their foundation. Eur. J. Oper. Res. **253**, 1–13 (2016)

EPRI: Cost Benefit Analysis for Maintenance Optimization. Palo Alto, CA (1999)

Goebel, K., Daigle, M., Saxena, A., Sankararaman, S., Roychoudhury, I. Celaya, J.: Prognostics, NASA (2017)

IAM 2015: Asset Management - an anatomy V3 (2015)

ISO 2009: IEC/ISO 31010 - Risk management – Risk assessment techniques (2009)

ISO 2014: ISO 55000, 55001 and 55002 Asset Management Standards. BSI Standards (2014)

Katina, P.F., Keating, C.B.: Critical infrastructures: a perspective from systems of systems. Int. J. Crit. Infrastruct. **11**, 316–344 (2015)

Komljenovic, D., Abdul-Nour, G., Boudreau, J.-F.: Decision-making in asset management under regulatory constraints. In: 12TH World Congress on Engineering Asset Management, Brisbane, Australia (2017)

Lacroix, T., Stevenin, P.: Strategic asset management: a system driven approach on electrical transmission systems. IEEE-Reliability Society (2016)

NASA 2010: Risk-Informed Decision Making Handbook. NASA Headquarters (2010)

NASA 2011: Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners, 2nd edn. NASA, Washington, DC, United States (2011)

Rowley, J.: The wisdom hierarchy: representations of the DIKW hierarchy. J. Inf. Sci. **33**, 163–180 (2007)

Wallace, D.P.: Knowledge management: historical and cross-disciplinary themes, Libraries unlimited (2007)

Zeng, Z., Zio, E.: A classification-based framework for trustworthiness assessment of quantitative risk analysis. Saf. Sci. **99**, 215–226 (2017)

Lifetime Prediction as Result of Calculated Component Loads from Available System Data

Lothar Wöll(✉), Georg Jacobs, and Achim Kramer

Institute for Machine Elements and Systems Engineering (MSE), RWTH Aachen University, Aachen, Germany
lothar.woell@imse.rwth-aachen.de

Abstract. Availability of offshore systems significantly affects operational costs and is therefore an important performance indicator. In order to ensure a high system availability, failures of equipment must be avoided. Condition monitoring for detection of upcoming failures usually requires an elaborate and therefore costly sensor infrastructure. This paper presents a method that aims to acquire condition information of gearbox components of ship-mounted offshore crane winches by using only already available and easily accessible external load information, such as motor torque, wave height and carried payload. For offshore cranes equipped with Active Heave Compensation (AHC) technology, the drivetrain loads depend significantly on the wave-induced motion of the ship and subsequently of the crane. The resulting motion of the load-determining crane tip is calculated using wave height predictions and the dynamic behaviour of the ship. With an appropriate system model, the local component loads can be calculated in a way that each component's remaining lifetime can be estimated by applying component lifetime models. Furthermore, critical components can be identified and monitored to avoid unplanned downtime of the equipment.

1 Introduction

For offshore operations, the availability of the employed equipment is crucial for the mission success. Typically, neither the offshore support vessel, nor the vessel-mounted offshore cranes with Active Heave compensation (AHC) for deep-water lifting are deployed redundantly, due to their high costs. Thus, unplanned downtime of a crane usually means either interruption or abortion of the entire mission. Due to this circumstance, omnipresent knowledge about the system condition is crucial to enable uninterrupted operations and a high availability. The most straight-forward method to gather information about equipment health is equipping as many components as possible with a variety of sensors. In combination with continuous analyses, the detection of deviations in the captured signals, which might announce upcoming failures, is possible. Since high costs of the necessary sensor infrastructure contradict that approach, condition monitoring systems that utilize only few additional sensors or only already available data are favourable to ensure high availability at low cost. For subsea cranes, external load information such as drivetrain torque, carried payload and wave height is already available or easily accessible. The reduction of remaining useful lifetime of mechanical components is usually the result of experienced component

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 267–274, 2020.
https://doi.org/10.1007/978-3-030-48021-9_30

loads and environmental influences. This paper describes an approach for the gathering of information about remaining useful lifetime of drivetrain components of an offshore crane winch based on available external operation data.

This paper is structured as follows: First, the investigated system of an offshore crane winch drivetrain and its operation conditions are described. Then, a method for calculating the system lifetime, based on external load data and system knowledge, is presented. Subsequently, exemplary results are presented and their benefits are pointed out. Finally, an outlook will be given, on how to employ the method's results for more economic designing and faster design process of similar systems.

2 System Description

Cranes for offshore applications are usually offered with Active Heave Compensation (AHC). This technology allows the crane operator to maintain a steady position of the payload despite movements of the vessel induced by ocean waves (see Fig. 1). In this paper, an offshore crane is investigated, that utilizes its winch to compensate wave motions.

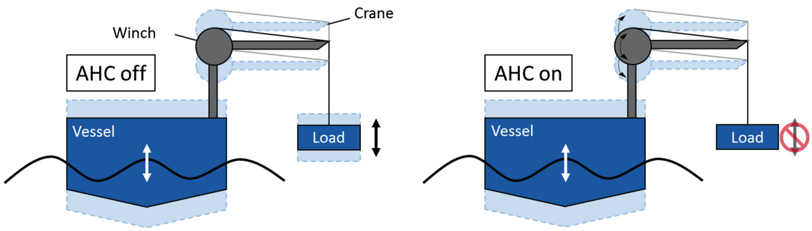


Fig. 1. Offshore crane with AHC, mounted on vessel

The design of the drivetrain of the winch subject to research in this paper is based on an existing winch of a typical 150 t knuckle-jib crane for subsea operations with a wire capacity of 3000 m. The drum is equipped with cogwheels on both sides, rotated by ten pinions each. All twenty pinions are connected to an individual hydraulic motor by a two-stage planetary gearbox (see Fig. 2). For the investigations, it is assumed, that all twenty sub-drivetrains are necessary for the proper machine operation. If one sub-drivetrain fails, the entire crane goes out of order.

During activated AHC mode, the experienced load on the drivetrain of the winch results from the static weight of the payload and dynamic loads caused by the acceleration of the drum and the coiled wire. When unrolled, an additional load has to be added to the payload due to the wire mass. Coiled, there is no additional static load, instead the moment of inertia of the wire is multiple times higher compared to the empty drum and has to be taken into account for the calculation of the dynamic loads. Furthermore, the lever arm for the load application is significantly higher for a coiled drum than for an uncoiled drum.

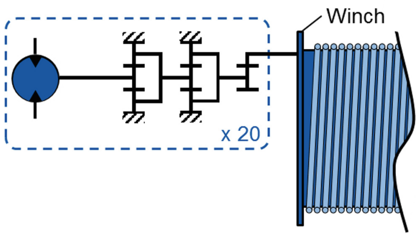


Fig. 2. Offshore winch drivetrain

3 Approach

A system's life expectancy is defined by the lifetimes of its components, which depend on their design and experienced load during operation. Acting loads on machine elements can be determined by measurement or simulation using external loads, such as torque at gearbox input. The method described in this paper gives an overview of the steps taken to calculate the lifetime of the present system. To provide a more general system analysis, not only the external loads, but the prediction of the ocean waves serve as basis for the load calculation, see Fig. 3.

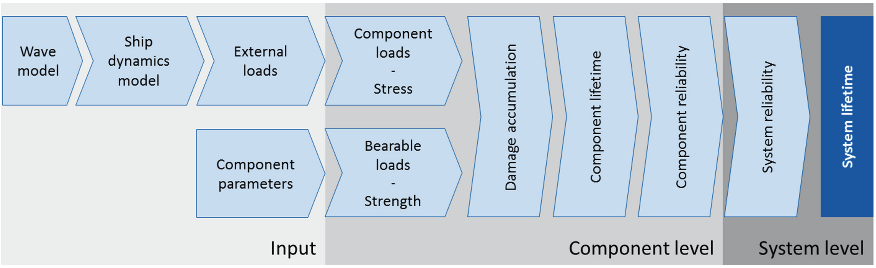


Fig. 3. Method for reliability evaluation approach of offshore winch

The external loads originate from static loads due to payload and dynamic loads. The dynamic loads are caused by dynamically accelerating inertias to compensate the crane tip motion induced by the irregular ocean waves. To determine the dynamic loads, it is necessary to describe the motions of the waves mathematically. For this, the spectral energy density formulation measured during the "Joint North Sea Observation Project" (JONSWAP) is used in this work (Hasselmann et al. 1973). Using this prediction model, the vertical wave motion can be calculated.

The vertical motion of the crane tip is decisive for calculating the loads in the drivetrain. Since ocean waves can cause rolling and pitch ship motions in addition to a heave motion, the ship dynamics have to be considered for determining the vertical crane tip speed and acceleration caused by the wave motion. Among the available methods to quantify the ship motion, in this project, the strip theory is selected to create a response amplitude operator (RAO) (Tasai 1969). The RAO can be used to transform the wave motion to a ship motion in frequency domain, from which subsequently the time-dependent vertical wave position, vertical wave velocity and vertical wave acceleration can be derived. Torque and rotational speed of the drives depend on the anticyclical acceleration and deceleration of the winch drum to compensate the vertical crane tip motion. The local component loads, necessary for the lifetime calculation, can be determined using static transfer functions or multi-body simulations of the winch and its gearboxes. The loads and speeds of the machine elements depend on the imposed torque and speed due to crane motion and payload.

Strength of machine elements, determined by its design, can be expressed by component Wöhler curves, which consist of the number of bearable load alternations or rotations at a specific load. For a system as the drivetrain subject to research, the components considered most relevant for the system reliability are gears and bearings. The number of bearable load alternations can be determined using available standards, such as ISO 281 for rolling bearings or ISO 6336 for gears. Since these standards consider only fatigue, other damage mechanisms cannot be considered using these standards. By comparing the number of experienced rotations for specific loads with the number of bearable rotations for the respective loads, the damage, caused by the load alternations, can be determined for each machine element. In combination with the sequence time, the lifetime can be extrapolated from the damage, see Fig. 4.

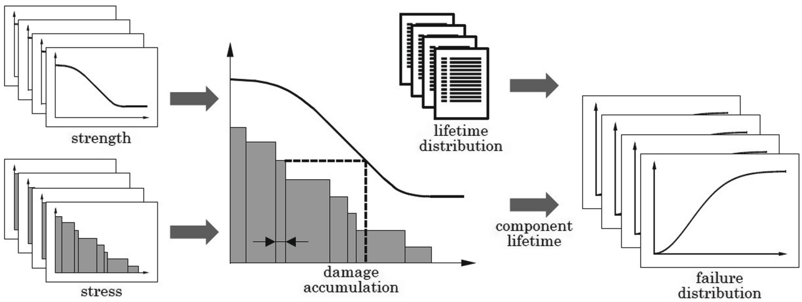


Fig. 4. Calculation of component failure probabilities

Since the lifetime of a system is not necessarily equal to the life expectancy of the weakest component, but is the result of the overall statistical failure probability of all system elements, the system failure distribution needs to be determined, from which subsequently the system lifetime can be derived. This is especially important for the investigated drivetrain, since the drivetrain consists of 20 identical gear-boxes, with assumingly similar lifetimes. For this, the probability of failure of each machine element needs to be available in the form of component failure distributions. With the results of the lifetime calculations available, the failure distributions can be approximated by Weibull distributions using parameters based on typical values of similar components from literature (Bertsche 2008; Neumann et al. 2016). Based on the assumption that the entire system fails if one machine element fails, which is the usual case for systems without redundant components, the overall system failure distribution is determined by applying the Boolean system theory for a serial reliability system (Bertsche 2008). Thus, the probability of survival of each components, as complement to the failure probability, is multiplied with each other to yield the resulting survival probability and the failure probability. The ultimate result is the system lifetime, derived from the overall failure probability (Neumann et al. 2016).

4 Case Study

The immediate result of the previously described method is the failure distribution for the entire system, from which the lifetime can be derived. Figure 5 shows exemplary failure probabilities for the investigated system. The loads for this analysis are determined for a use case of a carried payload of 150 t during a significant wave height of 1.65 m (Beaufort 5). The AHC mode maintains a steady vertical position of the payload and causes dynamic loads in the drivetrain. For comparison it is assumed that this use case lasts the entire service life until the components have failed.

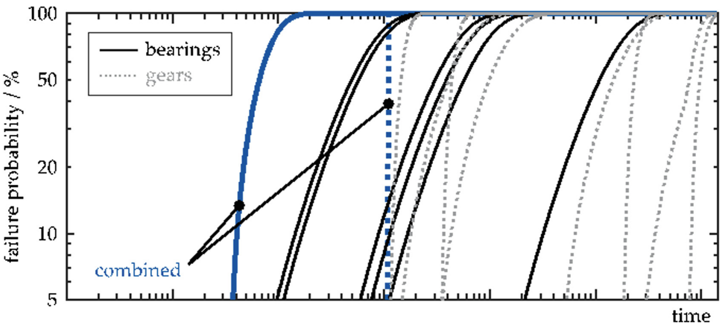


Fig. 5. System reliability results (Wöll et al. 2017b)

The solid black lines represent the failure probabilities of individual bearings of each of the twenty identical gearboxes. The dashed grey lines represent individual gears per gearbox. The solid blue line and the blue dashed line, respectively, give the likelihood of failure of all bearings and gears in combination of all twenty gear-boxes. It can be observed, that the overall failure probability of the system is much higher compared to the weakest component. Furthermore, bearings can be identified as the critical component group of the system.

Furthermore, by performing sensitivity analyses based on existing systems, crucial system parameters can be identified. Exemplary given in Fig. 6 are the results of sensitivity analyses regarding lifetime influence of different number of coiled layers on the drum of the winch for a constant static load (payload and rope weight). It can be observed, that an increasing number of coiled layers cause a significantly lower lifetime. This is due to the fact, that the static load is applied by a larger lever arm for an entirely coiled winch drum, compared to an unrolled winch drum, thus causing higher torques in the drivetrain. Other system parameters subject to sensitivity investigation have been, among others, wave height, amount of payload, and lubricant temperature (Wöll et al. 2017a and c).

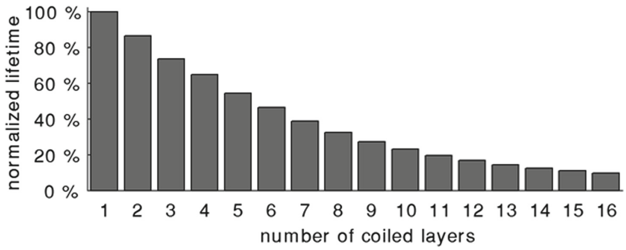


Fig. 6. Exemplary sensitivity analysis concerning (Wöll et al. 2017a)

The presented results will allow the implementation of cost-saving selective monitoring for such systems. In this way only the most critical components, identified by the reliability analysis, would be equipped with monitoring sensors, since they are likely to fail first and thus determine the system's life expectancy. Thus, unnecessary sensors for uncritical components can be avoided. Furthermore, by deriving simple real-time capable surrogate models, the fatigue condition of the system's machine elements can be calculated directly on the crane's IT-infrastructure, based on the fatigue status of one or more monitored components or external loads. This allows for immediate actions in case one or more components appear to be in need of repair or replacement, thus avoiding unplanned downtime during an operation.

The application of the presented method is not limited to offshore crane winches in use. Since the component loads can also be determined using wave height predictions, instead of using measured loads, the life expectancy can already be estimated in early design stages of similar systems. By building on the available results during future work, influences on the lifetime can be described and quantified by performing elaborate load analyses of existing systems, e.g. using multi-body simulations. These influences can include, among others, different drive technologies, such as hydraulic drives or electrical drives, or load amplifications due to internal gear excitations. These quantified influences can be transferred to comparable systems, e.g. in form of factors which describing load amplifications, (Wacker 2013). In this way, for designing similar systems, only simple calculations in combination with the quantified influences need to be performed to receive more accurate results, instead of redoing elaborate simulations every time. This enables machine designers to design economical systems with shorter development times.

5 Summary and Outlook

The condition of drivetrain components is usually assessed with condition monitoring systems, which require an elaborate and expensive sensor infrastructure. In this work, a general approach has been presented that allows the evaluation of the remaining lifetime of offshore crane winch drivetrains without additional sensor data. Modern cranes allow the easy access to external load information, such as wave height, payload and motor torque. Based on this already available data, the component loads can be calculated, which determine the component lifetimes. The results of the method, consisting of the lifetime and failure distribution of the individual components as well as the entire system, allow the identification of critical components and the estimation of the remaining useful lifetime. By creating surrogate models based on the results of the method, real-time evaluation of the machine condition will become possible. However, this method is not limited to the application of existing systems in use but also to future systems that are to be designed. Due to the possibility of a general load assessment, which allows the calculation of loads based on wave height predictions, the expected lifetime can already be evaluated in the design phase to allow an early and thus efficient economic design of such systems and its components.

Since technical systems can also fail due to causes other than fatigue and wear, condition monitoring systems cannot be avoided entirely to detect abnormal behaviour. However, the identification of critical components allows selective and therefore low-cost condition monitoring. The combination of such condition monitoring strategies with the ability to determine the remaining lifetime based on the experienced load history, enables the scheduling of maintenance activities with a higher precision and with a lower frequency. Thus, more efficient maintenance strategies would cause lower maintenance costs and higher safety for the maintenance crew.

Acknowledgements. The research presented in this paper has received funding from the Norwegian Research Council, SFI Offshore Mechatronics, project number 237896.

References

Bertsche, B.: Reliability in Automotive and Mechanical Engineering: Determination of Component and System Reliability. Springer, Berlin (2008)

Hasselmann, K., Barnett, T.P., Bouws, E., Carlson, H., Cartwright, D.E., et al.: Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP), Hamburg (1973)

Neumann, S., Wöll, L., Feldermann, A., Strassburger, F., Jacobs, G.: Modular system modeling for quantitative reliability evaluation of technical systems. MIC (Model. Identif. Control) **37** (1), 19–29 (2016). https://doi.org/10.4173/mic.2016.1.2

Tasai, F.: On the swaying yawing and rolling motions of ships in oblique waves. Sel. Papers J. Soc. Naval Archit. Japan **3**, 92–108 (1969)

Wacker, M.: Einfluss von Drehungleichförmigkeiten auf die Zahnradlebens- dauer in Fahrzeuggetrieben. Ph.D. thesis, Stuttgart University, Stuttgart (2013)

Wöll, L., Jacobs, G., Feldermann, A., Neumann, S., Strassburger, F.: Einfluss der Wellenhöhe auf die Zuverlässigkeit einer Offshore-Winde. In: Antriebstechnisches Kolloquium 2017: Tagungsband zur Konferenz, ATK 2017, pp. 59–70, Aachen (2017a)

Wöll, L., Schick, K., Jacobs, G., Kramer, A., Neumann, S.: Reliability evaluation of drivetrains: challenges for off-highway machines. In: Volosencu, C. (ed.) System Reliability. InTech (2017b)

Wöll, L., Feldermann, A., Jacobs, G.: Sensitivity analysis on the reliability of an offshore winch regarding selected gearbox parameters. Model. Identif. Control **38**(2), 51–58 (2017c). https://doi.org/10.4173/mic.2017.2.1

Method to Solve a Privacy and Security Issue in Cloud for Energy Informatics

Jayachander Surbiryala(✉) and Chunming Rong

Department of Electrical Engineering and Computer Science, University of Stavanger, Stavanger, Norway
{jayachander.surbiryala, chunming.rong}@uis.no

Abstract. Usage of the electronic devices has become part of the human life and with the advancement of technology, people are using more and more electronic gadgets which are getting connected to the Internet to provide better services to the customer. Evolution of Internet of Things (IoT) around customer usage for their daily applications is gaining a lot of attention from various service providers. Now, it is very common to find many of the services are connected with some sensor and data is transmitted to the service providers or they have made smart applications to help the customers are using products even more. But, in any case, this information is available or stored somewhere in the cloud. Energy based companies are also using the smart meters which will be communicating the user's behaviour and usage at regular intervals. Minute details regarding the energy usage can be used for the energy forecasting or load grid management for better resource availability for the customers. But this user level profiling can lead to serious security concerns to the customers. If this information is not well protected that will lead privacy issue, and customers using the services might face serious security issues in some situations. The service providers need to be care full where this data is stored and how this information is processed.

1 Introduction

Most of the development of the humans is based on technological innovations which are based Energy to achieve various aspects. If we look back the last couple of decades consumption of energy has been increasing at the rapid because of various aspects at the global level. This global demand for the energy can be because of the globalization, industrialization, increase in population, or rapid adoption of the modern technologies. Growth in the consumption energy might increase in coming years because of the evolving technologies like the Internet of Things (IoT), Cloud computing, Big Data Analysis and Social Networking sites. As most of these modern technologies need a lot of energy to support various services to run the servers, devices, and so on. Energy for everyone is generated from various sources such as fossil fuels, nuclear power, and renewable energy (such as hydro, wind, and solar). Still, there are some places which don't have enough electricity to support every household usage in their region or community. Some of the countries rely on the fossil fuels for electricity generation, but many countries are moving towards the renewable electricity generation (or sources) [1].

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 275–282, 2020.
https://doi.org/10.1007/978-3-030-48021-9_31

The various technologies are influencing a lot on the energy consumption by customers in various forms being part of their daily life and usage. For better serving, the customer's energy companies are trying to analyse the user behaviours and consumption patterns using Artificial Intelligence [2]. For this purpose, most of the modern houses are being fitted with the smart electric meters, which will transmit the electric usage of the customer at regular intervals. This information will be stored by the electric service providers somewhere on the Cloud. As the electricity usage of all the customers is being collected and stored at regular intervals of time. This leads to Big Data associated with the customer usage and their behaviour. These organizations will be using this information to do some analysis to provide better services to their customers. Analysis of this data helps them to better understand the usage of the customers and they can forecast the future requirements for each customer. For this purpose, there is a lot of research which has been carried out in this domain predict the electricity consumption using various techniques.

Proper predictions of the energy consumption, organizations can plan how much energy they are going to need, how much power they need to produce. They can make the plan for long-term without any problems. Present energy systems will be able to predict the future requirements based on the information collected by these smart meters. Based on these predictions electricity provider can easily manage the community or grid level load without any outages to the customers.

But using the customer or user behaviour and their patterns gives away crucial information. Which can be used against the customers to target them and its make them vulnerable to new security issues in their own neighborhood. To address this problem, we would like to suggest some different methods to forecast and how this help in anonymising the data and providing privacy for the users and securing the customers from the problems which might come in to picture.

The main reason for the prediction analysis is, many of the times electricity providers face problems with meeting the demands of the customers during the peak hours or during abnormal timings (other than peak hours). This depends on various factors such as climate, working or holiday, vacations, and climate [3]. Customer consumption can vary throughout the day for various other factors as well. Which will be hard for the providers to guess these aspects in real time to meet the requirements.

In this article, we have identified a privacy and security issue which arise by storing the energy consumption in a cloud environment and proposed a simple method to solve the problem. Rest of the paper is organized as follows: Related works in energy informatics are presented in Sect. 2. In Sect. 3, we have presented the database used in this work and generic energy consumption analysis is presented in Sect. 4. In Sect. 5, a new method to store the energy consumption in the cloud is presented and finally, we draw the conclusion in Sect. 6 with some future work.

2 Related Works

There has been a lot of research in the energy forecasting to predict the usage of customers by various researchers and electricity suppliers. Some of them have studied the clustering techniques for load profiling [4]. In [5, 6], they have analysed the applications for the load balancing and demand response of the customers. Few articles are focused on energy forecasting using long short term memory (LSTM) [7], conditional restricted boltzmann machines (CRBM) and factored conditional restricted boltzmann machines (FCRBM) [8] based on the deep learning models.

There are many research articles which discuss many other problems associated with energy informatics such as, loss and challenges in power grids associated with the smart meters [9, 10]. Load forecasting is one of the important aspects which can be achieved with the smart meters. These smart devices help to better predict the energy forecasting. Some of the forecasting techniques are summarized in [11, 12].

In [13], they have discussed the various challenges associated with security and privacy aspects of energy informatics. Our paper differs from [13] as we are focusing on privacy and security which can arise based on the energy consumption pattern of the customer which are stored in the cloud for energy forecasting. Whereas in the paper mentioned above, they have discussed the challenges and opportunities of energy forecasting based on Big Data analysis. This paper tries to identify a crucial problem which can arise with energy forecasting and propose a simple solution to solve those problems.

2.1 Problem

With the rapid changes in the electricity usage by customers, electricity suppliers or producers started analysing the energy usage of customers based on their previous usage patterns. For that purpose, they are keeping tracking of the energy usage of customers in cloud servers to predict the future requirements. As we know, there are many security issues in the cloud which need to be addressed otherwise storage and processing of the energy consumption data will lead to privacy and security issue for the customers. For example, as shown in the Fig. 1, it is evident that energy consumption in this particular household is at a minimal level in morning time from *8:30*

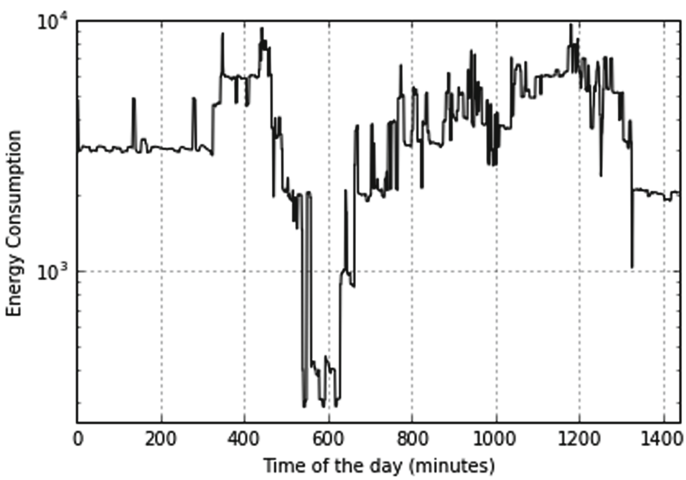


Fig. 1. Energy consumption of a particular household from Database

(*510* min) till *11:00* (*660* min). If someone gets access to this information leads to privacy and security issue for the customers who information has been stored in the cloud environment.

3 Database

The database consists of electricity consumption of households over the period of *4* months between *February 2017* and *May 2017* for *119* days. Dataset consists of smart meter reading of electricity consumption of a house for every *10* s accumulating to *8640* (*24*60*6*) for each day. A total number of data points for a household over the period of *4* months is equal to *119*8640 = 1028160*. Smart meter data is collected from *40* households from a locality and this information is been used to evaluate the energy informatics research purposes.

As smart meter data includes the various patterns and behaviour of the customer, proper care has taken in anonymizing the data to protect the customer anonymity with some large random sequences to represent them such as *'06749…2b8b'* with *64* hexadecimal characters. For better management of the data to view, we have calculated the mean consumption of energy for every one minute giving us the *24*60 = 1440* reading for each household in a day.

3.1 Missing Data

Some of the data in this dataset was missing, we did not investigate the reasons for missing data. To handle the missing data we had close look at the data set. Most of the data points were missing like one or two entries in a minute with adjacent entries available. To calculate the missing entries, we have taken the adjacent entries and calculated the average value as shown in Eq. (1).

$$E_{missing} = \frac{E_{missing-1} + E_{missing+1}}{2} \tag{1}$$

Where, $E_{missing}$ is energy consumption missing at some point, $E_{missing-1}$ is energy consumption before missing entry and $E_{missing+1}$ is energy consumption after missing entry.

4 Energy Consumption Analysis

Energy consumption is analysed to provide the better services to the customers. Energy consumptions for various consumers in household vary based on their daily application usage. Figure 2 shows the usage of four different household usages patterns for a particular day. Energy usage for a household is different for each day of the work as shown in the Fig. 3. It might be because of their working patterns or because of their kids didn't go for school. It is hard to understand the reason for volatility in their consumption, but organizations are using these details to predict their future

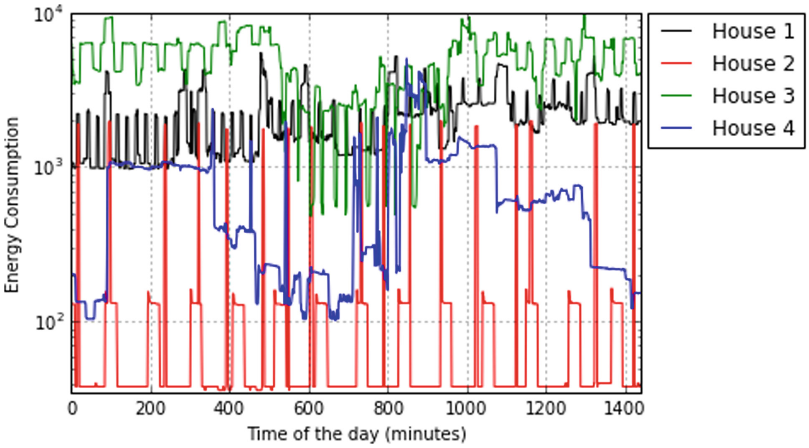


Fig. 2. Overview of energy consumption in 4 household

consumption based on their behavioural patterns. These predictions will help the energy companies to provide the better services with proper load analysis based on the load profiling of all the customers.

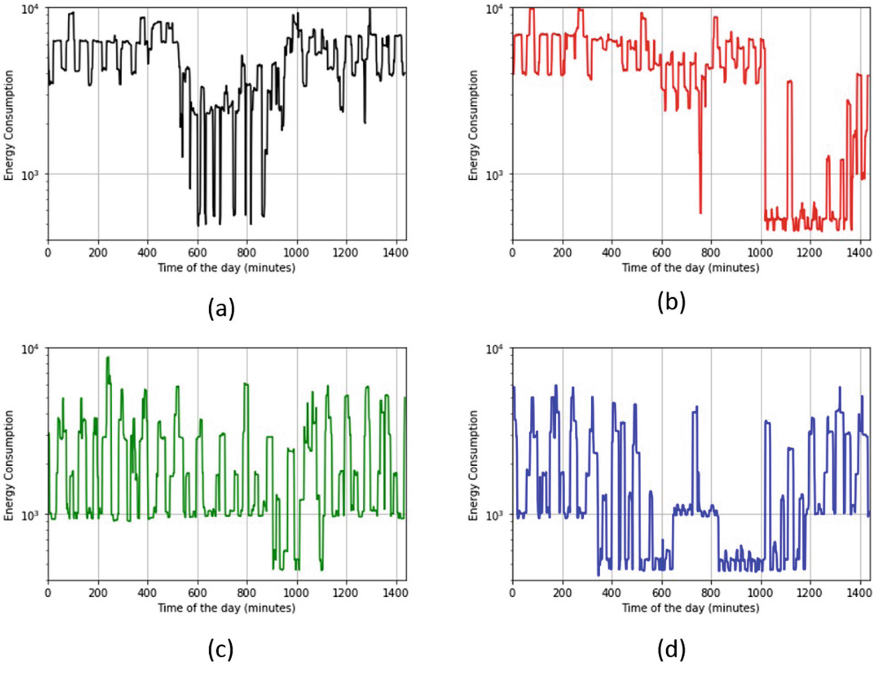


Fig. 3. Energy usage of four consecutive days of same house. (a) Day 1 (b) Day 2 (c) Day 3 and (d) Day 4

Load profiling means the process of classification of customers based on their energy consumption patterns. Many of the researchers have worked on load profiling based on clustering techniques such as, Direct clustering (K-means [4], dynamic clustering with modified k-means [14], self-organizing map (SOM) [5], and hierarchical clustering [6], multi-layered clustering [15]), Indirect clustering (Principal component analysis (PCA) [16, 17]).

In general, as shown in Fig. 4, the energy forecasting process can be divided into 5 steps: Energy Data, Energy Preprocessing, Energy clustering, Evaluation of clustering method, and Results.

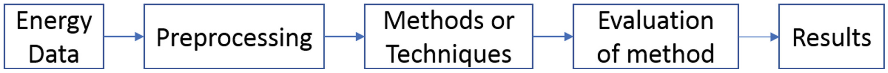


Fig. 4. Overview of energy forecasting process

Step 1 **Energy Data:** The energy consumption of the customers are gathered by energy providing companies or suppliers. This information will be stored somewhere for further processing.
Step 2 **Energy Preprocessing:** The stored Energy data will be processed to filter the bad information which might lead to false results. If some of the information is missing, they will add the information based on their knowledge.
Step 3 **Proposed method:** On this preprocessed energy data, they will apply different clustering techniques or their own techniques to find customers patterns to fit the data and they will fine tune the method to get the better results based on different parameters for those techniques which are been deployed or used.
Step 4 **Evaluation of proposed method:** In this step, the proposed method is evaluated based on the several factors associated with the data. In the evaluation, we will be able to find how the proposed method or technique is performing on the data. This will be measured or calculated on the complete data based on the customer usage patterns.
Step 5 **Results:** The evaluation results will be usually compared with the existing state of the art algorithms. In this step a comparison will be drawn between the proposed method and state of the art to give the reasoning, how much better or worst this new method works compared to other methods. This will give a better understanding of the results for energy profiling on customers usage.

5 Proposed Method

If we look closely at Fig. 2, which provides usage of the various customers on a single day. It clear that none of the households are having the same amount of usage on that particular day. As we can see in Fig. 3, the energy usage of the same household on four consecutive days has completely different patterns. To solve the security issue which can arise by storing the customer energy usage patterns in the cloud environment. We would like to propose a simple data storage method for energy usage in a cloud environment.

We would like to propose that instead of storing the energy usage patterns of each customer separately we would like to store the small group of houses or communities total as one record. If we aggregate the usage of various households and store it in the cloud, even if some gains access to the customer energy usage patterns they will not be able to draw the conclusion which household is using how much energy at any particular time.

To check the applicability of this proposed techniques, we have considered the same four houses which are used in Fig. 2 and calculated the aggregation of their usage for four households same day is presented in Fig. 5. From Fig. 2 and Fig. 5 it is clear that there is no direct relation to any of the households but the energy forecasting companies still can predict the future requirements for this small group of households without losing the quality of the information required by the energy companies.

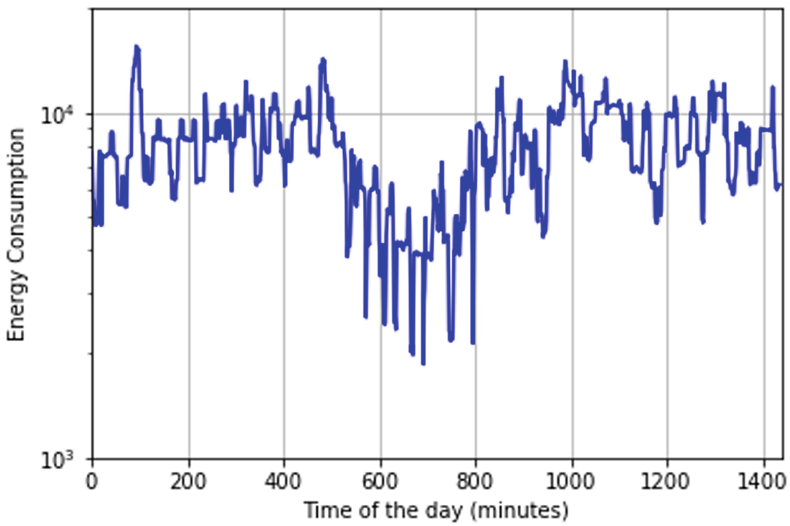


Fig. 5. Total consumption of four households every minute

6 Conclusion

In this paper, we have discussed privacy and security issue which can arise with the storage of energy consumption patterns of the customers in the cloud. If the data is not protected properly or someone else gets access to the data. To solve this problem, we have proposed a simple method to aggregate several households in a community to remove the energy data patterns of the individual households but still represents the same energy consumption pattern of the community or group that has been aggregated. In the future work, we would like to check the proposed method applied and check the predictability of various techniques on the aggregated data.

References

1. Dresselhaus, M., Thomas, I.: Alternative energy technologies. Nature **414**(6861), 332 (2001)
2. Dounis, A.I.: Artificial intelligence for energy conservation in buildings. Adv. Build. Energy Res. **4**(1), 267–299 (2010)
3. Ghelardoni, L., Ghio, A., Anguita, D.: Energy load forecasting using empirical mode decomposition and support vector regression. IEEE Trans. Smart Grid **4**(1), 549–556 (2013)
4. Chicco, G.: Overview and performance assessment of the clustering methods for electrical load pattern grouping. Energy **42**(1), 68–80 (2012)
5. Wang, Y., Chen, Q., Kang, C., Zhang, M., Wang, K., Zhao, Y.: Load profiling and its application to demand response: a review. Tsinghua Sci. Technol. **20**(2), 117–129 (2015)
6. Yang, S.-L., Shen, C., et al.: A review of electric load classification in smart grid environment. Renew. Sustain. Energy Rev. **24**, 103–110 (2013)
7. Marino, D.L., Amarasinghe, K., Manic, M.: Building energy load forecasting using deep neural networks. In: 42nd Annual Conference of the IEEE Industrial Electronics Society, IECON 2016, pp. 7046–7051. IEEE (2016)
8. Mocanu, E., Nguyen, P.H., Gibescu, M., Kling, W.L.: Deep learning for estimating building energy consumption. Sustain. Energy Grids Netw. **6**, 91–99 (2016)
9. Glauner, P., Boechat, A., Dolberg, L., Meira, J., State, R., Bettinger, F., Rangoni, Y., Duarte, D.: The challenge of non-technical loss detection using artificial intelligence: a survey arXiv preprint arXiv:1606.00626 (2016)
10. Ahmad, T.: Non-technical loss analysis and prevention using smart meters. Renew. Sustain. Energy Rev. **72**, 573–589 (2017)
11. Deb, C., Zhang, F., Yang, J., Lee, S.E., Shah, K.W.: A review on time series forecasting techniques for building energy consumption. Renew. Sustain. Energy Rev. **74**, 902–924 (2017)
12. Hong, T., Fan, S.: Probabilistic electric load forecasting: a tutorial review. Int. J. Forecast. **32**(3), 914–938 (2016)
13. Hu, J., Vasilakos, A.V.: Energy big data analytics and security: challenges and opportunities. IEEE Trans. Smart Grid **7**(5), 2423–2436 (2016)
14. Benitez, I., Quijano, A., Diez, J.-L., Delgado, I.: Dynamic clustering segmentation applied to load profiles of energy consumption from Spanish customers. Int. J. Electr. Power Energy Syst. **55**, 437–448 (2014)
15. Al-Jarrah, O.Y., Al-Hammadi, Y., Yoo, P.D., Muhaidat, S.: Multi-layered clustering for power consumption profiling in smart grids. IEEE Access **5**, 18459–18468 (2017)
16. Chelmis, C., Kolte, J., Prasanna, V.K.: Big data analytics for demand response: clustering over space and time. In: 2015 IEEE International Conference on Big Data (Big Data), pp. 2223–2232. IEEE (2015)
17. Koivisto, M., Heine, P., Mellin, I., Lehtonen, M.: Clustering of connection points and load modeling in distribution systems. IEEE Trans. Power Syst. **28**(2), 1255–1265 (2013)

A Systematic Approach to Determine the Amount of Data Required for Asset Management Decisions

Brendan Maestas, Sean Stuntz, Joey Applebee, William Bentley, Jared Breuker, Andrew Davenport, and Andrew Hoisington(✉)

Air Force Institute of Technology, Wright-Patterson AFB, Dayton, OH, USA
andrew.hoisington@afit.edu

Abstract. Facility, infrastructure, and asset related data is being generated at an unprecedented rate, usually without specific purposes or goals. Data is collected in large amounts for exploratory science, achieving significant statistical power, due to the relatively cheap cost of storing data in the cloud. In many cases however, organizations do not consider the negative issues with indiscriminate data collection to include diminishing returns to reduce uncertainty in asset management decisions and the cumulative costs of the data. This paper proposes a novel 4-step frame- work for determining the correct amount of data required for asset management decisions. The framework is built upon the following steps: 1) identify the problem, 2) establish context, 3) verify/collect data, and 4) analyze/decide (IEVA). The IEVA framework can be used as a baseline that orients asset managers to collect decision-focused data and make data-informed decisions.

1 Introduction

Advances in computing and communication technology have propelled the world into an unprecedented information age. For example, the computer system that supports the Large Hadron Collider (LHC), the world's largest particle physics laboratory, can process approximately one petabyte of data every day (European Organization for Nuclear Research (CERN) 2017), equivalent to 210,000 DVDs. The server grid that supports the LHC is only able to actually store 45 petabytes of data, and must rely on networked computers around the world for an additional 15 petabytes (European Organization for Nuclear Research (CERN) 2017). Meaning scientists at the LHC are unable to store 84% of the data that they process, which says nothing about the overwhelming amount of data that is being generated and not processed. Although the magnitude of data processed and ignored by the LHC is certainly an extreme, the trend of collecting more data than feasible to use is common. In the construction and asset management industries, large amounts of data are generated, collected, and stored that is never actually analyzed (Hammad et al. 2014). Examples like these show how data managers around the world, are "drowning in data while thirsting for information" (Herrmann 2001).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 283–289, 2020.
https://doi.org/10.1007/978-3-030-48021-9_32

Datakleptomania is an unconscious desire to collect increasing amounts of data under the premise that more data is better (Hellawell 1991). However, for many businesses, success is the direct result of collecting meaningful data and extracting useful information to support strategic decisions (Woldesenbet et al. 2016). There are some significant statistical benefits of collecting more data. Through increasing the number of observations, researchers are able to achieve greater confidence that their sample is representative of the population. When a sample can be said to be representative of the population, then conclusions about the sample are more likely to be valid for the population (Cohen 1992).

The practice of datakleptomania is enabled by the relatively cheap cost of storing data. As of this writing, Google's cloud service offers the ability to store 100 gigabytes of data for a subscription cost of $12/year (Google 2017). With the low cost and high accessibility of cloud storage, additional companies are storing data and doing business in the cloud. Because of this low cost of data storage, and the technologies that make collecting data even easier, the world is trending toward unquestioned data collection with less scrutiny regarding its necessity (Hanley 2012).

Since data is a resource that can provide value to an organization, data should be purposefully managed using asset management principles. In doing so, organizations will need to view their data within a lifecycle analysis context that considers costs, conditions, and performance of their data from creation to the deletion. However, many data collection activities aren't designed to support decision making processes (Flintsch and Bryant 2009). Data should be collected in such a way that the intent for its use is clearly defined, ensuring that the methods for collection, analysis, and use can be appropriately tailored from the start (Hanley 2012).

2 Problems with Big Data

Beyond the diminishing returns of additional data, the quantity of data that is being collected can be problematic. Datakleptomania can be useful in exploratory sciences where little is understood and questions will be developed later. However, the field of asset management may not be the best example of exploratory science. Instead of collecting every piece of information available, infrastructure asset managers should focus on collecting information that helps answer known questions. Indeed, collection of large quantities of data can lead to 'analysis paralysis' where decision makers have so much data they don't how to move forward with decisions.

Another problem is the time that it takes to process vast quantities of data can still be a challenge. Reducing the amount of data reduces the amount of time it takes to process and analyze that data. This in turn allows decisions to be made faster. Pat Helland explains another problem with analyzing too much data. He states that when the time it takes to process large amounts of data exceeds the window of time in which the decision must be made, the data being processed be- comes obsolete by the time the decision is made (Helland 2011). Helland recommends that in such situations, approximating a good answer can be more valuable than taking time to develop the perfect answer.

Labovitz et al. (1993) developed a notional rule to describe the cost of process correction based on the quality of management to that process. The rule was meant to be applied to a wide range of applications and is known as the 1-10-100 rule (Labovitz et al. 1993). Doyle (2014) tailored this rule to apply it to the costs of da- ta collection and management, namely, that as the quality of data management efforts decreases, the costs of using the data later increase significantly. Data collected early and deliberately with the intent to be applied to later decision-making efforts yield relatively low cost to the organization. The organization has failed to implement effective data management strategies and will more than likely result in poorly informed decisions which must be corrected down the road at a high cost to the organization. Data collected without proper forethought will more than likely result in time intensive analysis and require a large cost in human capital.

Once data is collected, an organization may desire a storage capability to use the date at a later date. Although the storage cost of data has decreased steadily as technology and management tools improve, it is still a portion of the costs (LaChapelle 2016). This data storage capability requires either data infrastructure in- vestment or regular storage fees for use on the cloud. Several factors influence data storage requirements. The organization must first determine the useful life of the data stored and also decide on a meaningful format for analysis and communication. When optimizing data management, an organization must also consider the lifecycle cost of storage. The cost of keeping a single gigabyte of data indefinitely can cost $100 (Omaar 2017), a cost far greater than the monthly cost of using the Google cloud or similar service. This cost is further compounded for organizations whose security policies prohibit the use of the cloud for storage purposes. Redundancy and security also play a heavy role in this process. The organization must build into their storage plan their requirement of backing up the data and ensuring its security from outside threats. They must decide how much risk they are willing to take and in which areas they are willing to accept this risk.

A final problem with big data occurs during data transformation, which is the process of transitioning from a set of stored raw data to usable data ready for analysis. Although it is recommended that the data is already in a useable format when collected, it may not be possible until new equipment or systems are installed. In this transformation effort, there are generally aspects of the raw data which are lost due to translation or conversion issues inherent with diverse formats and equipment systems. The goal is to perform these transformation efforts with minimal costs to the organization by way of lost or mistranslated data. Establishing a consistent transformation strategy to minimize these losses is critical to an organization ensuring their efforts are effective and useful. This transformation cost can also be incurred if the organization decides to upgrade operating systems or interfaces. In these situations, a well-developed plan and strategy are critical to successful transformation of data with minimal loss of information. This effort can be costly and time consuming in and of itself, but will be worth it in the future.

3 Notional Framework for Right Size Data

Any organization which seeks to optimize their decision making and data collection strategies benefit from a simple and effective framework to guide them through that process. A decision-making framework allows organization to simplify their data collection and analysis process. This in turn will enable trend tracking and decision making in the future by eliminating shortfalls in their data, eliminating unnecessary data, reducing decision making time and cost, and do this all to the organization's self-set standards. The 4-step IEVA framework presents a strategic level approach for an asset manager to identify the problem, establish the context of the problem, verify and collect appropriate data, and finally analyze and make a decision. Each step of this strategic asset management decision making framework can be seen in Fig. 1 and is explained in further detail throughout this section.

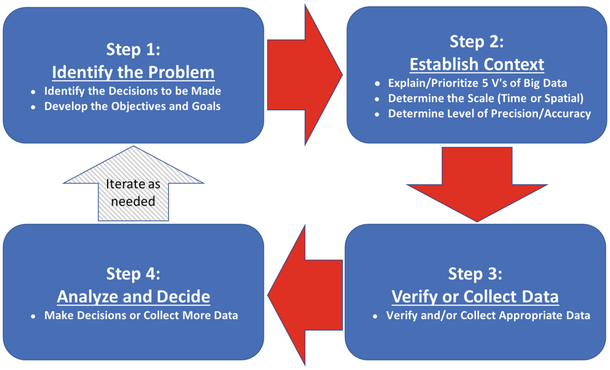


Fig. 1. IEVA framework steps

3.1 Step 1: Identify the Problem

The initial step of this strategic framework is determining the overall decision to be made. Asset managers and their respective teams must brainstorm during this step to narrow down and identify the overall decision to be made. Once the decision is understood, the next step is to develop objectives and goals, so that the decision can be fully developed. Finally, this decision should be emphasized throughout the 4-step framework.

Now that a decision has been identified, set goals and objectives specifically for this decision. These goals will be specific to each asset manager and their respective objectives. For example, if the decision is to save energy costs across an installation, the goal could be to reduce electrical usage by twenty percent in the next five years. To increase effectiveness, goals and objectives should be specific, measurable, attainable, relevant, and time bound (SMART) (Mind Tools Content Team 2014). A specific goal must be clear and well defined, and this goal should guide decision making throughout the rest of the framework. A measurable goal is one that sets a clearly defined objective. In addition to measurable, the goal should be attainable and realistically accomplished.

A relevant goal will help keep the decision in mind and keep the goals focused. Finally, a time- bound goal will create a sense of urgency and allow for an achievement if and when that goal is met.

3.2 Subheading: Step 2: Establish Context

Once the problem is understood, a user should prioritize the five V's of data for the specific objective and goal (Marr 2015). The five V's are volume, velocity, variety, veracity, and value. Volume is the vast amount of stored data which is gathered every day from multiple sources. The value of data is defined simply as the ability of a person to convert raw data to performance metrics. When applying a value determination to this framework's steps, it is important to ensure that the data which will be leveraged is being collected in the first place. Velocity refers to the speed at which data is gathered, which can be dependent on the equipment or data storage capacity. Veracity is the trustworthiness of the data. Lastly, variety refers to the multiple different types of data which can be collected (Marr 2015).

Overall, examining the five V's of data provides insight on how data is collected and measured (Marr 2015). It is important for users of this framework to ask these data questions before going ahead in their decision making. The order at which the five V's are considered are dependent upon the questions being asked and they should be weighted accordingly. The overall goal of this step is for the user to look at the question they are asking and mold the data collection for this purpose. In many cases where the five V's are not considered, the available data forces decision makers to alter the question they originally wanted to ask (Marr 2015).

Understanding the scale of the data will aid in this process to determine the right amount of data needed to make a decision. This step is to ensure there are boundary conditions for the problem being solved. As an organization steps through the process, it can be easy to add scope to the original question. Data scales can either be spatial or time related. A spatial scale would evaluate or collect data based on regional zones, facility types, utility infrastructure, or rooms in a facility. A time scale refers to where the beginning and end of the useful data occurs. Overall, any scale could be used by the decision-making team or individual, and picking one and justifying it will help make the overall decision and help identify the amount of data needed for it.

In addition to scale, precision and accuracy are important factors to consider to create a useful framework. Precision can be modelled by looking at the variance of the data set. As the variance increases in the data set, the data is less precise. On the other hand, accuracy is the trueness of the data set. The data's velocity can change the accuracy and precision of the data due to the number of samples that are usable. Once the five V's have been adjudicated, the next step is to collect data.

3.3 Step 3: Verify or Collect the Data

Deciding how to collect and aggregate data contributes to data management life-cycle costs and data quality. There are many software options which offer a degree of automation and analysis. It is therefore pertinent for each asset manager to research these options and balance the potential for errors during collection with software acquisition costs. In addition to software, some data collection may be dependent upon manpower, also contributing to increased data acquisition costs. Overall, there are many different ways to collect data, and each asset manager must balance which way is most efficient and practical for the decision they are making using this model.

3.4 Step 4: Analyze and Decide

After the appropriate data has been collected and consolidated, the next step of the framework is data analysis. This analysis will be dependent upon the objectives and questions which were developed earlier in the framework. These objectives will aid in distinguishing which statistical measurements and tests are required for a proper conclusion to be made. Each asset manager's analysis will be completely dependent to the overall decision being made. There are multiple methods of statistical analysis that can be simple or complex requiring different levels of mathematical foundation. Understanding the desired method and using any available aids, such as excel and JMP software, will decrease the time required and increase the useful output of the analysis effort.

The final step of this framework is to make the decision. At this point, the user of the framework has collected the data and analyzed it accordingly. The data has been specifically gathered to ensure the overall question will be answered in an effective and efficient manner. However, if the situation presents itself where data is insufficient, the team should return to Step 2 to re-evaluate what needs to be measured to accomplish the goal. This is designed to ensure that the data being collected truly relates to the original objective set forth by the team. If the data is sufficient and executable, the team shall move to the decision-making.

These steps were created to ensure that data collection is closely monitored and gathered with a specific purpose. The best-case scenario of this framework is that the user finishes the steps and comes to a data-validated answer to their specific question. They have the ability to make a decision and influence positive changes in their organization. In a worst-case scenario, the user will complete the frame- work with a better understanding of how to effectively return to the prior steps to ask a question better suited to the desired results.

4 Conclusion

Modeling a data management framework is an important tool for asset managers to sift through the large data quantities being collected. The 4-step IEVA framework provides a method to give realistic goals and objectives for asset man- agers to utilize while limiting the effects of uncertainty, quantity of data, and cost factors. The 4-step IEVA framework is the first attempt to address the issues of big data and utilizing a system to tailor the data for effective analysis. For future research, this framework can be applied to other asset management decisions regarding condition assessments for facilities and utility systems. Regardless of future applications, the 4-step IEVA process provides a framework for determining the amount of data required for asset management decisions.

Disclaimer
The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States Government.

References

Cohen, J.: Statistical power analysis. Curr. Dir. Psychol. Sci. **1**(3), 98–101 (1992)

Department of Defense: DoD Building Code (General Building Requirements). Washington (2016)

Doyle, M.: Why Data Should Be a Business Asset – The 1-10-100 Rule (2014)

European Organization for Nuclear Research (CERN): Computing | CERN (2017). https://home.cern/about/computing. Accessed 8 Dec 2017

Flintsch, G.W., Bryant, J.: Asset management data collection for supporting decision processes asset management data collection for supporting decision processes, pp. 2–97 (2009)

Google: Google Drive Pricing Guide (2017). https://www.google.com/drive/pricing/. Accessed 8 Dec 2017

Hanley, J.R.: Marine environmental monitoring programs: tips on design, measurement and interpretation. APPEA J. **52**(1), 317 (2012)

Helland, P.: If you have too much data, then 'good enough' is good enough. Commun. ACM **54** (6), 40 (2011)

Herrmann, K.R.: Visualizing Your Business: Let Graphics Tell the Story. Wiley, Hoboken (2001)

Hubbard, D.: How to Measure Anything: Finding the Value of Intangibles in Business. Wiley, Hoboken (2014)

Labovitz, G.H., Chang, Y.S., Rosansky, V.: Making Quality Work: A Leadership Guide for the Results-Driven Manager. Harper Business, New York (1993)

LaChapelle, C.: The cost of data storage and management: where is it headed in 2016? Data Cent. J. (2016)

http://www.datacenterjournal.com/cost-data-storage-management-headed-2016/. Accessed 5 Dec 2017

Marr, B.: Why only one of the 5 Vs of big data really matters. IBM Big Data & Analytics Hub (2015). http://www.ibmbigdatahub.com/blog/why-only-one-5-vs-big-data-really-matters. Accessed 8 Dec 2017

Mind Tools Content Team: Five Golden Rules for Successful Goal Setting (2014). https://www.mindtools.com/pages/article/newHTE_90.htm. Accessed 8 Dec 2017

Omaar, J.: Forever isn't free: the cost of storage on a blockchain database. (2017). https://medium.com/ipdb-blog/forever-isnt-free-the-cost-of-storage-on-a-blockchain-database-59003f63e01. Accessed 5 Dec 2017

Top-Down Practical Methodology for the Development of Asset Information Requirements

J. Heaton(✉) and A. K. Parlikad

Institute for Manufacturing, University of Cambridge, Cambridge CB3 0FS, UK
{jrh212, aknp2}@cam.ac.uk

Abstract. During the last decade Building Information Modelling (BIM) has emerged as a key enabler within the construction industry to increase productivity. While the benefits have been mostly realised within the Design/Construction phase, the adoption of BIM within the operational phase has been significantly limited. One of the critical challenges to utilising BIM within the operational phase is for asset owners/maintainer to develop information requirements that support the BIM process. UK BIM standard PAS 1192-3 provides information requirements frame that states that an organisation should develop Organisational Information Requirements (OIR) that then aid in generating individual Asset Information Requirements (AIR). While it is critical that the OIR generate the AIR, it can be witnessed that this is too much of a leap for most organisations. This paper proposes a practical Top-Down methodology that aids in bridging this *"gap"* by utilising an organisational point of view of assets as the functional output they support and proposes the development of Functional Information Requirements (FIR) to bridge this gap.

1 Introduction

The importance of information management within the engineering asset management domain is gaining momentum both in academic literature and industry applications. Practical applications for information management processes are being guided by an array of industry standards that solely focuses on information management within the individual life-cycle phase of engineering assets (British Standards Institute, 2007, 2013, 2014a, 2014b, 2015). Most noticeably PAS 1192-3:2014 (British Standards Institute, 2014b) developed by the British Standards Institute (BSI)[1] looks at the use of Building Information Modelling (BIM) exclusively within the operational phase of an asset's lifecycle. BIM has emerged as an effective solution for collecting, validating and storing asset related data within a collaborative environment for the whole-life of a given asset. BIM has most successfully been adopted and implemented within the design and construction phase, with widespread evidence demonstrating a reduction of design/construction cost, increase in productivity and improved management of risk (Azhar et al. 2008; Zhou *et al.* 2009; Smith 2014). Despite this success, the adoption of

[1] https://www.bsigroup.com/.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 290–301, 2020.
https://doi.org/10.1007/978-3-030-48021-9_33

BIM within the operational phase is relatively limited (Waterhouse and Philp 2016). Furthermore, there has been a noted increase in the amount of research effort aimed at the adoption of BIM in the operational phase, but this is having little impact in practical applications (Pärn et al. 2017).

The complexity of the adoption of BIM within the operational lifecycle of an asset is multipronged. A recurring theme is a fundamental ability to articulate and demonstrate the value of BIM adaptation within an organisation (Barlish and Sullivan 2012). This is partly because asset owners struggle to identify the information requirements during the BIM development stage, leading to the development of BIM processes that generate little to no value for the organisation. Organisational Information Requirements (OIR), if developed at all, are generally in the form of technical guidance/support (including statute requirements) and don't consider the organisational context within their development (Cavka et al. 2017). Furthermore, if the OIR is not lead by organisational requirements, then the development of the Asset Information Requirements (AIR), which is generated from OIR will not be fit for purpose for use within organisational processes such as capital investment decisions, customer management, maintenance management, etc. Indeed, due to the highly complex and diverse nature of the asset management industry, it can be seen that the effort to extend the BIM information management concept (within PAS 1192-2) into the operation phase has fallen short of providing the required knowledge for industry adoption (Kiviniemi and Codinhoto 2014). Therefore there is a clear need to provide a robust and practical methodology that provides asset owners with the tools for development of organisational lead information requirements, while still maintaining quality management requirements.

The objective of this paper is to provide the findings from a research effort that investigates the development of organisation lead AIR that supports BIM enabled processes within the operational phase of an asset. The principal objective is to understand the challenges in adopting BIM within the operational phase that generates direct value for the organisation. The research starts by formalising a conceptual approach that supports the creation of a relationship between the organisational requirements and broader context with the engineering assets within the organisation portfolio. Furthermore, a process is proposed that utilise the organisational point-of-view of an asset as the functional output it provides and proposes the intermediate step of developing Function Information Requirements (FIR) that aims to provides the bridge for an organisation for generating its OIR into individual AIR. Finally, a methodology is proposed that aids in the practice development of the FIR by utilising the information management processes with the UK BIM Standards and asset classification within UNIClass 2015 (Delany 2015). This paper uses BIM as both the data model (Building Information Model) and the process of developing the information model itself (Building Information Modelling). Furthermore, the research also explores the possibility of BIM defined as Building Information Management, the process of managing the information model throughout its whole life cycle.

This paper consists of 4 sections including this introductory section. Section 2 provides a comprehensive review of the use of BIM within the operational phase including current challenges and state of the art. Section 3 introduces the practical Top-

Down methodology for the development of AIR, Sect. 1 summaries the key findings and proposes future research opportunities.

2 Background of BIM in the Operational Phase

There is a greater understanding that the adoption of BIM can deliver significant economic, environmental and social benefits. As an asset spend the vast majority of its life within the operation and maintain (O&M) phase, ultimately the client and end user are the primary beneficiaries of BIM (Eastman *et al*. 2008). The UK government 2016 construction strategy references the use of BIM and the development into BIM level 3 with a whole-life cycle approach as a critical enabler to increase productivity and reduce operational costs while proving the platform for smart cities development (The Infrastructure and Projects Authority 2016). Furthermore, from a whole-life costing point-of-view, it can be seen that the O&M phase contributes to 60%–80% of the overall cost, significant economic saving can be achieved by focusing on the O&M phase. This is future enforced by (British Standards Institution, 2008) which states that 80% of the O&M costs can be influenced by the first 20% of design, reinforcing the need for consultation of the O&M phase within Design and Construction.

A significant portion of academic research is firmly focused on solving the technical challenges, including integrating BIM related data within a computerised As-set Management System. BuildingSMART[2], a not-for-profit industry lead organisation has developed an open source file format for the exchange of BIM related information but has noted limitations within the operational phase (Becerik-Gerber et al. 2011; ISO 2013). The adoption of Construction Operations Building Information Exchange (COBie) standard (initially published in the US 2007 and adopted as a British Standard 2014) by the UK Government, aids to standardise the transfer of newly built assets into an asset management system by providing a basic Common Separator Value (CSV) template as data a transfer protocol (British Standards Institute 2014a). While COBie provides a spreadsheet environment that is well understood and used for many organisational functions, it has limitations in visualising relationships between different data points and complexity in populating the request information over many stakeholders and within a complex supply-chain (Mehmet Yalcinkaya 2016). Furthermore, the integration of COBie and related BIM geometry within a computerised Asset Management system is limited.

While it is well cited that the data integration and transformation challenge is hindering the developing of BIM within the operational phase, it has not stopped a set of novel tools being developed. (Pärn and Edwards 2017) Has developed an Application Program Interface (API) that embeds operational related data directly into geometry placed within the BIM model. The using of linking Radio Frequency Identification (RFID) has been demonstrated within the construction industry; originally it has been used as a tracking and mapping tool for building elements (Ergen et al. 2007). More recent use of RFID has seen them used as a link between the physical asset and

[2] https://www.buildingsmart.org/.

the BIM environment by linking the RFID codes within the BIM model, this allows for whole-life tracking of digital information linking to an asset and visualisation of that asset within the BIM model (Meadati et al. 2010; Fung 2013). (Motamedi et al. 2014) takes advantage of the visualisation aspect of BIM by linking the BIM model to a knowledge root course problem database that classifies and visualises operational failures directly within the BIM model.

In parallel to the above technical research, there is a growing focus on the management of BIM progresses from an organisations point-of-view on how best to adopt the BIM principles within a whole-life organisational context. While not directly related to BIM, ISO 55000 provides the management framework for managing physical assets throughout an organisations whole-life cycle (ISO 2014). Due to the complex nature of asset management which encompassing multidiscipline activities there is an extensive need for the development of information requirements that integrate the operational and organisational requirements (Ashworth *et al.* 2016). (Becerik-Gerber et al. 2011) Attempts to provide some structure by developing a hierarchy framework to structure nongeometric data. Furthermore, a framework is presented that groups the data as per its use within the project lifecycle and the responsible stakeholders. From an owners' requirements point-of-view, (Cavka et al. 2017) developed an information requirements landscape from the owners' perspective and classification of owners requirements, this is further developed into computable requirements for helping to identify what information to include within the BIM model.

While there is unmistakable evidence of the benefits of adopting BIM for Asset Management and Facility Management, there is a lack of understanding in how BIM can provide practical tools for use within the O&M phase. An annual survey conducted by the National Building Specification of BIM professionals shows that 65% see using BIM within the O&M phase will aid in gained efficiencies, but 72% of clients do not understand the benefits of BIM (RIBA Enterprises 2017). Furthermore, lack of client demand can be seen as one of the leading barriers to the adoption of BIM in the O&M phase.

3 Bridging the Gap: Practical Top-Down Methodology

As highlighted in Sect. 1 and 2 of this paper, there is a fundamental lack of understanding of how to harness the value of BIM within the O&M phase. One of the key challenges is for organisations to develop OIR that aid in the development of individual AIR. Often AIR our developed from technical documentation and not organisation requirements, this creates a siloed effect from the organisational context and its physical asset portfolio. Furthermore, it can also be witnessed that the OIR is often developed in isolation and does not align with the organisational strategies, plans and objectives.

A high-level methodology was developed to support the formalisation and alignment of organisation information requirements, Fig. 1 provides this overview. Furthermore, the proposed methodology also supports the conceptual relationship between the organisational context and its physical asset portfolio. The methodology incorporated information management frameworks from asset management standards ISO 55000 and BIM related standards PAS/BS 1192. The methodology is diverted into two discrete parts. Firstly, identifying and classifying the engineering assets within the organisation. Secondly, identify, align and validating the OIR.

1. Identify and classify the range of functional outputs that are supported by the asset portfolio. Following this, the key asset systems that support the functional output will be identified
2. investigate and identify key organisational requirements by reviewing and extracting requirements from key data sources such as the strategic asset management plan, growth (financial) strategy, environmental strategy and organisational objectives.
3. Identify and capture information that aids in the organisation to create informed decisions around organisational requirements and objectives, e.g. OIR.
4. Identify and capture functional information requirements as identified in step 1, in alignment to the OIR developed in step 3.
5. Identify and capture AIR against the asset systems as identified in step 1, in alignment to the OIR developed in step 3.
6. Validate the capture of AIR is efficient to support the organisation requirements.

Asset Classification – Firstly, it is required for the organisation to develop a classification system within its asset portfolio. Traditionally this is completed by focusing on the individual instances of assets, this is a daunting task as it can quickly start to identify thousands of assets. It is proposed the first task of classification should focus on the functional output that the assets provide, e.g. a radiator will provide the function of heating. A recent update to the UNIClass classification system (Table E/F) provides a table of 76 functions that are supported by asset systems and instances (Delany 2016). Organisational requirements are focused on the functional performance of a given asset system, the vast majority of stakeholders interact with the assets functional output and not the asset systems or products. Despite this, as an array of different asset systems supports most functions, it is required to capture information at the asset system level to support the organisational requirements. UNIClass Table S/s classifies 2085 asset systems that support the 76 functional outputs. In some unique organisational requirements, it might be required to capture information at the asset produce level, such example might be for legal and statute requirements. The organisational value of classifying at the product level should be strongly validated, as it is a costly and complicated task that could see thousands of products classified within individual asset systems that don't support an organisation requirement. If it deemed appropriate to classify at the product level, UNIclass provides Table P/r that classifies 6869 products that link back to asset systems and functional output. The critical part of the asset classification process is to start from the function output downwards and not from the product upwards. A summary of the relationship between asset classification and key stakeholder is provided in Fig. 2.

Identify Organisational Requirements – the second step is to identify the key organisational requirements, depending on the size and the context of the organisation this could be spread over several documents within discrete and often siloed departments with individual requirements. The data collecting strategy including both quantitative and qualitative data sources. A sample of some data sources that should be analysed includes annual reports, design handover requirements, environmental impact

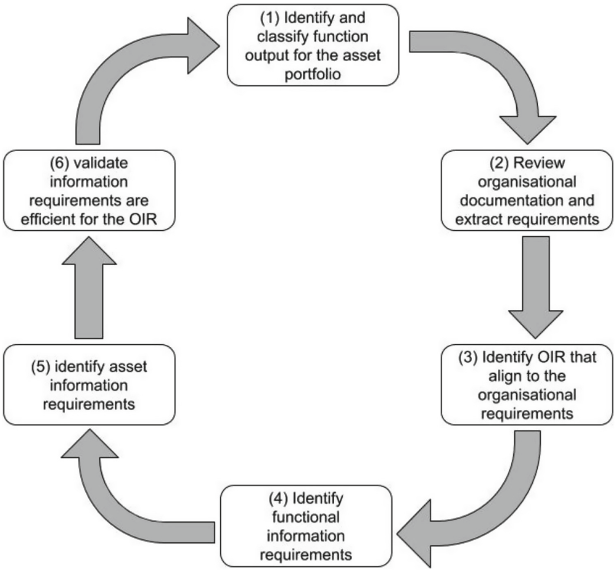


Fig. 1. Information requirements methodology

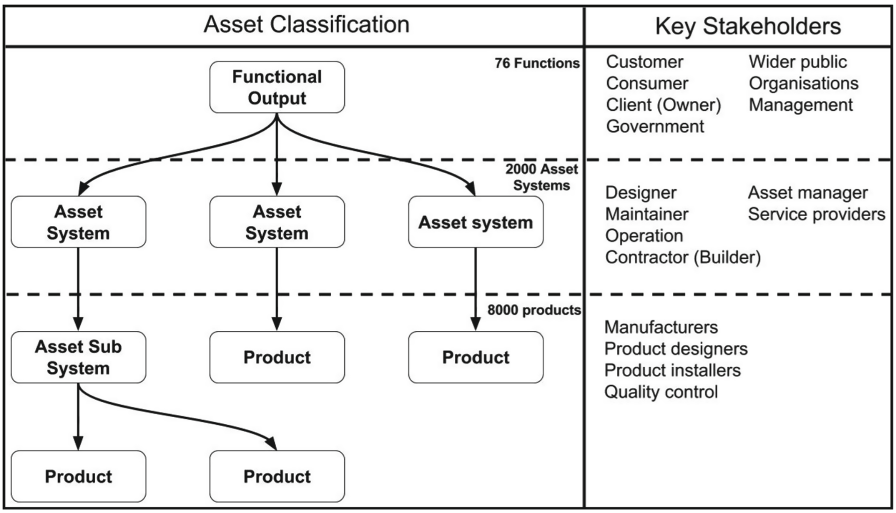


Fig. 2. Stakeholder's interaction with different classification levels of a physical asset

strategy and organisational reports. If the organisation has established an asset management system as per ISO 55000, then the documented produced from this process should be prioritised. Specifically, both organisational and asset management objectives should be identified as they provide measurable organisational requirements. Once the organisational requirement has been extracted, there is a need to categorise them into one of 4 categories including Financial, environmental, operational and reputational. This process allows organisational to structure their requirements; this is more critical for complex organisations that could easily have over 100 requirements. Furthermore, it will also highlight any gaps in the requirements identifying the process. The principal data sources are summarised below in Table 1.

Organisational Information Requirements – the third step, is to identify the information requirements that are needed to validate that the organisation is meeting (or not) the identified organisational requirements. Using the organisation requirements extracted from step two we can align information requirements directly with an organisation requirement. The OIR should not focus on specific asset functions or types, but be general high-level requirements. When all the OIR are aggregated together they should help to create greater informed decisions for the organisation requirement, if a bit of information is not helping to aid that decision, then it should not be collected. To support the development of the OIR, an information requirement matrix capture has been developed. The matrix utilises the extracted organisational requirements, categories organisational requirements, the classified asset functional outputs and information requirements categories (financial, technical and managerial). The matrix is completed within interviews and workshops with key personnel within the organisation that has extensive knowledge of the organisational objectives and key goals. The information requirements categories are derisive from BIM standards and provide a structured approach to devolving the capture of information requirements throughout the whole organisation. To support the link between organisation requirements to OIR and AIR, we highlight the top three asset functions defined in step 1 that will have the most significant impact on achieving (or not) the organisational requirement. The number of functions aligned to the organisation requirements can increase from three if required but should be limited to no more than five to ensure it is achievable within a reasonable timeframe. An example of the information requirements matrix is shown in Fig. 3.

Functional Information Requirements – The fourth step, is to identify the information requirements as per the functional outputs defined in step one and aligned to the organisational requirements in step three. FIR our captured utilising the same style information requirements capture matrix as presented in the OIR step. Within the FIR matrix we are not focused on the organisation requirement but the functional output that was highlighted in the OIR exercise. One added process to developing the FIR is to utilise the framework of whole-life management including the design, construction, operational and maintenance and disposal/renew phase. Whole-life is a well-defined concept and is used by many organisational to manage the lifecycle of their physical assets. The primary objective of the FIR is to provide the bridge between the OIR and the AIR. It should be generic to the point that it does not focus on asset systems or individual produces but solely on the functional output which is being analysed. To support this, it is required to engage with personnel within the organisation that has a good understanding of the asset portfolio and the primary functional output it provides. Creating the link from the FIR to the AIR, it is required to capture the asset systems that support the functional output as defined in step 1, this supports the direct link from the OIR to the FIR then finally the AIR. Figure 4 shows an example template for the FIR capture matrix.

Table 1. Organisational data sources

Sources of requirements	Description
Strategic asset management plan	Key documentation of ISO 55000 asset management requirements. Containing asset management objectives aligned to the organisational objective
BIM execution plan	Contains design and construction requirements that can be used in the operational phase
Environmental strategy	Organisational environmental framework and objectives to minus impact on the natural environment
Organisational business plan	Key strategic documentation that outlines the business financial growth plan and objectives
Customer engagement strategy	Provides the framework and key objectives for engagement with customers and end users
Information/technology strategy	Highlights vital objectives of digitalisation in-line with the organisational requirements

Information Requirements	Organisational Requirement		Category
Managerial Information			
Technical Information			
Financial Information			
Asset Functions	Function 1	Function 2	Function 3

Fig. 3. Organisational information requirement capture matrix

Asset Information Requirements – the fifth step, is to identify the information requirements as per the asset systems as defined in step one and alignment to the functional asset output. The AIR uses the same information requirements matrix as per step four, for development of the FIR. The information captured should focus on the specific asset system and not the asset functional output or individual products. All information should align with the original OIR and the associated organisation requirement. To support the development of the AIR is it requested to engage with personnel within the organisation that have a strong technical understanding of the specific asset systems being analysed. This task should be repeated for any asset products that are identified within the asset system as highlighted in step one.

Validation – the final step is to validate that all of the required information has been captured. This is achieved by developing a prototype of a sample set of objectives,

Function 1		Objective	
Life Cycle	Managerial Information	Technical Information	Financial Information
Design			
Construction			
Operational / Maintenance			
End-of-Life			
Asset Systems	System 1	System 2	System 3

Fig. 4. Functional information requirement capture matrix

OIR and AIR. The prototypes are discussed during a set of workshops with key stakeholders including asset managers, senior management, maintenance team, commercial management and customer management. Feedback from the workshops is built into updated versions of the prototype. This step is repeated until all members of the workshop are satisfied

4 Conclusion

It can be witnessed within the literature review that the barriers to BIM adoption within the operational phase is multiprong. One of the critical challenges is the organisation's ability to harness the value of BIM within the operational phase, this is because the organisational context and requirements are not considered within the BIM development processes. To address this challenge, a practical top-down methodology is proposed that aids to support the development of AIR that are aligned with organisation requirements. Utilising the organisations point-of-view of assets as the functional output they support, the methodology proposes the developed of Functional Information Requirements. The FIR act as a bridge between the OIR and AIR, allowing direct line-of-sight from organisational requirements, OIR, FIR and AIR. Depicting assets as a functional output allows for none technical organisational leading personnel to greater understand the broader impact assets have on the organisational requirements and therefore develop associated information requirements. Fundamentally, the proposed methodology allows for organisations to create greater informed decisions in relations to their requirements by distilling these requirements into individual AIR that address an OIR via the FIR.

A key development of this methodology is allowing the adoption of BIM within the operational phase of assets. This is achieved by bridging the gap between the OIR and the AIR. Furthermore, the concept of whole-life is incorporated within the developed of

FIR, supporting the development of information requirements throughout the whole lifecycle of an asset.

As an example, an organisational objective to reducing operational cost by 5% could be aligned to the function of heating, which would produce specific heating functional information requirements such as airflow measures, performance requirements and criticality. Furthermore, the mutable heating systems that support the function of heating such as gas heating systems, heat exchanges and underfloor heating systems, will each have their own specific asset information requirements.

Future work will use the information requirements capture within this methodology to develop an Asset Information Model (AIM) that support whole-life asset management and the use of BIM processes within the operational phase. Furthermore, this model will define the relationship between the physical asset, the asset life-cycle and alignment to the organisation requirement. One possible approach is to demonstrate the AIM as an ontology, this has the flexibility of defining the relationships between different sets of information and not how that information should be structured. Several ontologies have been developed within the BIM domain (Park *et al.* 2013; Lee et al. 2014). A more modelled based approach could be achieved by using Unified Modelling Language (UML). UML is a general-purpose modelling language that is used in software engineering, which provides a standardised way to design and visualise a system or network. One key advantage within UML is the use of behaviour diagrams that are used to describe the functionality of a system with activities and use case diagrams.

References

Ashworth, S., Tucker, M., Druhmann, C., Kassem, M.: Integration of FM expertise and end-user needs in the BIM process using the employer's information Requirements (EIR). In: Proceedings of CIB World Building Congress 2016, vol. 5 (2016)

Azhar, S., Nadeem, A., Mok, J.Y.N., Leung, B.H.Y.: Building Information Modeling (BIM): a new paradigm for visual interactive modeling and simulation for construction projects. In: Proceedings of the First International Conference on Construction in Developing Countries (2008)

Barlish, K., Sullivan, K.: How to measure the benefits of BIM - a case study approach. Autom. Constr. **24**, 149–159 (2012). https://doi.org/10.1016/j.autcon.2012.02.008

Becerik-Gerber, B., Jazizadeh, F., Li, N.: Application areas and data requirements for BIM-enabled facilities management. J. Constr. Eng. Manag. **138**(March), 431–442 (2011). https://doi.org/10.1061/(ASCE)CO.1943-7862.0000433

British Standards Institute: BS 1192-2007 +A22016: Collaborative production of architectural, engineering and construction information (2007)

British Standards Institute: PAS 1192-2:2013 Specification for information management for the capital/delivery phase of construction projects using Building Information Modelling, vol. 1, pp. 1–68 (2013). http://www.carillionplc.com/media/105185/building_information_modelling.pdf

British Standards Institute: BS 1192-4: 2014 Collaborative production of information Part 4 : fulfilling employer's information exchange requirements using COBie - Code of practice, British Standards Institute, p. 58 (2014a). http://shop.bsigroup.com/forms/BS-1192-4/

British Standards Institute: PAS 1192-3:2014 - specification for information management for the operational phase of assets using building information modelling, British Standards Institute, vol. 1, pp. 1–44 (2014b)
British Standards Institute: PAS 1192-5-2015_Specification for security- minded building information modelling, digital built environments and smart asset management, British Standards Institute (2015)
British Standards Institution: Bs iso 15686-5:2008 Buildings and constructed assets—Service-life planning—Part 5: Life-cycle costing, p. 56 (2008)
Cavka, H.B., Staub-french, S., Poirier, E.A.: Automation in construction developing owner information requirements for BIM-enabled project delivery and asset management. Autom. Constr. **83**(June), 169–183 (2017). https://doi.org/10.1016/j.autcon.2017.08.006
Delany, S.: UNICLASS Calssification, NBS. https://toolkit.thenbs.com/articles/classification Accessed 15 Nov 2016
Eastman, C., Liston, K., Sacks, R., Liston, K.: BIM Handbook (2008). 2007029306
Ergen, E., Akinci, B., Sacks, R.: Tracking and locating components in a precast storage yard utilizing radio frequency identification technology and GPS. Autom. Constr. **16**(3), 354–367 (2007). https://doi.org/10.1016/j.autcon.2006.07.004
Fung, M.A.: Application of BIM and RFID in Public Housing Projects, Hong Kong Housing Authority (2013)
ISO: ISO 16739:2013 - Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries (2013)
ISO: BS ISO 55000 series - Asset Management (2014). https://doi.org/10.1177/0032329211420047
Kiviniemi, A., Codinhoto, R.: Challenges in the Implementation of BIM for FM—Case Manchester Town Hall Complex, Computing in Civil and Building Engineering (2014), pp. 665–672, September 2014. https://doi.org/10.1061/9780784413616.083
Lee, S., Kim, K., Yu, J.: Automation in Construction BIM and ontology-based approach for building cost estimation. Autom. Constr. **41**, 96–105 (2014). https://doi.org/10.1016/j.autcon.2013.10.020
Meadati, P., Irizarry, J., Akhnoukh, A.K.: BIM and RFID integration: a pilot study. In: Second International Conference on Construction in Developing Countries (2010)
Mehmet Yalcinkaya, V.S.: Evaluating the usability aspects of Construction Operation Building Information Exchange (COBie) standard evaluating the usability aspects of Construction Operation Building Information Exchange (COBie) Standard, June 2016
Motamedi, A., Hammad, A., Asen, Y.: Knowledge-assisted BIM- based visual analytics for failure root cause detection in facilities management. Autom. Constr. **43**, 73–83 (2014). https://doi.org/10.1016/j.autcon.2014.03.012
Park, C.S., Lee, D.Y., Kwon, O.S., Wang, X.: A framework for proactive construction defect management using BIM, augmented reality and ontology-based data collection template. Autom. Constr. **33**(August), 61–71 (2013). https://doi.org/10.1016/j.autcon.2012.09.010
Pärn, E.A., Edwards, D.J.: Conceptualising the FinDD API plug-in: a study of BIM-FM integration. Autom. Constr. **80**, 11–21 (2017). https://doi.org/10.1016/j.autcon.2017.03.015
Pärn, E.A., Edwards, D.J., Sing, M.C.P.: The building information modelling trajectory in facilities management: a review. Autom. Constr. **75**, 45–55 (2017). https://doi.org/10.1016/j.autcon.2016.12.003
RIBA Enterprises, National BIM Report, vol. 5, pp. 1–28 (2017). https://doi.org/10.1017/cbo9781107415324.004
Smith, P.: BIM & the 5D project cost manager. Procedia – Soc. Behav. Sci. **119**, 475–484 (2014). https://doi.org/10.1016/j.sbspro.2014.03.053

The Infrastructure and Projects Authority, Government Construction Strategy 2016-20, p. 19, March 2016

Waterhouse, R., Philp, D.: National BIM Report, National BIM Library, pp. 1–28 (2016). https://doi.org/10.1017/cbo9781107415324.004

Zhou, W., Heesom, D., Georgakis, P., Nwagboso, C., Feng, A.: An interactive approach to collaborative 4D construction planning. J. Inf. Technol. Constr. **14**(March), 30–47 (2009)

The Use of Relational and NoSQL Databases in Industrial Asset Management

Jaime Campos[1](✉), Pankaj Sharma[2], Mikel Cañizo[3], Erkki Jantunen[4], David Baglee[5], Santiago Charramendieta[3], and Angel Conde[3]

[1] Linnaeus University, Kalmar, Sweden
jaime.campos@lnu.se
[2] Indian Institute of Technology, New Delhi, New Delhi, India
pankajtq@gmail.com
[3] IK4-Ikerlan Technology Research Centre, Gipuzkoa, Spain
{mcanizo,scharramendieta,aconde}@ikerlan.es
[4] VTT Techical Research Centre of Finland Ltd., Espoo, Finland
erkki.jantunen@vtt.fi
[5] University of Sunderland, Sunderland, UK
David.Baglee@sunderland.ac.uk

Abstract. The advancements concerning the development of ICT systems including Internet of Things (IoT), big data, cloud computing and NoSQL databases provide new opportunities and challenges for industrial asset management. The use of NoSQL databases has emerged due to the limitations of the relational databases, in particular, the inability to scale-up horizontally and to manage the data that is constantly generated by industry. The current work highlights the key aspects of both relational and NoSQL databases. The paper provides a review of the database technologies mentioned above. In this context, in order to demonstrate the effectiveness and adequacy of NoSQL databases, a real industrial case study is presented. The authors also discuss the different database technologies and their suitability in the domain of interest.

1 Introduction

Large volumes of data often referred to as big data and the developments of the Internet of Things (IoT) requires larger storage and computational resources. Consequently, companies today are generating, gathering and storing large amounts of data, the size of which is increasing exponentially. This data is an excellent tool for analysing the problems in the machines. A large number of researchers have worked in the field of using big data for asset maintenance. Trappey et al. (2016) incorporated data warehouse technology to a fault diagnosis system for a fleet of transformers. However, selecting the right database is not always easy, as there is a wide variety of databases ranging from relational to non-relational databases, which have different limitations and strengths. Moreover, the number of databases to use is also a key factor in selecting only one database may not be beneficial as it may not be able to satisfy all requirements. On the other hand, the more databases are used, the more complex the maintenance of the system will be. Thus, there is a need for analysing the existing database

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 302–308, 2020.
https://doi.org/10.1007/978-3-030-48021-9_34

types and how they can be combined in order to maximise the performance of industrial applications, based on their concrete needs. This work presents the characteristics, limitations and strengths of different relational and NoSQL databases and how they can be combined to meet the multiple needs of an industrial solution. The rest of this article is structured as follows: Sect. 2 discusses the features of relational database management systems. Section 3 focuses on NoSQL databases. In Sect. 4, a use case is presented. Section 5 discusses the positive and negative aspects of the different solutions and provides a conclusion.

2 Relational Database Management Systems

Relational Database Management Systems (RDBMS) are databases where data is stored in tables. Each table is a database object composed of rows and columns where the former are instances or records containing the values of the data and the latter define their attributes. RDBMS are based on relational models and thus data stored in one table can be related with any data persisted in another provided that there is a common attribute shared between them (Bordoloi et al. 1994). The most powerful feature of RDMS is the standardised Structured Query Language (SQL) (Chantham 2012), which is used for querying the database to extract information from one or multiple tables. The strength of SQL lies on its ease of use, due to its use of English standard statements. Another key feature of relational databases is that they adhere to ACID, which is a set of properties, namely, Atomicity, Consistency, Isolation and Durability (Mohamed et al. 2014). In the atomicity, everything in a transaction must happen successfully in order to be considered as valid. Otherwise, the database state is left unchanged. The consistency is a result of a transaction the database will only change from a valid state to another. The isolation is a transaction that will not affect to another by changing data that it is using. The durability changes made by a completed transaction must be preserved. RDBMS are also CAP theorem compliant. Eric Brewer (1998) introduced the theorem according to which there is a fundamental trade-off between consistency, availability, and partition tolerance. They are defined as follows Consistency: (do not confuse with the previous term) means that all the clients have the same view of the data. Availability refers to the ability of the database to ensure that a client can always read and write. Partition Tolerance of the system enables the database to work adequately despite network and machine failures (Mohamed et al. 2014). Only two of these requirements must be selected as long as it is not possible to achieve all of them at the same time. RDBMS implement both consistency and availability (CA). In order to ensure consistency of the data, a change in a single record temporarily halts the use of that table such that data modifications can be completed before the data is given out or another user can make another change on it. However, the complex network of references between data items in RDBMS makes it very hard to distribute relational data across several servers and can lead to performance issues both when reading and writing data (Bazar and Iosif 2014). Thus, relational databases suffer from partition tolerance. Availability is achieved by replicating the database and having the replicas in standby mode. If the active one fails, one of the replicas will be activated in order to maintain the proper operation of the database. In addition, relational databases are not well suited for modern web applications that can support millions of concurrent users (Nance et al. 2013). New applications such as text processing, image processing, office information systems, geographic information systems and robotics are difficult to capture in a flat, record-oriented model (Bordoloi et al. 1994). RDBMS has proved beneficial in handling structured data, but there are a large number of drawbacks of RDBMS in the management of large size, number and variety of unstructured data. Storage of large amounts of unstructured data in RDBMS significantly increases the demand for hardware resources. Future growth of data can be addressed through vertical scaling, which further burdens resources and requires complex replication, which is costly (Stevic et al. 2015). Companies added servers with a superior capacity to store and analyse data quicker in an attempt to improve scalability in relational databases. Servers with higher capacity are highly complex, proprietary, and disproportionately expensive (Nance et al. 2013). There were other attempts to scale down the existing relational schemas in order to fit large amounts of data into existing servers. Companies also tried denormalising the schema, relaxing durability and referential integrity, introducing query caching layers, separating read-only from write dedicated replicas, and data partitioning. However, these techniques were unable to address the main issues and ended up adding additional overhead and technical tradeoffs (Nance et al. 2013).

3 The Rise of NoSQL Databases

Thus, the storing and handling large amount of data is one big obstacle of the current relational DBMS (Pereira et al. 2018). Therefore, the NoSQL databases have emerged because of the need to offer solutions with the needed features to tackle the large amounts of data with higher performance. Consequently, NoSQL databases have become popular because of increased scalability and availability requirements of the different solutions (Van der Veen et al. 2012). NoSQL databases deal with challenges associated with web-based applications through their property of scalability and that they are schema-less with the capability to maintain large amounts of data (Barbierato et al. 2014; Stevic et al. 2015). They are distributed by design and intended for efficient management of large amount of unstructured data using horizontal scaling, thus removing obstacles stemming from the future growth of data. Moreover, schema-free NoSQL databases are much better in handling unstructured data (Hellerstein et al. 2007). NoSQL databases contain denormalised data where each record contains a complete set of information without external reference. When a record is moved from one server to another, all the information is transferred complete and with ease. There is no concern that some parts of the record from other tables will be omitted (Bazar and Iosif 2014). Whenever the record is updated, the document itself is modified without the need to change entries in multiple tables. This increases the consistency of the database. There are four main data modelling techniques in NoSQL, i.e. Key-value, Document-oriented, Column-oriented and Graph. The Key-value is a schema that consists of a key and a value. The key is a string that is associated with only one value that can be of any type. Notice that the value is converted to a so-called Binary Large Object (BLOB), a collection of binary data stored as a single entity in a database management system. In this way, the performance of the database is improved since there is no need to index the data. However, a value can only be obtained by its key as long as the value is opaque and thus it cannot be filtered. Common examples of key-value store databases are RIAK (basho.com), DynamoDB (aws.amazon.com/dynamodb/?nc1=h_ls), Redis (redis.io/), InfluxDB (www.influxdata.com) and MemCacheDB (memcachedb.org). Document-oriented databases store data in document-like structures that encode information. Documents can be stored in a binary form (i.e. PDF, DOCX) or an ASCII format (i.e. XML, YAML, JSON, BSON) (Nance et al. 2013). The common uses of document databases include content management and monitoring web and mobile applications. CouchDB (couchdb.apache.org), MongoDB (www.mongodb.com), Elasticsearch (https://www.elastic.co/) are some of the document databases. The column-oriented are also called as table style databases. They store data across tables that can have a large number of columns. Commonly, these are used for internet searches and large web applications. BigTable (cloud.google.com/bigtable/?hl=en), Cassandra (cassandra.apache.org/), HBase (hbase.apache.org) are some of the wide-column databases. The Graph databases are based on a schema-free nature and emphasise connections between data elements. It uses nodes, edges, and properties to represent data. It stores related 'nodes' in graphs to accelerate queries. One of the big advantages to this model is that data can be represented and stored inherently; whereas, in the relational model data has to be transformed from non-tabular to tabular data (Robinson et al. 2015). AllegroGraph (allegrograph.com/) and Neo4j (neo4j.com/) have commonly used graph databases. The very strength of NoSQL databases comes from non-adherence to ACID properties which make them suitable for large-scale applications. Most of NoSQL databases lose on consistency to gain availability and partition tolerance as highlighted by Brewer's CAP theorem that states about the capabilities of the different databases, i.e. distributed database systems have the alternative to select two of the three CAP properties (Indrawan-Santiago 2012). In this way, NoSQL databases adhere to BASE in which an application works all the time (basically available). It does not have to be consistent all the time (soft state), but the storage system guarantees that if no new updates are made to the object, eventually all accesses will return the last updated value (Pokorny 2013). However, not all NoSQL databases choose availability over consistency as some are more focused on achieving more consistency than availability. However, some of them can contain both availability and partition tolerance (AP) or consistency and partition tolerance (CP) depending on how they are configured.

4 Industrial Use Case

In this real manufacturing monitoring use case, hundreds of industrial machines placed around the world are monitored in real-time. A cloud-based solution has been developed to fulfil data capture and storage requirements. Figure 1 shows an overview of the architecture and the data flow of the cloud-based solution. Briefly stated, data generated by the industrial machines is relayed through a distributed messaging system to the cloud where it is processed in real-time. If an anomaly is detected, an alert is sent to the dashboard to create a visualisation. Furthermore, advanced analytics are performed using historical data.

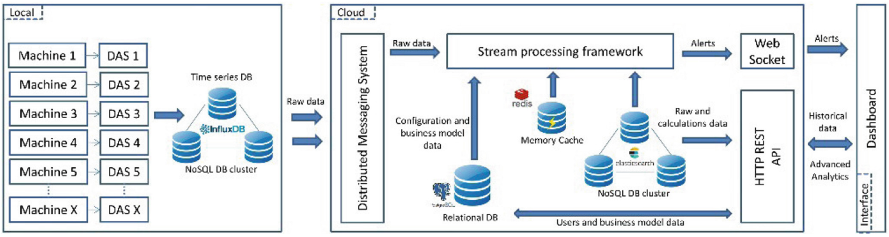


Fig. 1. Cloud-based real-time monitoring architecture

One of the main challenges of this architecture is the management and persistence of the data as they have different natures and needs. Thus, several databases are used to satisfy them. On the one hand, huge data volumes generated by the industrial machines, which are constantly generated as they operate non-stop, must be stored. Therefore, a database with high write throughput, scalable, fault-tolerant and capable of querying stored data for advanced analytics is needed. As the industrial machines work non-stop, the database must also be highly available. To satisfy these requirements, two AP compliant NoSQL database clusters are used: one located in each plant (local side) and the other one in the cloud. As a consequent, it ensures that the databases will always be available even if some nodes fail, adhering in this way to the BASE. In addition, the cluster design provides horizontal scalability by adding more nodes in case more resources are required to store newly gathered data. In this case, consistency (C) is not that critical as the advanced analytics which is not performed in real-time. Thus, the probability of having different views of the data is lower. Referring to the local side, a Direct Attached System (DAS) storage, which can be considered as an additional hard drive for the machine, is connected to each machine to collect the data generated. The data of each DAS is centralised in a key-value database cluster so as it can then be sent to the cloud. For this purpose, InfluxDB is employed as it can handle large data volumes and is specifically designed to manage time series data, making it well suited for the management of collected raw data. With regard to the cloud, a document-oriented database cluster is implemented to store raw data and performed calculations. It is implemented with Elasticsearch, which is a distributed, document-oriented, RESTful search and analytics engine that is capable of persisting data, fulfilling the established requirements. On the other hand, user, business model and configuration data must be stored. This data volume is low and is not expected to grow significantly. However, the management of this data is vital as an error at the time of making a transaction or consulting the database can be critical. Therefore, a relational database is selected as its ACID properties satisfy described requirements, and there is no need to scale it horizontally. To this end, PostgreSQL is selected as it is a powerful engine and therefore suitable for this purpose. Finally, a memory cache is implemented in the cloud to enhance the performance of the real-time processing application as it uses the configuration data stored in the relational database to work. The memory cache is much faster than querying the relational database each time it requires some data. Thus, it only queries the relational database once to collect all the required data and afterwards, it caches that data to query it faster the following times.

5 Discussion and Conclusions

Within the industrial domain, there are multiple machinery and industrial systems which require complicated monitoring systems. Each system has its characteristics and needs. Therefore the data involved in the process is different. As shown in the previous sections, both RDBMS and NoSQL databases have their strengths and weaknesses. It is not possible to use only one type of database to achieve all the different needs. So, why use only one instead of using a combination of database technologies? For instance, asset management systems and the organisations should consider migrating to a storage architecture that takes advantage of the strengths of both NoSQL and relational databases. This hybrid use of the databases is called as "*polyglot persistence*" to express the idea that applications should be written using a mix of languages and database technologies to take advantage of the fact that different databases are suitable for tackling different problems (Schaarschmidt et al. 2015; Sadalage and Fowler 2012). Similarly, they have different types of compliance with ACID properties and the CAP theorem. Sensitive data must be handled through encryptions, and therefore RDBMS must be used. In this way, complex queries or multi-table transactions can be performed while preserving their atomicity, consistency, isolation, and durability. This ACID assurance gives companies the ability to use both SQL and NoSQL interfaces without worrying about many of the security fears associated with the use of NoSQL. It is vital for any company to consider a number of variables, in particular economic considerations, which must be realised financially as well as to the performance of the individual database. Numerous use cases where a high database performance is required may not perform well with SQL databases and can cost more regarding lost revenue. However, giving up on RDBMS and setting up NoSQL databases is also costly and may dissuade some companies from continuing with safer but slower RDBMS. In an asset management application, various types of RDBMS and NoSQL databases have different applicability. In addition, wide-column NoSQL databases are suitable for data mining and analytics. These databases can be used for fault diagnosis and prognosis applications. Additionally, because of the aspects of the document NoSQL databases are useful for content management and storing information about various stakeholders of the asset management system. The Graph databases have their application in analysing 'Users' Network' as also for 'Recommendations' to the users regarding action to be taken to tackle a machine fault. The inventory and costs of the individual items are suitable to be stored in RDBMS. In conclusion, it is crucial to understand the characteristics of the different ICTs, in this case, the DBMS and the NoSQL databases to be able to understand how they fit into different processes and activities of the domain of interest. The storage architecture using polyglot persistence or similar solutions are convenient, i.e. the use of a hybrid system for different kind of data that is generated in asset management. Consequently, it is crucial to have the state of the art knowledge from both academia and industry to comprehend the strength and weakness of the different kind of database systems that exist or are emerging to be able to decide the most suitable solution to implement taking into consideration the aspects of the domain.

References

Barbierato, E., Gribaudo, M., Iacono, M.: Performance evaluation of NoSQL big-data applications using multi-formalism models. Future Gener. Comput. Syst. **37**, 345–353 (2014)

Bazar, C., Iosif, C.B.: The transition from RDBMS to NoSQL: a comparative analysis of three popular non-relational solutions: cassandra, mongoDB and couchbase. Database Syst. J. **V**(2), 49–59 (2014)

Bordoloi, B., Agarwal, A., Sircar, S.: relational or object- oriented or hybrid?: a framework for selecting an appropriate database management system type in a computer integrated manufacturing setting. Int. J. Oper. Prod. Manag. **14**(9), 32–44 (1994)

Chatham, M.: Structured Query Language By Example - Volume I: Data Query Language. p. 8 (2012). ISBN 978-1-29119951-2

Hellerstein, J.M., Stonebraker, M., Hamilton, J.R.: Architecture of a database system. Found. Trends Databases **1**(2), 141–259 (2007)

Indrawan-Santiago, M.: Database research: are we at a crossroad? Reflection on NoSQL. In: 2012 15th International Conference on Network-Based Information Systems (NBiS), pp. 45–51. IEEE (2012)

Mohamed, M.A., Altrafi, O.G., Ismail, M.O.: Relational vs NoSQL databases: a survey. Int. J. Comput. Inf. Technol. **03**(03), 598–601 (2014)

Nance, C., Losser, T.: Iype, R., Harmon, G.: NoSQL vs RDBMS - why there is room for both. In: SAIS 2013 Proceedings, pp. 111–116 (2013)

Pereira, D.A., Ourique de Morais, W., Pignaton de Freitas, E.: NoSQL real-time database performance comparison. Int. J. Parallel Emergent Distrib. Syst. **33**(2), 144–156 (2018)

Pokorny, J.: NoSQL databases: a step to database scalability in web environment. Int. J. Web Inf. Syst. **9**(1), 69–82 (2013)

Robinson, I., Webber, J., Eifrem, E.: Graph Databases: New Opportunities for Connected Data. O'Reilly Media Inc, Sebastopol (2015)

Sadalage, P.J., Fowler, M.: NoSQL Distilled - A Brief Guide to the Emerging World of Polyglot Persistence. 1st edn., Addison-Wesley Professional (2012)

Schaarschmidt, M., Gessert, F., Ritter, N.: Towards automated polyglot persistence. Datenbanksysteme für Business, Technologie und Web (BTW 2015) (2015)

Stevic, M.P., Milosavljevic, B., Perisic, B.R.: Enhancing the management of unstructured data in e-learning systems using MongoDB. Program **49**(1), 91–114 (2015)

Trappey, A.J.C., Trappey, C.V., Ma, L.: Incorporating data warehouse technology into asset information management systems for larger assets. In: Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015), pp. 601–612. Springer, Cham (2016)

Van der Veen, J.S., Van Der Waaij, B., Meijer, R.J.: Sensor data storage performance: SQL or NoSQL, physical or virtual. In: 2012 IEEE 5th international conference on Cloud computing (CLOUD), pp. 431–438. IEEE (2012)

Leveraging Asset Management Data for Energy Recovery and Leakage Reduction

Gideon Johannes Bonthuys[1](✉), Petrus Blom[1], and Marco van Dijk[2]

[1] i@Consulting, 472 Botterklapper St, Die Wilgers, Pretoria 0184, South Africa
{deon, petrus}@iatconsulting.co.za

[2] Department of Civil Engineering, University of Pretoria, Pretoria 0002, South Africa
marco.vandijk@up.ac.za

Abstract. Excess pressure and increases in pressure within water distribution systems correlate directly to increased water losses from pipe leakages. Pressure management in water distribution systems, is one of the most influential, most important and most cost-effective interventions that can be implemented in order to reduce leakage. Excess pressure available in water distribution systems can be used as a renewable, low cost clean energy alternative for energy production with no significant environmental impacts. Similar to the use of pressure reducing valves, hydraulic energy recovery devices can be used to reduce excess pressure in a water network, but with the added benefit of converting the excess pressure into electric power rather than to dissipate the energy. Asset management systems, inclusive of asset management plans and asset registers, provides unabridged information on water distribution systems with regards to spatial, technical, operational and financial data. Analysis of this data with respect to energy recovery highlights the hidden opportunities. The energy consumption in urban water supply and distribution systems represents 7% of the world's energy consumption. The paper investigates and proposes the leveraging of asset management data for optimizing pressure management to enable maximum energy recovery and reduce the leakage rate from the system. Energy recovery decreases the carbon foot-print of the water distribution system while simultaneously either generating a stream of revenue for the operator or resulting in an energy cost saving. This increase in revenue or cost saving impacts on the dynamic of an asset management plan and provides for a more sustainable system.

1 Introduction

The four widely excepted fundamental management practices for leakage reduction or the constraint of physical losses from water distribution systems are pressure management, speed and quality of repair of existing leaks, active leakage control, and asset management (Samir et al. 2017). This paper demonstrates the potential of leveraging the asset management data to identify areas in which to implement pressure management through conduit hydropower installations. The conduit hydropower installation not only serves as a pressure management practice but recovers energy from the water distribution systems.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 309–317, 2020.
https://doi.org/10.1007/978-3-030-48021-9_35

The South African Water Services Act requires every municipality to have an asset management plan. Both the Department of Water and Sanitation and the Institute of Municipal Engineering of Southern Africa have developed practices to assist municipalities to compile and comply with the asset management requirements of the Water Services Act (Van Zyl 2014). Similarities in the steps of asset management plan development as proposed by the Department of Water and Sanitation and the Institute of Municipal Engineering of Southern Africa is described in Table 1.

Table 1. Development of an asset management plan as proposed by the Department of Water and Sanitation and the Institute of Municipal Engineering of Southern Africa (Van Zyl 2014).

Phase	DWS	IMESA
Technical assessment	1. Asset register	1. Asset register
	2. Condition Assessment	2. Condition Assessment
		3. Remaining Useful Life
Financial assessment	3. Current & future needs	4. Levels of service & demand
	4. Costing analysis	5. Valuation & Life cycle cost
		6. Business risk exposure
Asset management practices	5. Operational plan	7. Operation & Maintenance plans
		8. Capital investment validation
	6. Maintenance plan	9. Future expenditure model & Funding
		10. Asset Management Plan to suit budget

In 2014 the ISO 5500x series of standards were published and ushered in a new era for asset management. Soon after the South African Bureau of Standards followed suit and published SANS 5500x (Boshoff and Childs 2016). ISO 5500x defines asset management as the coordinated activity to deliver value from assets. Using energy recovery from water distribution systems for pressure management delivers value from assets through direct energy cost savings and economic and socio-economic benefits resulting from leakage reduction. Energy recovery should therefore form part of asset management plans where viable potential within the water distribution system exists.

Since the late 1970s it has been widely recognised that pressure has a fundamental influence on average leakage rates within water distribution systems and that leakages can be reduced by limiting operating pressures. Increasing energy costs has guided pressure management controls in water networks from using pressure reducing valves towards incorporating hydro turbines to utilize pressure dissipation as a mechanism for energy recovery from the system. Tricarico et al. (2014) highlights the potential revenue or cost saving from energy recovery while simultaneously reducing leakages within the water distribution system.

Each step in the asset management plan development process from Table 1 provides data that can be leveraged to obtain a model for energy recovery within the water distribution system. The method in which this data is processed to identify potential pressure management zones suitable for energy recovery; the potential leakage reduction from pressure management through energy recovery in these zones; and how energy recovery improves the reliability of the water distribution system, is described below.

2 Asset Management Data

The concept of leveraging asset management data centres on the utilization of existing, measured and documented data to add energy recovery and leakage reduction potential to the asset management plan value chain. The flow of asset management data in the energy recovery and leakage reduction process is shown in Fig. 1.

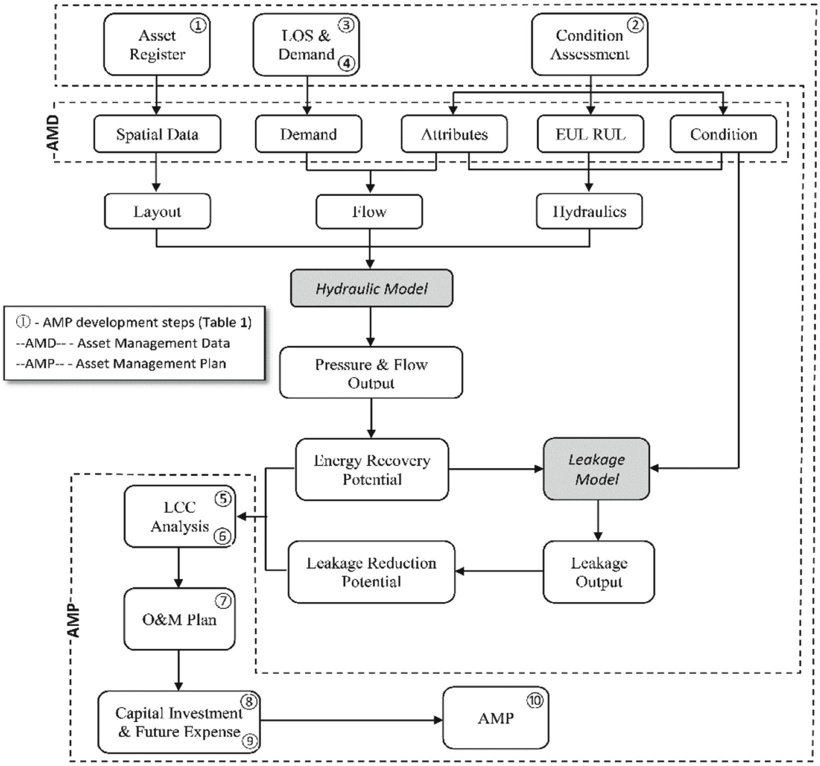


Fig. 1. The flow of asset management data in the energy recovery and leakage reduction process.

Spatial data of linear and point components within the asset register forms the layout of the hydraulic model for the specific section under investigation. Data from the asset management plan with regards to customer profiles and levels of service (LOS) populates the flow requirements within the hydraulic model. The variables that influences the hydraulic behaviour of the system are derived from data extrapolated from the condition assessment of the components within the asset register as outlined in the asset management plan. Asset management data relating to a specific section is thus consolidated into a hydraulic model, which can calculate potential excess pressure within a system (or section thereof) and highlight an opportunity for energy recovery as explained in Sect. 4.

The condition assessment of the section under investigation in terms of unaccounted for water is used to calibrate a leakage model. Additional information on system hydraulics such as pressure and flow can be obtained through installed telemetry used in preventative maintenance regimes described within the asset management plan. Additional information from telemetry increases the confidence rating of the leakage model. Pressure and flow output from the hydraulic model relating to energy recovery potential is used in conjunction with the calibrated leakage model to calculate the potential for leakage reduction from energy recovery, as discussed in Sect. 3.

3 Leakage Reduction

The International Water Association and the American Water Works Association recommend the Infrastructure Leakage Index (ILI) as the most detailed performance indicator for nonrevenue water and real operational losses. The Infrastructure Leakage Index is the dimensionless ratio between the current annual real losses (CARL) and the unavoidable annual real losses (UARL) within a system. The unavoidable annual real losses are the lowest technically achievable annual real losses and can be calculated using Eq. 1 (Samir et al. 2017).

$$UARL = (18\,Lm + 0.8\,Nc + 25\,Lp) \times H \tag{1}$$

Where Lm = mains length (km); Nc = number of service connections; Lp = the total length of underground pipe between the edge of the street and customer meters (km); H = average operating pressure (m).

Liemberger (2010) developed a physical losses assessment to classify leakage levels in both developed and developing countries into 5 categories as shown in Table 2. Pressure has a fundamental influence on average leakage rates within water distribution systems as leakage is, to an extent, driven by pressure. Leakage volume and new leakage frequency is reduced through the reduction and stabilization of pressure within a water distribution system. Equation 2, known as the FAVAD equation, is the preferred relationship between leakage and pressure (Kabaasha et al. 2016).

$$Q = cd\,\sqrt{2}(A_0 h^{0.5} + mh^{1.5} \tag{2}$$

Where Q = leakage rate (m^3/s); Cd = discharge coefficient; g = gravitational acceleration (m/s^2); A_0 = initial leak opening without any pressure in the pipe (m^2); h = pressure head (m); m = slope of the pressure area line (m).

Table 2. Physical Loss Assessment (Liemberger 2010) (Adapted)

Category	ILI	Description
A1	<2	Potential for further NRW reductions is small
A2	2–4	Further NRW reduction may be uneconomic unless there are water shortages or very high water tariffs
B	4–8	Potential for marked improvements
C	8–16	Poor NRW record
D	>16	Highly inefficient; a comprehensive NRW reduction program is imperative and high-priority

The FAVAD equation is used to estimate the leakage from a water distribution system for any given pressure within the system. The total leakage reduction in a water distribution system is the difference between the current leakage (base scenario) and the estimated leakage after a change in system pressure. The change in pressure for the purposes of this paper refers to the pressure head associated with the energy recovery from the water distribution system.

4 Energy Recovery Potential

In both a pumped and a gravity water distribution system, pressure in the system must be lower than the prescribed limit to prevent water losses through leakages and pipe bursts (Gaius-obaseki 2010). In addition, too little pressure will cause unsatisfactory levels of service to consumers and therefor water distribution systems have limitations for pressure. Leakage volume is reduced though a reduction in pressure within a water distribution system. The maximum leakage reduction potential therefore exists when the pressure within a water distribution system is reduced to the minimum acceptable operating conditions. Any pressure within a water distribution system over and above the minimum acceptable operating pressure is potentially excess pressure which could be converted into energy through energy recovery. The residual pressures within South African water distribution systems should be within the limits shown in Table 3 or specific local authority requirements.

Table 3. Residual pressure in South African water distribution systems (CSIR 2005)

Types of development: dwelling houses	Minimum head under instantaneous peak demand (m)	Maximum head under zero flow conditions (m)
House connections	24	90
Yard taps + yard tanks	10	90

The typical approaches to pressure control within water distribution systems are to install pressure reducing valves. Energy dissipated in water distribution systems can be recovered by replacing pressure reducing valves with hydro turbines. In South Africa there are 257 municipalities which all own and operate water distribution systems which can be equipped with hydro turbines for energy recovery and supplement or reduce the requirements for pressure reducing valves (Van Dijk et al. 2012). Whenever excess energy is present within a water distribution system, hydro turbines can be installed for energy recovery.

To calculate the potential energy recovery from a water distribution system, the available excess pressure head, along with the associated flow is used. The flow and head available is used in Eq. 3 to calculate the energy recovery capacity in certain scenarios. Total potential energy recovery is calculated as the product of the energy recovery capacity and the period over which it occurs.

$$P = \rho g Q H \mu \tag{3}$$

Where P = power output (watt), ρ = density of fluid (kg/m^3), g = gravitational acceleration (m/s^2), Q = flow rate (m^3/s), H = available pressure head (m) and μ = turbine system efficiency.

5 Case Study – Ext 47, Ekurhuleni, South Africa

The concept of energy recovery and leakage reduction through leveraging asset management data was tested through a desktop study conducted on a section within the Ekurhuleni Metropolitan Municipality in South Africa, called Extension 47. From the asset register of Ekurhuleni and Ext 47 data was extracted to construct the layout of the hydraulic model. A combination of data from the condition assessment of assets pertaining to the water supply at Ext 47 were used in calculating the hydraulic characteristics of the model.

Ext 47 is a residential suburb consisting of a development of 107 retirement units. The type of development is classified by the Guidelines for Human Settlement Planning and Design (CSIR 2005), as dwelling houses (residential zone 1), and has a theoretical annual average daily water demand of between 600 and 1200 l/day based on an erf size of less than 400 m^2. The actual demand of Ext 47 was measured during the month of June 2107 and extrapolated to estimate an annual average daily water demand of 518 ℓ/day and correlates closely with the standard guidelines for residential zone 1. Figure 2 shows an extract of the flow measurements recorded at the supply inlet pipe to Ext 47 during the month of June 2017. These measurements along with billing information and estimates within the asset management plan were used to set up the flow conditions within the hydraulic model.

Figure 2 also indicates a minimum night flow of approximately 0.76 m3/h which is indicative of possible water leakage. This indicative leakage flow along with the non-revenue water calculations within the asset management data of the asset management plan, were used to calculate the International Leakage Index of Ext 47 as well as to calibrate the leakage model. Figure 3 shows the graphical representation of the hydraulic model for Ext 47 set up in EPANET, using asset management data.

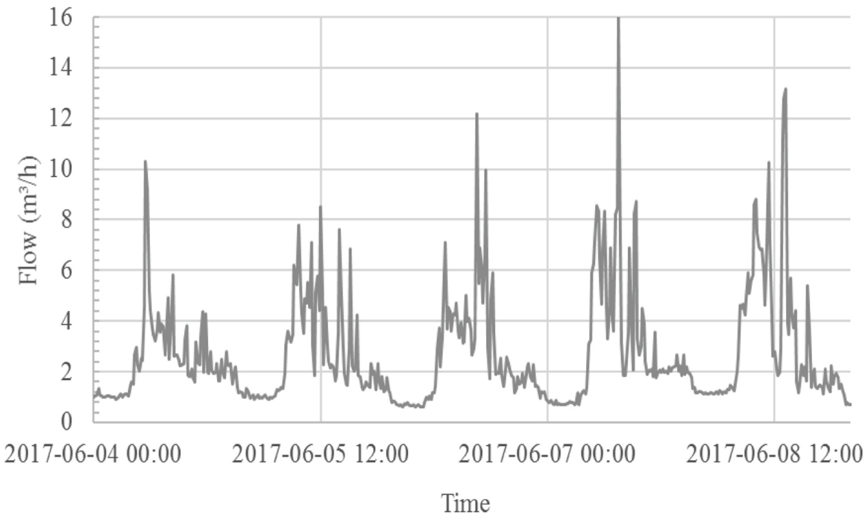


Fig. 2. Ext 47 - Flow measurements - June 2017

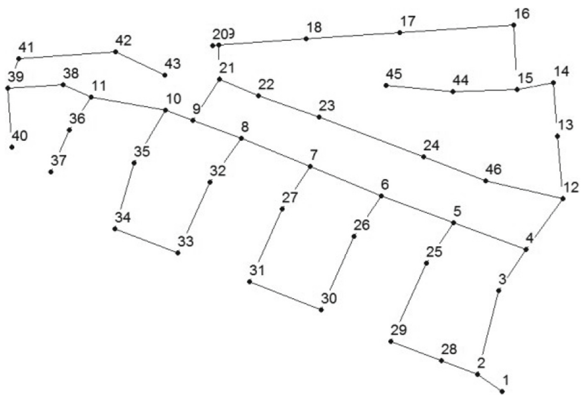


Fig. 3. Ext 47 - Hydraulic Model – EPANET

6 Results

The unavoidable annual real losees and current annual real losses for Ext 47 were calculated as 79 ℓ/day/connection and 122 ℓ/day/connection respectively, resulting in an International Leakage Index of 1.5. According to Table 2 this places Ext 47 in category A1 and highlights that potential for reduction in non-revenue water is small.

From the hydraulic model it was evident that the minimum pressure in the system exists at node 43. Reduction in system pressure is therefore limited by node 43. Due to the small extent of Ext 47 the optimum scenario for energy recovery was modelled using a singular location at the main inlet pipe to Ext 47 at node 1. In larger sections a combination of several locations may have the largest impact. The available flow for energy recovery was calculated as the total flow at the main inlet pipe to Ext 47 minus the accumulated leakage rate of 0.76 m3/h. Figure 4 shows the available pressure head and flow for energy recovery for a 24 h period as per the average demand pattern calculated from the June 2017 flow data. This graph is an output from the hydraulic model as was used to calculate the total daily energy recovery potential as 5.2 kWh.

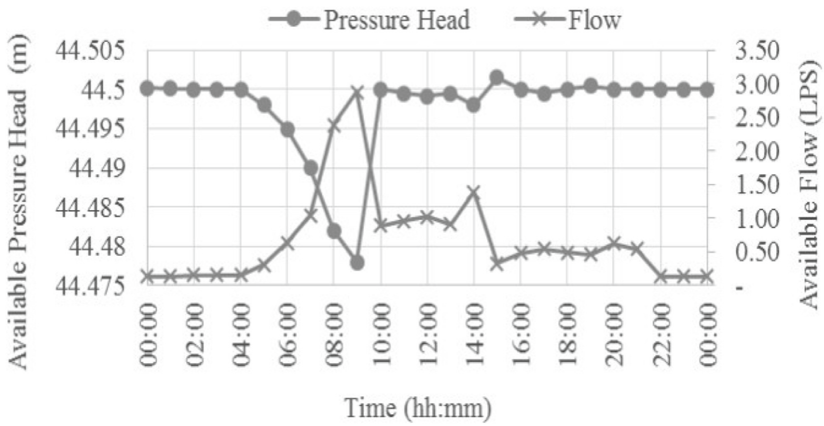


Fig. 4. Total available pressure head and flow for energy recovery from Ext 47

The leakage model was calibrated using the current annual real losses and the flow and pressure results from the hydraulic model. The leakage model was run using the pressure both before and after potential energy recovery. Results from the model showed an average reduction in leakage from 122 ℓ/day/connection to 31 ℓ/day/connection. This constitutes a 74% reduction in unaccounted for water within Ext 47 when energy recovery is used for leakage reduction. The energy recovery and leakage reduction for Ext 47 as per the study amounts to an average annual saving of R61 650, based on electricity and water unit tariffs.

7 Conclusions

The following conclusion and recommendations emanated from the study:

- Asset management data can be leveraged to highlight opportunities for energy recovery and leakage reduction in municipal environments.
- By combining leakage reduction and energy recovery, a more sustainable system can be obtained, recovering 5 kWh of energy from the water distribution system in Ext 47 and a 74% reduction in unaccounted for water.
- The energy recovery and leakage reduction are not one directly a product of the other but linked through the dissipation of excess energy in the system.

- It is recommended that larger areas be included in the study to investigate the impact of higher demands on energy recovery and leakage reduction.
- It is recommended that the study be developed into an add-in tool to identify and spatially report energy recovery and leakage reduction through existing asset management software.

References

Boshoff, L., Childs, R.: Urban infrastructure asset management: the cities infrastructure delivery and management system toolkit (CIDMS). SAAMA, Somerset West, Cape Town 2016

CSIR: Guidelines for Human Settlement Panning and Design, vol. 2. Capture Press, Pretoria (2005)

Gaius-obaseki, T.: Hydropower opportunities in the water industry. Int. J. Environ. Sci. **1**(3), 392 (2010)

Kabaasha, A.M., Van Zyl, J.E., Piller, O.: Modelling pressure: leakage response in water distribution systems considering leak area variation. Computing and Control in Water Industry, Amsterdam (2016)

Liemberger, R.: Recommendations for initial non-revenue water assessment. International Water Association, Sau Paolo (2010)

Samir, N., Kansoh, R., Elbarki, W., Fleifle, A.: Pressure control for minimizing leakage in water distribution systems. Alexandria Eng. J. **56**(4), 601–612 (2017)

Tricarico, C.: Optimal water supply system management by leakage reduction and energy recovery. Procedia Eng. Osa/vuosikerta **89**, 573–580 (2014)

Van Dijk, M., Van Vuuren, S.J., Bhagwan, J.N.: Conduit Hydropower Potential in a City's Water Distribution System, IMESA, George (2012)

Van Zyl, J.E.: Introduction to operation and maintenance of water distribution systems. Water Research Commission, Pretoria (2014). 1 toim

Part IV: Asset Decisions and Decision Support

Application of a Value-Based Decision-Making Process to an Industrial Water Supply System

Manuela Trindade[1], Nuno Almeida[2(✉)], Matthias Finger[3], Jaime Gabriel Silva[4], Luís Ghira[5], and João Vieira[6]

[1] Infraestruturas de Portugal, CERIS, Lisbon, Portugal
manuela.trindade@tecnico.ulisboa.pt
[2] IST - University of Lisbon, CERIS, Lisbon, Portugal
nunomarquesalmeida@tecnico.ulisboa.pt
[3] EPFL – École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
matthias.finger@epfl.ch
[4] Instituto Superior de Engenharia do Porto, Águas de Santo André, Santiago do Cacém, Portugal
g.silva@adp.pt
[5] Águas de Santo André, S.A., Santiago do Cacém, Portugal
l.ghira@adp.pt
[6] IST – University of Lisbon, Lisbon, Portugal
joaopcvieira@outlook.com

Abstract. The quality requirements of industrial water are often not as strict as those of water for human consumption, but failure to fulfil the needs of industrial processes may result in extremely high economic impacts. Águas de Santo André (AdSA) is a Portuguese utility responsible for the industrial water supply system of the Industrial and Logistics Zone of Sines (IWSS-ILZS). This paper uses the IWSS-ILZS as a case study to test the application of a proposed value-based decision-making process for asset intensive organizations. The proposed process highlights the benefits of competing alternatives for renewing the infrastructure. It contributes to higher engagement in value-based thinking over the lifecycle of the asset base of AdSA, enabling the balancing of performance, cost and risk, and an increased alignment with corporate objectives and stakeholder needs and expectations.

1 Introduction

This paper addresses the complexities of decision-making in an industrial water supply system of a Portuguese utility. This industrial water infrastructure is used as a case study to test a prototype value-based decision-making process designed by the authors (Trindade and Almeida 2016; Trindade et al. 2017). The process includes the following steps: (i) communications and consultation; (ii) establishing the context; (iii) opportunity assessment (opportunity identification, opportunity analysis and opportunity evaluation); (iv) opportunity implementation; (v) recording and reporting; and, (vi) monitoring and review.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 321–329, 2020.
https://doi.org/10.1007/978-3-030-48021-9_36

The systematic application of the proposed process is expected to promote proactive lifecycle thinking and enable the consideration of short and long-term impacts towards the delivery of the best value for the organization and all stakeholders.

2 Methodology

The methodology adopted in the case of study is as shown in Fig. 1: (i) review of literature and standards applicable to the water sector; (ii) analysis of documented information on AdSA and the IWSS; (iii) application of the proposed decision-making process to upgrade/renewal decisions of a critical gravity pipe installed in 1980; (iv) multidisciplinary discussion and critical review of results; and v) validation of results prior to operationalization of the decision.

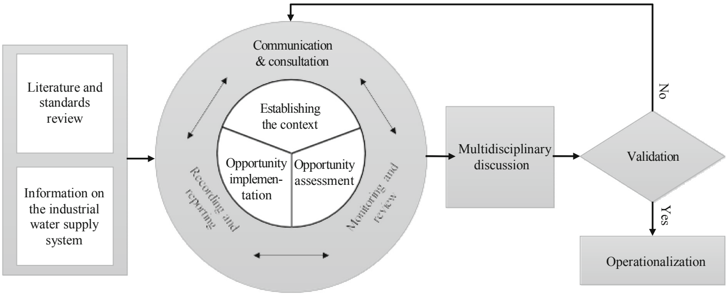


Fig. 1. Methodology of the case study

3 Case Study

Águas de Santo André (AdSA) is a water utility in Portugal whose mission is to manage and operate water and wastewater systems, including an industrial water supply system comprising catchment, adduction, treatment, storage and distribution processes (AdSA 2015). The needs of the industrial clients of the ILZS determine the service levels of the system. AdSA has to deal with the complexities of the ILZS and uncertainties about the future situation of this industrial and logistical zone.

Aiming at reducing infrastructure risk, AdSA is planning the upgrade/renewal of one of its critical assets, a gravity pipe, which is the focus of the case study discussed in this paper. The case study includes a sensitivity analysis for several demand scenarios given the uncertainty of growth expectations for ILZS in the short/medium term, taking also into consideration the effects of demand in the asset deterioration rate, amongst other parameters affecting the decision-making process.

The period of analysis of the case study is 20 years (2019–2038). The analysis covers both short-term (5 years) and long-term (15 years) effects.

4 Application of the Proposed Process

4.1 Communication and Consultation

Successful management of opportunities involves sound decisions and informed choices. This depends on effective communication and consultation with relevant stakeholders.

In this case study, the authors used engagement sessions and workshops with an *ad hoc* asset management group of AdSA, involving the CEO and the head and technicians of Engineering, Maintenance, Operations and Finance. This enabled the gathering of the documented information needed to assess the competing options or investment opportunities, plus the multidisciplinary discussion leading to validation before the operationalization of the decisions made.

4.2 Establishing the Context

Scope of decisions

The context of the decision-making is a renewal or upgrade of a 10 km long gravity pipe with diameters ranging from 1000 to 1500 mm. This pipe connects the water treatment plant at an elevation of 70 m to a 50 000 m^3 storage tank at 53 m of elevation that precedes the distribution network. The pipe material is reinforced concrete in general and steel in the valve chambers. Inspections reported a poor structural condition of these valve chambers and leaks of treated water.

The gravity pipe, in service since its construction in 1980, presently delivers about 13 000 000 m^3/year, with an expected increase in demand up to 20 000 000 m^3/year in the short term. The installed capacity of this pipe is approximately 42 000 000 m^3/year (AdSA 2014).

This gravity pipe is one of the most important and critical assets. There is no redundancy for this part of the supply system and its failure implies an unacceptable interruption in the supply to the industrial clients. There is thus a need for an upgrade or renewal.

Value drivers and impacts

AdSA's mission states three high-level concerns: economic, social and environmental efficiency and sustainability. Its vision is to be a national reference in the environmental sector, in terms of quality of the public service provided and an active partner for the economic development of the region. These are the guidelines or inputs for establishing the strategic objectives of the organization, alongside with the customers' expectations and needs and the legal and regulatory requirements (LNEC & ERSAR 2017; Alegre et al. 2014).

Value drivers are essential factors for making asset-management decisions as it focuses attention on the key elements that need to be considered in the decision-making process, such as the asset and its functionality as well as the needs and expectations of different stakeholders (Haddadi et al. 2016). Value drivers help asset managers to focus on those activities that will have the greatest impact on value for the organizations and its stakeholders and broadly encompass factors related to cost, risk and performance

(Parlikad and Srinivasan 2016). For example, in this case study value-drivers include the impacts on service (quality and capacity), costs (opex and capex) and reputational risks (e.g. safety and environment).

In the proposed value-based decision-making process the competing renewal and upgrade options of the gravity pipe are assessed against: i) impacts in the organization (direct and indirect business impacts of the project on the strategic objectives of AdSA), and; ii) impacts on stakeholders (impacts of the project on those that are directly or indirectly affected, taking into account social concerns beyond financial aspects).

The demand scenario considered is 20 hm^3/year, 7 hm^3/year higher than present consumption levels. A sensitivity analysis is carried out to evaluate the impacts of changes in demand.

4.3 Opportunity Assessment

Opportunity identification
Four options are considered: Option 0 - Status quo (42 hm^3/year); Option 1 - Rehabilitation of the existing pipe (42 hm^3/year); Option 2 - Construction of a new pipe with higher capacity (50 hm^3/year) and disposal of existing pipe; Option 3: Construction of a new pipe (20 hm^3/year) and rehabilitation of the existing pipe (42 hm^3/year).

Opportunity analysis
The analysis of the competing options ensures that the selected capital project delivers value to the stakeholders and meets the objectives set for the project and the organization (GFMAM 2016).

The analysis is based on nine criteria with corresponding value metrics, four of which pertain to the performance dimension (Fig. 2), two to the cost dimension (Fig. 3) and three to the risk dimension (Fig. 4). All risks are monetized and included as financial costs.

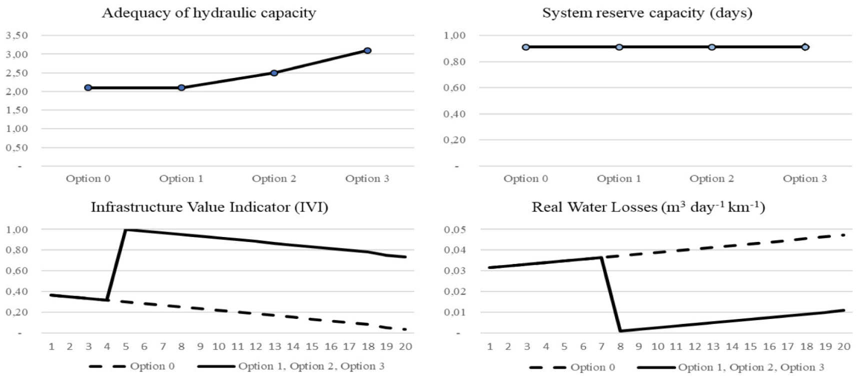


Fig. 2. Performance analysis

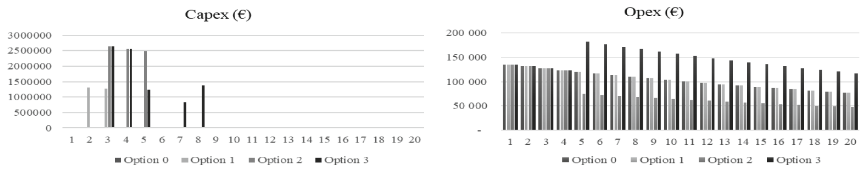


Fig. 3. Cost analysis

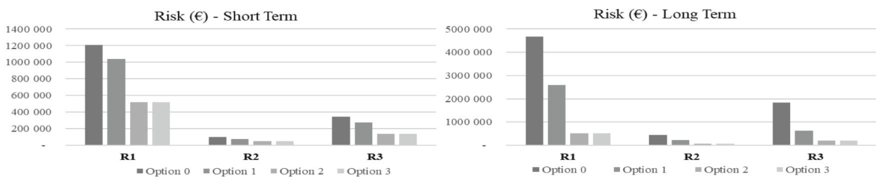


Fig. 4. Risk analysis

Opportunity evaluation

Evaluating implies a judgement on the value deriving from the competing opportunities and choosing those that are to be implemented (Trindade et al. 2017). This involves the estimation of the value of each individual opportunity or group of opportunities.

The value delivered to the organization and the various stakeholders can be estimated by evaluating the impacts (organization and stakeholders) of each competing option.

In this case study, future cash-flows are considered to worth less than current cash flows (Parlikad and Srinivasan 2016). A reference discount rate of 3% was considered following the Guide to Cost-Benefit Analysis of Investment Projects (European Commission 2014).

The following techniques were used to quantify organization impacts (Fig. 5): Net Present Value Analysis (Eq. 1), Return On Investment (Eq. 2) and Benefit Cost Analysis (Eq. 3).

In the equations, the financial costs include OpEx, CapEx and monetized risks, r is the discount rate, i is the year of the period of analysis, and N is the total period of analysis (20 years).

$$NPV\ (€) = \sum_{i=1}^{N} \frac{(Business\ revenues - Financial\ costs)}{(1+r)^i} \tag{1}$$

$$ROI\ (\%) = \frac{\sum_{1=1}^{N}(Business\ revenues - Financial\ costs)}{\sum_{1=1}^{N}(Financial\ costs)} \tag{2}$$

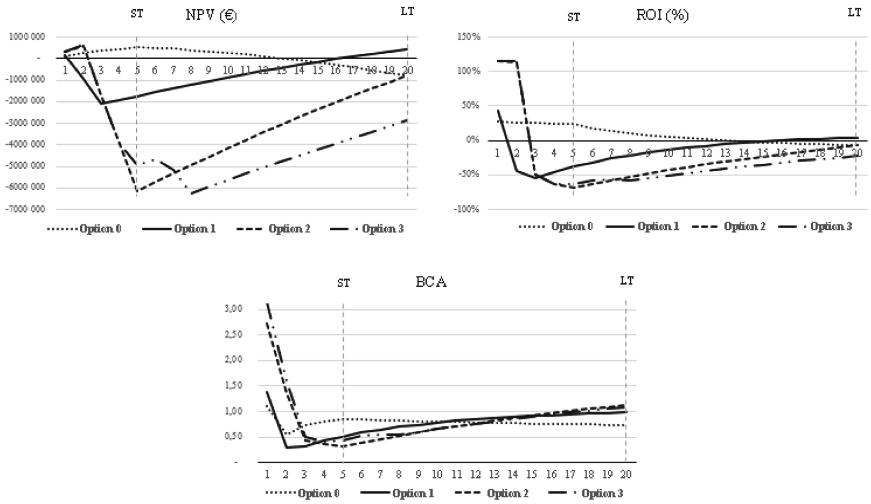


Fig. 5. Cumulative results for NPV, ROI and BCA methods

$$BCA = \frac{\sum_{i=1}^{N} (Business\ revenues + Social\ benefits)}{\sum_{i=1}^{N} (Financial\ costs + Social\ costs)} \tag{3}$$

Considering the whole-life of the infrastructure and for an average demand of 20 hm^3/year, the NPV indicates that option 1, followed by option 2, delivers the best value for the organization in the long term.

The ROI for the entire period also indicates best value for option 1 followed by the option 2.

When the positive (benefits) and negative (costs) impacts of various options are considered and evaluated according the BCA technique, the best option is number 2 followed by option 3 (BCA > 1), options number 0 and 1 having a BCA < 1.

Value is generated by a water utility primarily through meeting the needs of its stakeholders. These needs will vary between stakeholders (Hughes and Moore 2017). A preliminary evaluation of the impacts on stakeholders was performed using a qualitative method of the perceived value of each competing option by the different stakeholders. A further detailed evaluation can be done using a multi-criteria analysis (MCA), namely by identifying and evaluating the positive and negative impacts of the different options in non-monetary terms (IIMM 2015). When evaluated the impacts on stakeholders considering the respective value drives defined, the best option was number 2, followed by the option 3. When stakeholders were considered option 1 was found unacceptable.

4.4 Opportunity Implementation

Opportunity implementation involves a set of feasible projects that could transform some or all of the relevant opportunities into realized benefits (Trindade et al. 2017). 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 and the needs of stakeholders.

4.5 Recording and Reporting

The proposed value-based opportunity management process assists decision makers with documented information about the choices they make (e.g. prioritise actions and select among alternative options).

In an asset management context, it is likewise relevant to succinctly demonstrate that the best value is being delivered from the asset base. And that an adequate balancing of performance, cost and risk was achieved, both in the short and the long terms.

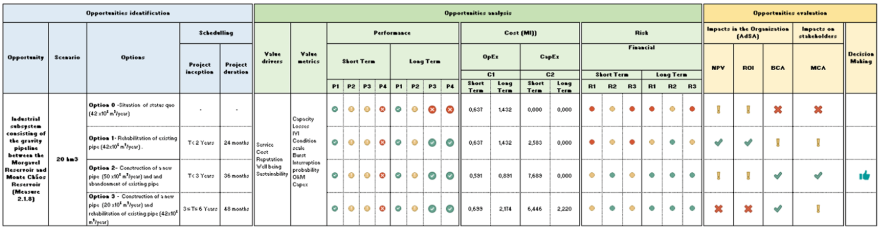


Fig. 6. Extract of a value record of the industrial water supply system

Different techniques (e.g. NPV, ROI, BCA, MCA) may be applicable to express the concept of value depending on the value drivers assumed by the organization. These assumptions should be clearly recorded and reported as part of the decision-making process. For example, a value record such as the one presented in Fig. 6 can be used in this regard. It summarizes the information generated throughout the decision-making process, enabling a quick perception on how the best possible result can be achieved.

4.6 Monitoring and Review

Business environments are constantly changing. Factors affecting the organization and its stakeholders are likely to change as well. The balancing of performance, costs and risks in asset management decisions is therefore an iterative process. The efficiency and effectiveness of the decision making proposed process thus depends on monitoring and revision activities, so that the opportunity implementation plan remains relevant and updated.

Because of the uncertainty in the estimated for demand of industrial water, a sensitivity analysis for five different scenarios was undertaken (13 hm^3/year, 14 hm^3/year, 20 hm^3/year, 25 hm^3/year and 50 hm^3/year). Apart from the considerations related with stakeholders (e.g. when the BCA technique is used), option 1 seems to be the best solution when demand is under 25 hm^3/year and for every interest rate variation (Fig. 7).

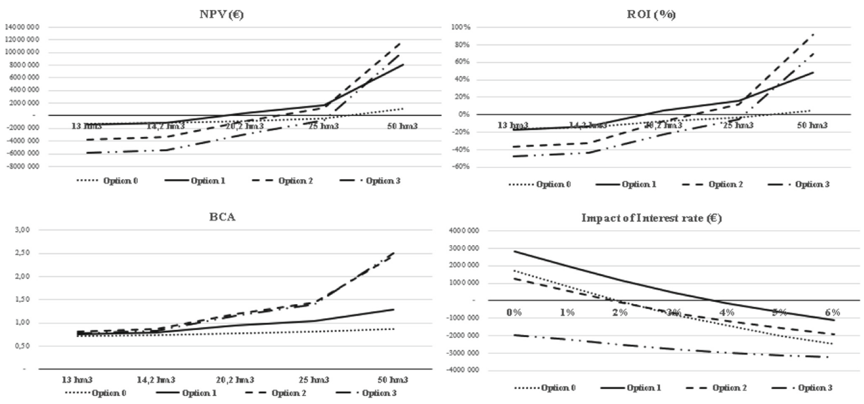


Fig. 7. Sensitivity analysis for water demand and interest rates

5 Discussion

A combination of techniques was used to calculate the value delivered by the industrial water supply system of AdSA, namely the NPV, ROI and BCA. NPV is a basic form of economic analysis that sums all discounted costs and revenues expected for the asset or investment alternatives through its lifecycle (IIMM 2015). ROI is the net profits (or savings) expected from a given investment expressed as a percentage of the return on investment. BCA is an extension of NPV, but considering organizational and stakeholders impacts, that involves identifying and evaluating the positive (benefits) and negative (costs) impacts of various options in monetary terms, over a given period of analysis.

If NPV and ROI are used to express value, option 1 is preferable, followed by option 2. However, if the social impacts are also factored in with the BCA analysis, the option with the highest value is option 2.

The BCA technique was found to be appropriate when decisions involve a balancing of performance, risk and cost (IIMM 2015; Cruz and Marques 2012), as required by the ISO 55000 standards on asset management.

In the case study, the stakeholders impact of each competing option was represented by value drivers. A qualitative assessment of these value drivers indicate option 2 as delivering the highest value.

6 Conclusions

This paper discusses the best option for intervening in a critical asset of the industrial water supply system of ILZS, considering both short and long-term views. Various methods were tested to find the best trade-off between performance, risk and cost. Best value was determined also by analysing the impacts of the competing options on the stakeholders of the industrial water supply system and the AdSA strategic objectives.

The industrial water supply system was found to be critical because its interruption due to failure or sudden deterioration can lead to extremely high economic losses for the industries in the ILZS. Given the criticality of the industrial processes involved, this would subsequently lead to unacceptable impacts at the national level as well. The revenues and reputation of AdSA were thus considered of paramount importance.

The main objective of the case study was to test the applicability of a proposed value-based decision-making process. The process shows potential to be a valuable asset management tool, but more testing is needed to conclude about the contexts in which the process valid (e.g. the maturity level of asset management in the organization) and what kind of improvements or adjustments are needed.

References

AdSA: Plan of intervention in the water supply subsystem of Santo André (in Portuguese). Phase I - Reference Situation Diagnosis, AdP, AdSA, Lisbon (2014)

AdSA: Plan of intervention in the water supply subsystem of Santo André (in Portuguese). Phase III - Action Plan, AdP, AdSA, Lisbon (2015)

Alegre, H., Coelho, S.: Infrastructure asset management of urban water systems. In: Water Supply System Analysis - Selected Topics, chap. 3, pp. 49–73, InTech (2013)

Alegre, H., Vitorino, D., Coelho, S.: Infrastructure value index: a powerful modelling tool for combined long-term planning of linear and vertical assets. Sci. Direct: Procedia Eng. **89**, 1428–1436 (2014). https://doi.org/10.1016/j.proeng.2014.11.469

Cruz, C.O., Marques, R.C.: The State and Public-Private Partnerships (in Portuguese). Editions Sílabo, Lisbon (2012)

European Commission: Guide to cost-benefit analysis of investment projects - economic appraisal tool for cohesion policy 2014–2020, European Commission (2014) http://europa.eu. ISBN 978-92-79-34796-2. https://doi.org/10.2776/97516

GFMAM: The Asset Management Landscape, 1st edn., Global Forum on Maintenance and Asset Management (2016). https://gfmam.org/. ISBN: 978-0-9870602-5-9

Haddadi, A., Johansen, A., Andersen, B.: A conceptual framework to enhance value creation in construction projects. Procedia Comput. Sci. **100**, 565–573 (2016)

Hughes, A., Moore, D.: Driving value from Logan city council's water assets. In: Logan: AMPEAK Conference Logan Water Infrastructure Alliance, QLD, Australia (2017)

IIMM: International Infrastructure Management Manual, IPWEA - Institute of Public Works Engineering, Australasia (2015)

LNEC & ERSAR:. Guide to quality assessment of water and waste services rendered to users. 3rd generation of the evaluation system (in Portuguese). Technical Guide n.22. LNEC e ERSAR, Lisbon (2017)

Parlikad, A., Srinivasan, R.: Whole - Life Value-Based Decision-Making in Asset Management. Cambridge Centre for Smart Infrastructure & Construction, London (2016). ISBN 978-0-7277-6061-6

Trindade, M., Almeida, N.: Value-based management of constructed assets (in Portuguese). In: 2nd Encounter on Quality and Innovation in Construction, QIC2016. LNEC, pp. 21–23, Lisbon (2016)

Trindade, M., Almeida, N., Finger, M., Ferreira, D.: Value-based opportunity process for asset intensive organizations. In: 12th World Congress on Engineering Asset Management & 13th International Conference on Vibration Engineering and Technology of Machinery, Brisbane (2017)

Serious Games for Decision-Making Processes: A Systematic Literature Review

Matti Rissanen(✉), Lasse Metso, Kalle Elfvengren, and Tiina Sinkkonen

Industrial Engineering and Management,
Lappeenranta-Lahti University of Technology, Lappeenra, Finland
{matti.rissanen,lasse.metso,kalle.elfvengren,
tiina.sinkkonen}@lut.fi

Abstract. In asset management one key element is the availability of information at right time for decision-making. Decision-making processes in organizations are often multi-actor problems. Studies to train and improve decision-making with serious games have been previously conducted. The term "serious game" describes an intention of the player or the developer of the game to include a purpose other than pure entertainment into the game. Serious games communicate their purpose to the player through immersive and fun gameplay. While being engaged in gameplay, the players are knowingly or subconsciously more receptive to learning and skill acquisition. The number of articles about serious games in decision-making has increased during recent years. In this study, a systematic literature review is performed by using the Scopus database. The purpose of the paper is to categorize and analyse the content of existing literature of serious games for decision-making processes in organizations. The paper also raises some points on what the current games lack considering organizational and technological trends.

Keywords: Serious games · Decision-making · Decision support · Simulation games

1 Introduction

Physical asset management can be seen as a process of identification, design, construction, operation and maintenance of assets. Management of critical assets requires reliable information for the decision-making process. (Faiz and Edirisinghe 2009) Better decision-making can be done by using better and more efficient information management (Vanier 2001). Serious games and games for learning in general are mostly used for knowledge acquisition but they have been successfully used also for example for skill acquisition and behavior change (Boyle et al. 2016).

Serious games are an approach for incorporating learning or training purpose in a fun and engaging game. Zyda (2005) defines a serious game as a mental contest where through entertaining gameplay the players' achieve a learning purpose built into the game. Playing any kind of a game is based on making series of decisions in the environment of the game. Abt (1970) mentions that testing different choices of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 330–338, 2020.
https://doi.org/10.1007/978-3-030-48021-9_37

decisions is too expensive and risky in many industrial and governmental settings due to complexity and incomplete information and serious games provide a way to compare alternatives. With serious games, complex decision-making processes can be simulated, trained and improved.

Conducting a preliminary literature review to support the design of a collaborative serious game for supporting inter-organizational decision-making processes in business ecosystems proved to be difficult. This systematic literature review investigates how serious games are used to support decision-making processes in organizations and aims to either support further research in the field of serious games in inter-organizational collaboration between companies or to reveal a gap in academic literature regarding such games.

In the Scopus database, over 200 papers can be found about supporting decision-making with a serious game between the years 2003 and 2017. A general conclusion after reviewing the large number of articles is that serious games are relevant for educational purposes as well as training decision-making in organizations. Serious games offer a possibility to learn and test complex decision-making situations in many different fields by simulating decision processes and studying the data of users' interactions. This literature review excludes serious game research concentrating purely on educational purposes because there is previous research on these topics with various other terms such as educational games or game-based learning.

2 Research Design

Scopus database was used to search with string "*TITLE-ABS-KEY ((“serious game” AND “decision making”) AND NOT (“education” AND NOT “training”))*". This produced 197 search results that were screened out down to a final sample for analysis. As serious game as a term is used to also describe games with a purely educational use, these articles were excluded from the search but separately processed. Other searches were conducted in Scopus, Google Scholar and Google search engine to get a wider understanding to the topic in hand. For example, including other game terms "simulation game", "video game" and "computer game" in the search increased the amount of results in Scopus to 926. The other game terms steered the results too far from the training of decision-making processes and therefore the presented search string was deemed suitable for the purpose of this research.

The 197 search results were first combed through on abstract level and 82 of them were excluded from the review due to not presenting a specific game or decision-making process in organizational context. The remaining 115 papers were skimmed through and 78 were removed with the same criteria as on abstract level including the removal of papers with target players being customers or patients of an organization instead of its employees, or not indicating its target players clearly. For example, a serious game dedicated to rehabilitating drug addicts might aim to improve the decision-making of the drug addict but does not directly affect the decision processes within the organization facilitating the game. Third and final screening removed 14 papers out of the remaining 37 based on the papers introducing a game too early in its concept phase for the purposes of this paper. The games too early in the concept phase would not communicate their target players or the decision-making processes included in the game and therefore making the categorizing of those papers impossible. Out of the 39 search results disregarded with the "*AND NOT ("education" AND NOT "training")*" part of the search string, 5 were skimmed through based on abstracts and one was chosen for the final sample. This search condition excluded papers dedicated to educational purposes but specifically included educational papers with a training purpose.

Qualitative content analysis focuses on meaning rather than quantification. Content analysis is a technique to identify reference models and to estimate parameters from textual data (Luna-Reyes and Andersen 2003). Content analysis is a systematic research method (Krippendorff et al. 1980; Downe-Wambolt 1992). Content analysis is used to analyse data, which is in textual form.

3 Findings

The usage of serious games to train tasks involving decision-making within organizations is gaining traction. Out of the 24 analysed papers, 17 were written during the past five years (2013–2017) and the rest between the years 2005 and 2012. The 24 papers presented 20 different serious games of which only 2 are non-digital and 18 playable with computers or mobile devices. The papers are numbered in coding by the games they present. SKYBOARD-game is presented in four papers (4a-d), D-CITE in three (7a-c) and Muller and van de Boer-Visschedijk (2017) introduces two games, BrainRun (16) and Casual Tactical Decision Game (17).

13 games have simulation as primary or secondary game genre and 7 have role play. Most primarily role playing games have simulation as a secondary genre. The distinction between simulation and role play comes from a player assuming a role, for example a physician, and playing according to the role. Simulation puts the player in the game as themselves and puts them through tasks that simulate ones from the real world. Simulation is a natural choice for training purposes as the player automatically makes the link between game and the purpose of the training in real life. Role play on the other hand is a genre that is used to better support the engagement and immersion to the game play, especially when the player does not practice exactly the presented task in reality. Other primary or secondary game genres recognized are quiz (3), action (1), adventure (1) and strategy (1). A quiz presents a series of questions the player answers and at the same time learns more about the topics included, action game involves fast paced situations that require swift decision-making, adventure involves different kinds of events in a game world, and strategy requires decision-making on a strategic level. All made categorizations are presented in Appendix 1.

The decision-making situations in the analysed serious games were compared with the four phases of rational decision-making: intelligence, design, choice and implementation (Turban et al. 2010). Most of the decision-making training in the analysed games belong to the design and choice phases. The design phase means that the decision-making processes under training involve inventing, developing and analysing different courses of action. In the choice phase, the player selects an action from the alternatives developed in the design phase. This is not surprising as most of the processes in the games present some form of vague emergency or critical situation, which require swift, stable and correct decision-making – for example medical complex surgery procedures or modelling decisions of a player during critical situation in supply chain management. In the intelligence phase, the player would search for conditions that require decision-making and in the implementation phase adapt the made decision. Occurrence of each phase within the analysed games is presented in Table 1. Most of the games require the player to act in more than one phase of rational decision-making.

Table 1. Occurrence of rational decision-making phases in analysed games.

Intelligence	Design	Choice	Implementation
4	16	14	4

7 of the 20 analysed games are primarily related to healthcare, 4 to military, 3 to managing a conflict situation and 6 to other including project management, airport management, infrastructure planning, sports and financial decision-making. The use of serious games in training of tasks in the design and choice phases of rational decision-making also reflects in the fact that 14 of the games are single player and 6 multiplayer. 17 of the games are intra-organizational and only 3 involve roles of players inter-organizationally. All three inter-organizational, two of which are board games, are collaborative games. Inter-organizational collaboration is key in future. Games recognized here reflect that poorly, especially since only one game supports inter-organizational training in digital form, which is much more accessible to play than physical board game.

The stage of the analysed game in the papers concentrates on prototype level. The scale is from a concept, a prototype, validated to released. A concept means the paper does not present any test results, only thoughts on the design of the game. A prototype includes a proper playable version of the game and a validated game has gone through tests within its intended target group or otherwise can draw reliable conclusions on achieving its purpose. A released game means that the game is commercially available. 4 papers present a concept, 13 a prototype, 5 a validated and 2 a released game.

Content analysis visualization was done by creating a word cloud using NVivo software (see Fig. 1). In the word cloud, word frequency from the 24 analysed papers was examined. The most used word was *game* (synonyms included) with 2419 times in the analysed papers. The size of text in the word cloud tells the frequency of the word, a bigger font meaning higher frequency. The next frequent words were: *learning* (1698), *making* (1607), *training* (1591), *design* (1456), *process* (1228), *work* (1170), *take* (1061), *results* (1060), *decision* (1021), *study* (993), *performance* (972), *player* (899), *control* (840), *management* (830), and *serious* (801). However, a word of interest, collaboration, was far from the most frequent (*collaborative* (307) and *cdm* (242)). The last word in list (100^{th}) was *interaction*.

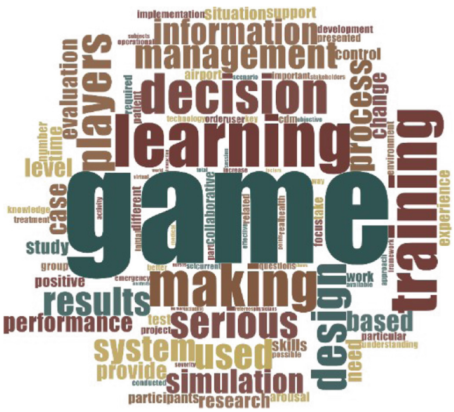


Fig. 1. Word cloud

4 Discussion and Conclusions

New and developing digital tools allow for better and more realistic simulations. Serious games including simulations in this study are still missing some elements, such as smell and feel. In "Auction Game" (Astor et al. 2013) biofeedback is used to give the player some of these missing elements but other games do not report similar usage. In addition, Anon (2010) notes that serious games do not provide stimulation for all senses using the lack of the smell of a battlefield as an example from their game.

The two most important findings were the lack of presence of the intelligence phase of the four phases of rational decision-making in the analysed games and the fact that the games concentrated on decision-making processes within a single organization. The intelligence phase includes determination of whether a problem requiring decision-making exists and the explicit definition of the possible problem. In the analysed papers, almost all of the presented games were already past this phase and the player knew the situation around which they had to make decisions. The lack of implementation phase is not as surprising since most of the games simulated this phase for the player after the player completed the choice phase. A quick glance at simulations used to optimize risk management procedures (e.g. a nuclear power plant, see Williamson et al. 2012) show that games to support the intelligence phase exist with other terminology (simulation game, simulation, gamification) than serious games, while fulfilling Zyda's (2005) definition of serious games–being entertaining and incorporating a learning purpose. Further research is required to connect the different terms to get a better understanding about using games with serious purpose to improve decision-making processes in organizations.

Inter-organizational decision process increases in complexity compared with one where the decision influences only a single organization. As Abt (1970) notes, complexity is one reason to use serious games to go through decision-making processes. However, only 3 out of the 20 analysed games involved inter-organizational decision-making. Here the further research directs towards designing and using serious games more in inter-organizational setting to test if they are indeed able to answer the issue of complexity of decision-making processes.

Decision-making in asset management requires relevant information for the basis for decisions. Serious games offer a tool for increasing the knowledge of the players about the topics they make decisions about in reality or develop the decision-making processes through simulation. Therefore, serious games are versatile in the way that they can both ensure the decision maker has the relevant information and is capable of making the decisions.

Coded Papers

1. Astor, P.J., Adam, M.T.P., Jerčić, P., Schaaff, K. and Weinhardt, C., 2013. 'Integrating biosignals into information systems: A NeuroIS tool for improving emotion regulation', *Journal of Management Information Systems*, Vol. 30 (3), pp. 247-277.
2. Basole, R.C., Bodner, D.A. and Rouse, W.B., 2013. 'Healthcare management through organizational simulation', *Decision Support Systems*, Vol. 55 (2), pp. 552- 563.
3. Bekebrede, G., Mayer, I., Van Houten, S.P., Chin, R. and Verbraeck, A., 2005. 'How serious are serious games? Some lessons from infra-games', *in Proceedings of DiGRA 2005 Conference: Changing Views - Worlds in Play, Vancouver, Canada*.

4a. Corrigan, S., Mårtensson, L., Kay, A., Okwir, S., Ulfvengren, P. and McDonald, N., 2015. 'Preparing for Airport Collaborative Decision Making (A-CDM) implementation: an evaluation and recommendations', *Cognition, Technology and Work*, Vol. 17 (2), pp. 207-218.

4b. Corrigan, S., McDonald, N., Zon, R., Maij, A. and Mårtensson, L., 2013. 'Collaborative learning & serious game development', *in Proceedings of the SESAR Innovation Days 2013, Stockholm, Sweden*.

4c. Corrigan, S., Zon, G.D.R., Maij, A., McDonald, N. and Mårtensson, L., 2015. 'An approach to collaborative learning and the serious game development', *Cogni- tion, Technology and Work*, Vol. 17 (2), pp. 269-278.

4d. Zon, R., Corrigan, S., McDonald, N. and Maij, A., 2012. 'A Learning, training & mentoring framework (LTM) & the role of serious games to facilitate sustainable change in the aviation industry', *in Proceedings of 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012, PSAM11 ESREL 2012, Helsinki, Finland*.

5. Crozier, M.D., Moore, A.P. and Verna, T.M., 2010. 'Serious games for first person 'thinker'', *in Proceedings of Spring Simulation Multiconference 2010, SpringSim'10, Orlando, FL, USA*.
6. Ferretti, I., Mazzoldi, L., Zanoni, S. and Zavanella, L., 2016. 'Learning effects in professional training for emergencies management in industrial plants with serious games', *in Proceedings of 5th International Workshop on Innovative Simulation for Health Care, IWISH 2016, Larnaca, Cyprus*.

7a. Freese, M., 2016. 'Game-based learning: An approach for improving collaborative airport management', *in Proceedings of the European Conference on Games-based Learning, Paisley, United Kingdom*.

7b. Freese, M. and Drees, S., 2016. 'D-CITE - A serious game to analyze complex decision-making in air traffic management', in De Gloria A., Veltkamp R. (eds) *Games and Learning Alliance. GALA 2015. Lecture Notes in Computer Science,* Vol. 9599, Springer, Cham, Switzerland.
7c. Schier, S., Freese, M. and Mühlhausen, T., 2016. 'Serious gaming in airport management: Transformation from a validation tool to a learning environment', in Bottino R., Jeuring J., Veltkamp R. (eds) *Games and Learning Alliance. GALA 2016. Lecture Notes in Computer Science*, Vol. 10056, Springer, Cham, Switzerland.
8. Graafland, M., Vollebergh, M.F., Lagarde, S.M., Van Haperen, M., Bemelman, W. A. and Schijven, M.P., 2014. 'A serious game can be a valid method to train clinical decision-making in surgery', *World Journal of Surgery*, Vol. 38 (12), pp. 3056-3062.
9. Gulec, U. and Yilmaz, M., 2016. 'A serious game for improving the decision making skills and knowledge levels of Turkish football referees according to the laws of the game', *SpringerPlus*, Vol. 5 (1).
10. Jarvis, S. and De Freitas, S., 2009. 'Evaluation of an immersive learning programme to support triage training: In-game feedback and its effect on learning transfer', *in Proceedings of the 2009 Conference in Games and Virtual Worlds for Serious Applications, VS-GAMES 2009*, Coventry, United Kingdom.
11. Kurapati, S., Kolfschoten, G., Verbraeck, A., Corsi, T.M. and Brazier, F., 2013. 'Exploring shared situational awareness in supply chain disruptions', *in ISCRAM 2013 Conference Proceedings - 10th International Conference on Information Systems for Crisis Response and Management*, Baden-Baden, Germany.
12. Lelardeux, C.P., Panzoli, D., Lagarrigue, P. and Jessel, J., 2016. 'Making decisions in a virtual operating room', *in Proceedings of 2016 International Conference on Collaboration Technologies and Systems, CTS 2016*, Orlando, FL, USA.
13. MacKinnon, L. and Bacon, L., 2012. 'Developing realistic crisis management training', *in ISCRAM 2012 Conference Proceedings - 9th International Conference on Information Systems for Crisis Response and Management*, Vancouver, Canada.
14. Mohan, D., Rosengart, M.R., Fischhoff, B., Angus, D.C., Farris, C., Yealy, D.M., Wallace, D.J. and Barnato, A.E., 2016. 'Testing a videogame intervention to recalibrate physician heuristics in trauma triage: Study protocol for a randomized controlled trial', *BMC Emergency Medicine*, Vol. 16 (1).
15. Monga, C., Jain, J., Kumar, S. and Sandeep, A., 2017. 'Context rich digital games for better learnability in the IT project management context', *in Proceedings of the 9th International Conference on Computer Supported Education CSEDU 2017*, Porto, Portugal.
16 & 17. Muller, T. and Van De Boer-Visschedijk, G., 2017. 'Mobile gaming for military: Two case studies', *in Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017*, Graz, Austria.
18. Ribeiro, C., Monteiro, M., Pereira, J.M., Antunes, T. and Hauge, J.B., 2014. 'Sepsis fast track: A serious game for medical decision making', in Ma M., Oliveira M.F., Baalsrud Hauge J. (eds) *Serious Games Development and*

Applications. SGDA 2014. Lecture Notes in Computer Science, Vol. 8778, Springer, Cham, Switzerland.

19. Vidani, A.C., Chittaro, L. and Carchietti, E., 2010. 'Assessing nurses' acceptance of a serious game for emergency medical services', *in Proceedings of 2nd International Conference on Games and Virtual Worlds for Serious Applications, VS-GAMES 2010*, Braga, Portugal.
20. Anon, 2010. 'Serious Games and Tactical decision making training', *in Proceedings of Spring Simulation Multiconference 2010, SpringSim'10, Orlando, FL, USA*.

Appendix 1. Categorization of Serious Games for Decision-Making.

Topic	Game platform		Amount of players		Training content for		Game genre				Subject of the game					Stage of the game (at the time of publication)			
	Digital	Non-digital	Single player	Multi player	Single organization	Inter-organizational	Simulation	RPG	Quiz	Other	Health care	Military	Conflict management	Airport management	Other	Concept	Prototype	Validated	Released
1	x		x		x			1							1			x	
2	x		x		x		1				1				2			x	
3	x			x	x		1								1	x			
4a		x		x		x	2	1						1		x			
4b		x		x		x	2	1						1			x		
4c		x		x		x	2	1						1				x	
4d		x		x		x	2	1						1		x			
5	x			x	x		1					1							x
6	x		x		x		2	1			2		1				x		
7a	x			x		x	2	1						1			x		
7b	x			x		x	2	1						1			x		
7c	x			x		x	2	1						1			x		
8	x		x		x				1		1						x		
9	x		x		x				1						1			x	
10	x		x		x		1				1							x	
11		x		x		x	2	1					1				x		
12	x			x	x		1				1						x		
13	x		x		x		1						1			x			
14	x		x		x			2		1	1						x		
15	x		x		x		2	1							1		x		
16	x		x		x				1			1					x		
17	x		x		x					1		1					x		
18	x		x		x		1				1						x		
19	x		x		x		1				1						x		
20	x		x		x		1					1							x

1 = primary, 2 = secondary

References

Abt, C.: Serious Games. Reprint, University Press of America, Lanham (1987)

Boyle, E.A., Hainey, T., Connolly, T.M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., Pereira, J.: An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Comput. Educ. **94**, 178–192 (2016)

Downe-Wamboldt, B.: Content analysis: method, applications and issues. Health Care Women Int. **13**(3), 313–321 (1992)

Faiz, R.B., Edirisinghe, E.A.: Decision making for predictive maintenance in asset information management. Interdisc. J. Inf. Knowl. Manag. **4**(1), 23–36 (2009)

Krippendorff, K.: Content Analysis: An Introduction to Its Methodology. Sage Publications, Newbury Park (1980)

Luna-Reyes, L.F., Andersen, D.L.: Collecting and analyzing qualitative data for system dynamics: methods and models. Syst. Dyn. Rev. **19**(4), 271–296 (2003)

Turban, E., Sharda, R., Delen, D.: Decision Support and Business Intelligence Systems, 9th edn. Prentice Hall Press, Upper Saddle River (2010)

Vanier, D.J.: Why industry needs management tools. J. Comput. Civil Eng. **15**(1), 35–43 (2001)

Williamson, R.L., Hales, J.D., Novascone, S.R., Tonks, M.R., Gaston, D.R., Permann, C.J., Andrs, D., Martineau, R.C.: Multidimensional multiphysics simulation of nuclear fuel behavior. J. Nuclear Mater. **423**(1–3), 149–163 (2012)

Zyda, M.: From visual simulation to virtual reality to games. Computer **38**(9), 25–32 (2005)

Equipment Life Cycle Management Based on Private Blockchain and Smart Contract

Guicang Peng[1(✉)], Songpu Ai[1(✉)], Li Zhang[2], Chunming Rong[1], and Tore Markeset[1]

[1] University of Stavanger, Stavanger, Norway
guicang.peng@live.com, {songpu.ai, chunming.rong, tore.markeset}@uis.no
[2] University of Electronic Science and Technology of China, Chengdu, China
zhangli240@uestc.edu.cn

Abstract. Equipment management is gradually becoming more decentralized, and, in many cases, the equipment owner, operator, maintainer and inspector are not the same legal entity. This slows down equipment data transmission between stakeholders and reduces business and technical process automation. In this paper, we discuss how the application of distributed ledger concept based on private blockchain and smart contract technology can resolve these challenges and create a more automated and surveillance-free equipment life cycle management process.

1 Introduction

Conventional equipment ledgers are hosted by the equipment owner often as a central database. It is the owner's responsibility to maintain the central database so that other stakeholders can retrieve updated data. However, different stakeholders may use different systems and protocols to store and share equipment data, and the different systems and protocols cause data transmission lags and inaccuracies between stakeholders. This would not only reduce business and technical process automation, but also create challenges for data verification and validation which contract execution and regulatory auditing largely depend on. A typical example of this deficiency is the aviation industry: It costs about 1 billion USD annually to transfer aircrafts between operators, with a large part of the expenses being caused by aircraft records transmission and approval (Canaday 2017).

To resolve these challenges, the use of a distributed equipment ledger can be the first step. As UK Government Chief Scientific Adviser Mark Walport described in his report: "A distributed ledger is essentially an asset database that can be shared across a network of multiple sites, geographies or institutions" (Walport 2016). However, a distributed ledger alone is not capable of solving the existing challenges; questions remain, such as: Based on a distributed ledger, how can we enhance business and technical process automation whilst ensuring stakeholder consensus and data authenticity? One answer is the private blockchain (Chowdhury et al. 2018). With its advantages, such as decentralization, transparency and anti-tampering, it provides a

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 339–348, 2020.
https://doi.org/10.1007/978-3-030-48021-9_38

solid foundation for distributed ledger implementation. Moreover, together with the smart contract, the distributed ledger can be used to automate business and technical processes without third-party surveillance (Walport 2016).

The rest of the paper is organized as follows: Sect. 2 will explain what the distributed ledger concept, private blockchain and smart contract are, and how they can be implemented together for general business scenarios. Section 3 will discuss how equipment life cycle management can be improved by the blockchain and smart contract. Section 4 will conclude on the pros and cons of the proposed solution. However, due to page limitations, we will not discuss the technical details of the private blockchain and smart contract to a research level; instead, we will only discuss how to customize and apply them to improve the equipment life cycle management process.

2 Distributed Equipment Ledger, Blockchain and Smart Contract

2.1 Distributed Equipment Ledger

A distributed equipment ledger is owned by all stakeholders as an auto-convergent database, in which every stakeholder maintains the part of the database that is relevant only to themselves, according to certain contract and regulation requirements. Equipment data can be recorded accurately and in a timely way by data generators themselves on their own premises; also, because a distributed ledger has to implement a unified data-storing and sharing protocol for collective usage, the unified protocols will further enhance data transmission efficiency and accuracy between stakeholders. The distributed equipment ledger concept is illustrated in Fig. 1, with boxes representing stakeholders, and arrows representing data flow between stakeholders and the distributed ledger.

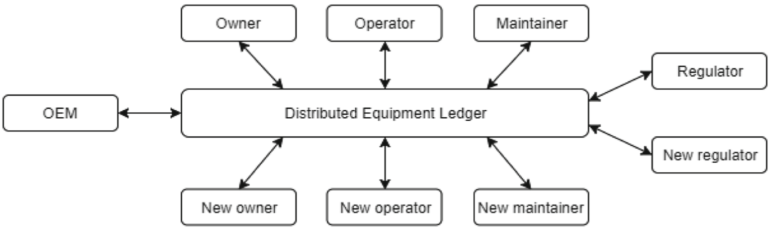


Fig. 1. Distributed equipment ledger

2.2 Private Blockchain

Blockchain: A blockchain is a decentralized ledger with record groups stored simultaneously on multiple computers. A group of records is called a block. Once a block is entered into the ledger, it is immutable and linked using cryptography (Yaga et al. 2018). A basic blockchain structure is illustrated in Fig. 2. Each block has at least two core elements: 1) A timestamp; 2) Previous block's hash value generated by cryptography function; and 3) Data stored in the block (Conte de Leon *et al*. 2017).

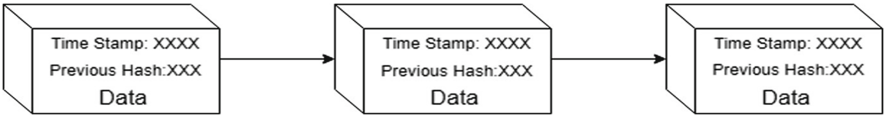


Fig. 2. Basic Structure of a blockchain

A sound blockchain has three features which are relevant to our discussion here: 1) The blockchain is immutable and data stored in the blockchain is not changeable; 2) Blocks can only be added to the blockchain after passing consensus mechanism; 3) After a new block is successfully added to the blockchain, the new blockchain will auto-synchronize itself across the network to ensure the updated blockchain copy prevails. These three features of the blockchain are the backbone of anti-tampering and timely convergent distributed ledger system (Conte de Leon *et al.* 2017).

Decentralized Network: The peer to peer network without central administration, anyone within the blockchain decentralized network keeps an updated copy of the blockchain. A blockchain decentralized network is illustrated in Fig. 3.

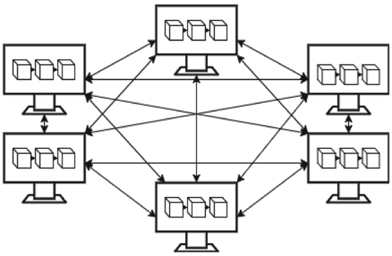


Fig. 3. Blockchain decentralized network

Consensus Mechanism: A consensus mechanism is a pre-defined protocol to maintain the consistency of the blockchain among the computers within the decentralized network. A proposed block can be added to a blockchain only after passing through the blockchain's consensus mechanism (Pilkington 2016).

Public Blockchain: Anyone connected by Internet can download a public blockchain and to be an "endorsing peer" to add new block to the blockchain. The decentralized network of a public blockchain is limitless and can be extended to any computer connected by Internet.

Private Blockchain: Access to the private blockchain network is governed and only consists of people who are within the same business process or generic endeavours. For example, nodes with certain rights and obligations bounded by relevant regulations, contracts and agreements can access a private blockchain (Jayachandran 2017).

2.3 Smart Contract

Smart Contract: A computer program intended to digitally facilitate, verify, or enforce the negotiation or performance of a contract (Tar 2017), smart contracts embedded in blockchain serve the same purposes as anywhere else, the only specificity of them is that they are executed automatically and impartially without needs for human intervention (Walport 2016). A typical process of private blockchain and smart contract integration are illustrated in Fig. 4:

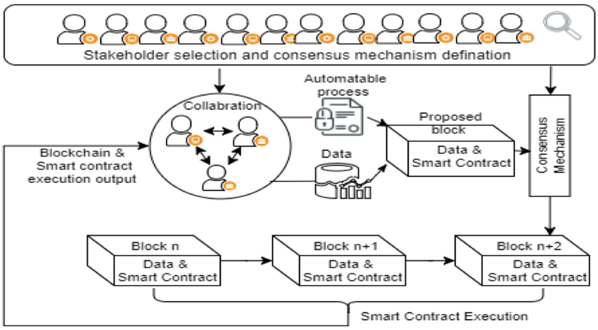


Fig. 4. Private blockchain and smart contract integration process

- **Stakeholder selection and consensus mechanism definition:** This step selects eligible stakeholders to establish private blockchain network according to relevant regulations, contracts and agreements. It is also practical to define or adjust the consensus mechanism based on the stakeholders being selected.
- **Pack smart contract and data into blocks:** Data generated during collaborations are packed into proposed blocks; automatable business and technical process are converted to smart contract and embedded into proposed blocks as computer programs. The proposed block will be added to blockchain only after passing pre-defined consensus mechanism.
- **Smart contract execution:** Smart contracts embedded into blocks are auto-executed based on block data and external inputs. Updated blockchain together with smart contract executing results is looped back to support stakeholder decision making and generating new blocks when necessary.

3 Equipment Life Cycle Management by Private Blockchain and Smart Contracts

Because equipment life cycle management is conducted by a limited number of stakeholders and is subject to confidentiality requirements in general, we follow a private blockchain and smart contract integration process in Fig. 2 to model the equipment life cycle management process. There are many blockchain platforms that can run smart contracts on blockchain, Ethereum Foundation and IBM Block- Chain Platform are two well-known examples. However, due to page limitations, we will only discuss major function requirements and the technology roadmap to realize them.

3.1 Stakeholder Selection and Consensus Mechanism Definition

There are in general eight types of stakeholders who participate in equipment life cycle management, their roles are described below:

1. The original equipment manufacturer (OEM) design and manufacture equipment, supplies spare parts, provides technical upgrades and overhaul services.
2. The regulator approves equipment's design, issues equipment market access certificates, establishes equipment operation, maintenance and scrapping baseline.
3. The dealer sells equipment and provides value-added services under distributor contract.
4. The owner owns equipment and may operate the equipment with their own personnel or lease it out. If the equipment is leased out it may be leased out including operators, or the leaser may have their own operators. The user may also buy the function of the equipment as in a functional product (Markeset and Kumar 2005).
5. The operator operates equipment under ownership, commission or lease agreements.
6. The service provider/contractor provides consultation, inspection, maintenance, repair and scrapping services etc. under service contracts. A service provider may also own or operate the equipment.
7. The supplier provides equipment consumables and spare parts.
8. Third-party auditor provides auditing services when necessary.

The eight types of stakeholders are selected and collaborate with each other according to equipment management regulations, contracts and agreements etc. The decentralized private blockchain network formed by them are illustrated in Fig. 5, where computer icons represent stakeholders, arrows between computer icon represent collaborations between stakeholders. Any stakeholder in the network is allowed to propose new blocks, stakeholders who generate data are generally responsible to generate the associated blocks. In situations where data and documents are generated by more than one parties, the associated block can be generated by any party since the consensus mechanism will ensure the data is accepted by all relevant stakeholders. When to generate blocks is mainly depending on when data is generated. For example, equipment condition monitoring data may be posted at some fixed period, then the associated block can be generated at end of each period when data is available – the schedule-based block generation. On the other hand, inspection and maintenance report is produced only after certain failure detection and repairing job is done, then the report associated block is generated based on a specific event the event-based block generation. Actual block generation is a mixture of both situations.

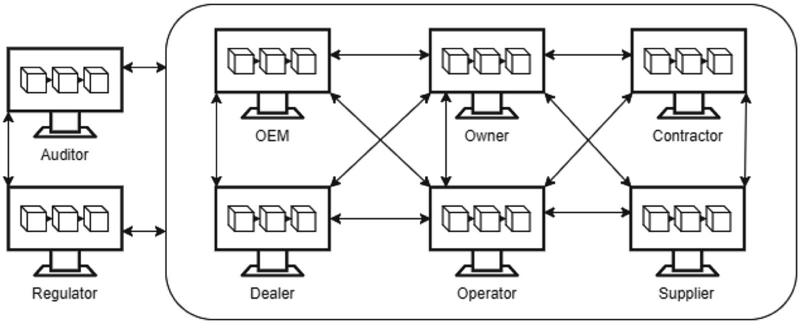


Fig. 5. Equipment lifecycle management private blockchain network

Only blocks passing consensus mechanism can be added into blockchain. In our case, the consensus mechanism depends on stakeholders' rights and obligations bounded by contracts and regulations within the equipment management process. For example, the equipment operator proposes a block with operation commission charges data to the blockchain, only equipment owner as commissioner can approve the proposed block rather than anyone else, this type of consensus mechanism is called "Selective Endorsement" (Lucas 2017). Selective endorsement relationship between stakeholders within equipment life cycle are illustrated in Table 1:

- N means having no rights for endorsement, for instance, stakeholders for specific projects usually do not have rights to endorse regulations, standards, statements and reports generated by regulators, similarly stakeholders have no rights to endorse their competitors or other stakeholders they have no collaboration with.
- ER means endorsing based on regulations, regulators endorse blocks based on relevant regulations and standards, auditors usually do the same when commissioned by regulators but also endorse blocks based on contracts as third party.
- EC means endorsing based on conditions such as regulatory, contractual or technical requirements etc. Contract-based endorse rights are always mutual between contracting parties.

Block endorsement process can be automated or manually executed. Many of nowadays business scenarios still need manual endorsement, yet automate endorsements are gradually taking holds because of technology advancement such as IoT and 5G etc. enable it. For example, block with equipment delivery confirmation can be automatically endorsed by data sent from warehouse's barcode scammer.

Table 1. Block selective endorsement relationship

Selective ensdorsement		Block endorsed by							
		Regulator	OEM	Dealer	Owner	Operator	Contractor	Supplier	Auditor
Block Generated by	Regulator	N	N	N	N	N	N	N	N
	OEM	ER	N	EC/N	EC/N	N	N	N	ER/EC/N
	Dealer	ER	EC	N	C	N	N	N	ER/EC/N
	Owner	ER	EC/N	EC/N	N	EC/N	EC/N	EC/N	ER/EC/N
	Operator	ER	N	N	EC/N	N	EC/N	EC/N	ER/EC/N
	Contractor	ER	N	N	EC/N	EC/N	N	EC/N	ER/EC/N
	Supplier	ER	N	N	EC/N	EC/N	EC/N	N	ER/EC/N
	Auditor	ER	EC/N	EC/N	EC/N	EC/N	EC/N	EC/N	N

3.2 Pack Smart Contract into Blocks

Smart contract can largely improve operation efficiency because it can be used to automate business and technical process without needs for third-party surveillance. For equipment life cycle management, it is possible to automate the following process by running a smart contract in combination with a blockchain.

1. **Regulatory compliance process:** a smart contract can automate regulation enforcement process such as: equipment market access certification based on third-party auditing results; equipment operation and service contractor qualification auditing based on stakeholder's online profiles; equipment environmental protection auditing based on IoT data.
2. **Business process:** a smart contract can automate contracts execution process such as: equipment operation commission charges approval based on equipment running parameters; equipment service contract payment approval based on equipment running parameters; spare parts and consumables procurement and logistic automation based on warehouse IoT data.
3. **Technical process:** a smart contract can automate technical process such as: equipment condition monitoring data processing and storage; activate equipment inspection, maintenance, repair or scrapping process based on equipment running parameters; suspense or continue equipment operations based on operating parameters; issuing equipment reports and notifications.

Smart contracts automating process described above are embedded into proposed block as relatively simple computer programs, together with other information within the proposed blocks, embedded smart contracts has to pass through consensus mechanism before execution on blockchain network.

3.3 Pack Data into Blocks

Equipment life cycle data is generated with different sizes and formats, to store all data on blockchain demands tremendous data traffics and storage spaces because blockchain synchronizing updated copies on all computers in the blockchain network. To overcome this challenge, an off-chain encrypted storage plus on-blockchain access management methodology is proposed based on the well-known Inter Planetary File System (Swan 2015). IPFS is a protocol and network designed to create a content-addressable, peer-to-peer method of storing and sharing hypermedia in a distributed file system (Finley 2016). Below are the IPFS features that suit for blockchain implementation:

1. Decentralized data storage and hash searchable: IPFS do not use central server to store and share data, instead data is stored on private computers as local copies and published to IPFS with only a unique hash, the hash is then used to search and download data from any computers with the published data copy. The decentralization of IPFS provides great storage scalability and do not post any data format restrictions. Meanwhile because both blockchain and IPFS are decentralized data storage system, they can collaborate on the same decentralized network seamless without extra network configuration. Stakeholders only need to pack the data's unique hash into blocks rather than the data itself, other stakeholders can download data based on its hash from IPFS only when they need it. These data hash packed blocks are coming in a small and similar size which make blockchain synchronization much quicker. It also reduces data traffic and redundant storages within the blockchain network.
2. Data anti-falsification: Data is published to IPFS with unique hash, any changes made to a data copy after publication will make the changed data copy unsearchable by the given hash anymore, this feature makes data falsification impossible after publication to IPFS. If stakeholder download data based on data hash packed in blockchain from IPFS, they will get the unchanged data copy associated with the hash for sure.
3. Data access control: For data which is ought to be open for everyone, its ok to publish data to IPFS without encrypting it first. However, for technical and business sensitive data, stakeholders can fully or partially encrypt them before published to IPFS, then control data assess by broadcasting data hash and decryption key only to authorized stakeholders through blockchain and smart contract, only those who get both the data hash and the decryption key can download and review data.

IPFS provides anti-falsification, multi-level access control and decentralized data storage network that copes well with blockchain implementation in our case. Typical equipment life cycle data access control is illustrated in Table 2:

- F means full access, for example, equipment related design and operation guidelines stored on regulators' premises are generally available to everyone.
- C means condition-based access, for example, OEM has full access to the data stored on its own premises but may give partially or no access of its data to other OEMs under specific conditions such as regulations and contractual requirements. Equipment owner generally gives condition-based access of equipment data stored on its premises to the equipment operator and verse visa.
- R means regulation-based access, depends on regional and international regulations and standards, the regulator has certain access to data stored on other stakeholders' premises.

Table 2. Decentralized data storage access control

Decentralized storage access control		Data accessed by							
		Regulator	OEM	Dealer	Owner	Operator	Contractor	Supplier	Auditor
Data stored on	Regulator	F	F	F	F	F	F	F	F
	OEM	R	F/C	C	C	C	C	C	C
	Dealer	R	C	F/C	C	N	N	N	C
	Owner	R	C	C	F/C	C	C	C	C
	Operator	R	C	N	C	F/C	C	C	C
	Contractor	R	C	N	C	C	F/C	C	C
	Supplier	R	C	N	C	C	C	F/C	C
	Auditor	R	C	C	C	C	C	C	F/C

- N means no access, for stakeholders do not have any collaboration with each other, there is generally no access to each other's on-premises data neither.

3.4 Smart Contract Execution

After a block passing through consensus mechanism, smart contract embedded in the block will auto-execute desired business and technical process based on block data and external inputs from authorized stakeholders. Typical outputs of equipment management-related smart contract execution can be: equipment maintenance notifications and reports, service payment statements and invoices, taxing and fining statements, payments confirmations, authorizations for suspending or continue further operations etc., these outputs will be automatically looped back to support management decision making and generation of new blocks.

4 Conclusion

This paper discussed how to apply private blockchain together with smart contracts and IPFS to improve equipment life cycle management process. The private blockchain provides decentralized business management framework, whilst the IPFS provides decentralized data storage, and the Smart contracts provide business and technical process automation. We believe the proposed solution would have the following advantages:

- Distributed equipment ledger enabled by private blockchain not only remove barriers for data transmission between stakeholders, it also provides a transparent and rigid collaboration framework without needs of third-party arbitration and surveillance.
- IPFS reduces tremendous data traffic and redundant data storage costs during equipment management process, IPFS also enable data storage and sharing much more efficient because it offers great scalability and posts no format restrictions for equipment data storage.

- Smart contracts enable highly rigid, transparent and surveillance-free technical and business process automation, which will increase equipment operation and maintenance performance considerably.

However, there are some disadvantages to the proposed solution as well. For example, IPFS is still at early stage of development, data access control is not a future-proof functionality of IPFS yet. For example, who gets the hush and download an encrypted data without decryption key cannot access the encrypted data, but maybe just for now since future computation power may be able to decrypt the encrypted data file anyway. This concern would intimidate many early adopters seriously. Moreover, blockchain and smart contract enable high degree of business and technical process automation, which is good for some scenarios, but maybe too rigid and robotic for the other scenarios where inherited deep uncertainties and huge risks require a certain level of human-interventions. Finally thanks Norwegian Research Council and Chinese Scholarship Association sponsor this research project.

References

Canaday, H.: Blockchain in MRO Could Happen Sooner Than You Think (2017). https://www.mro-network.com/big-data/blockchain-mro-could-happen-sooner-you-think. Accessed 15 July 2018

Conte de Leon, D., et al.: Blockchain: properties and misconceptions. Asia Pac. J. Innov. Entrepr. **11**(3), 286–300 (2017). https://doi.org/10.1108/APJIE-12-2017-034

Finley, K.: The Inventors of the internet are trying to build a turly permanent web (2016). https://www.wired.com/2016/06/inventors-internet-trying-build-truly-permanent-web/

Jayachandran, P.: The difference between public and private blockchain (2017). https://www.ibm.com/blogs/blockchain/2017/05/the-difference-between-public-and-private-blockchain/. Accessed 15 July 2018

Lucas, M.: The difference between Bitcoin and blockchain for business (2017). https://www.ibm.com/blogs/blockchain/2017/05/the-difference-between-bitcoin-and-blockchain-for-business/. Accessed 25 July 2018

Markeset, T., Kumar, U.: Product support strategy: conventional versus functional products. J. Qual. Maint. Eng. **11**(1), 53–67 (2005). Emerald Group Publishing Limited

Pilkington, M.: Blockchain technology: principles and applications. In: Research Handbook on Digital Transformations, p. 225. Edward Elgar Publishing (2016)

Swan, M.: Blockchain thinking: the brain as a dac (decentralized autonomous organization). In: Texas Bitcoin Conference. pp. 27–29, Chicago (2015)

Tar, A.: Smart Contracts, Explained (2017). https://cointele-graph.com/explained/smart-contracts-explained

Walport, M.: Distributed ledger technology: beyond blockchain, UK Government Office for Science (2016).https://www.gov.uk/government/news/distributed-ledger-technology-beyond-block-chain

Yaga, D., et al.: Blockchain technology overview, Draft NISTIR, 8202 (2018). https://csrc.nist.gov/publications/detail/nistir/8202/draft

Chowdhury, M.J.M., Colman, A., Kabir, M.A., Han, J., Sarda, P.: Blockchain versus database: a critical analysis. In: 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/Big- DataSE), 1–3 August 2018, pp. 1348–1353 (2018)

A Participatory Approach for Developing Decision Support Systems for Building Energy Plants

Samuel Patterson[1(✉)], Talara Berry[1,2], and Michael Campbell[3]

[1] Synengco Pty. Ltd., Highgate Hill, Australia
{sam.patterson, talara.berry}@synengco.com
[2] Queensland University of Technology, Brisbane, Australia
[3] Childrens Health Queensland Hospital & Health Service, South Brisbane, Australia
michael.campbell@health.qld.gov.au

Abstract. Building energy plants generate, consume and transfer a large amount of energy to deliver various services, such as heating, cooling, lighting and electricity. In large buildings, these energy plants can be complex to manage optimally. Decision support systems are useful tools for aiding managers, however, they too can also become overly complex. Participatory approaches for decision support system development and application are suggested in literature for overcoming this issue. This paper presents the foundations of a wider research project that is applying participatory techniques in the development of a decision support system at a hospital energy plant. A generic integrated building energy plant optimisation model is formulated and expressed using problem domain language with the aim of promoting participation from stakeholders, such as facility managers or maintenance personnel, that do not necessarily have modelling expertise. While the formulation is targeted towards modelling the plant's operational behaviour and decisions, the paper describes how it can be used in a decision support system for aiding short-, medium and long-term decisions of facility managers – highlighting again the model's flexibility to meet stakeholder requirements elicited using participatory techniques. The project's case study site, Lady Cilento Children's Hospital in Queensland, Australia, is introduced along with the planned methodology for participatory development of the decision support system. Finally, future directions are proposed for both further research and practical applications of the contributions.

1 Introduction

The high energy consumption of buildings is a well-recognised issue (Doukas et al. 2007; Michailidis et al. 2018). A variety of energy related services are required during a building's operation, such as hot water, chilled water, lighting and electricity. In large buildings, a system of assets delivers these services using one or more input streams, such as electricity, renewables, gas, or diesel (Chicco and Mancarella 2009). In this paper, these systems are referred to as *building energy plants*. Optimally managing the

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 349–357, 2020.
https://doi.org/10.1007/978-3-030-48021-9_39

operation and lifecycle of the plants from a whole- of-system perspective can be quite complex.

Decision support systems (DSSs) that make use of operational data and mathematical models are well recognised in literature for aiding managers of complex systems (Hung et al. 2007; Bennet and Bennet 2008). Though their development is a popular research field, they often fail to achieve their potential benefits. It is suggested that this can be due to many factors, such as failing to meet end-user requirements, high costs, and a lack of flexibility (Valls-Donderis et al. 2014). Participatory techniques, that involve consultation with stakeholders and end-users throughout the DSS development process, have been successful at improving the realised benefits of the developed DSS (Borenstein 1998; Breuer et al. 2008; Jakku and Thorburn 2010; Van Meensel et al. 2012; Valls-Donderis et al. 2014).

Several DSSs for building energy plants can be found in literature (Doukas et al. 2007; Lu et al. 2015; Michailidis et al. 2018). These represent significant research contributions with clear practical applications. However, there appears to be a lack of participatory techniques in the building energy plant DSS literature or at least a lack of published discussion if participatory techniques were used in their development. The apparent gap at the intersection between building energy plant problems, decision support system solutions and participatory techniques is the motivation for the research project introduced in this paper, as seen in Fig. 1.

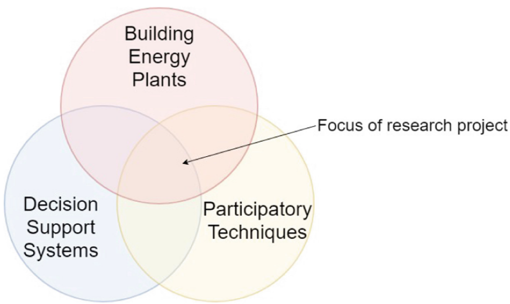


Fig. 1. Research project context and focus

The iterative and participatory validation process for DSS development described in Borenstein (1998), and used more recently in Van Meensel (2012), shown in Fig. 2, is used as the overarching research project's methodology. The term 'module' is used here in place of the Borenstein (1998) term 'subsystem' to avoid confusion with the use of 'subsystem' in the modelling framework also used in this paper. The scope of this paper covers first phase of prototype development, as highlighted in Fig. 2. This has been conducted in preparation for developing it for a hospital energy plant case study introduced later in this paper.

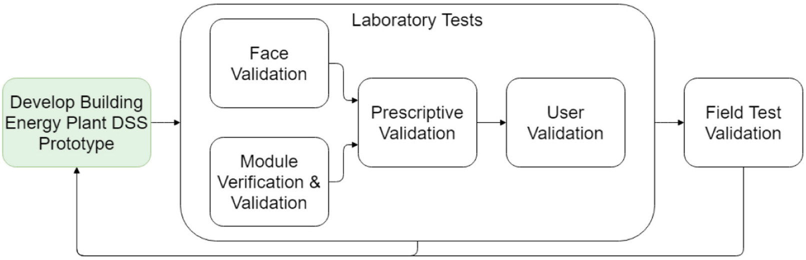


Fig. 2. Proposed participatory approach adapted from Borenstein (1998)

Section 2 describes the high-level architecture being proposed as the Building Energy Plant DSS prototype. The generic optimisation model used in the prototype is outlined in Sect. 3 using problem domain language. The case study hospital is introduced in Sect. 4 along with an outline of the initial DSS prototype and plan for the participatory validation steps. Finally, Sect. 5 summarises the paper's contributions and outlines the wider research project's next steps.

2 Proposed Decision Support System Architecture

As Fig. 2 suggests, the research approach requires an initial prototype to begin the process of iterative participation and development on a given energy plant. To serve as a generalised starting point for a prototype, a high-level DSS architecture is proposed in this section.

By conducting a cursory examination of the plant being studied, a prototype can be developing by specifying the basic functionality of the modules that are part of this high-level architecture. This initial system examination and prototype development will inevitably depend on the information available before the participation begins and so should be tailored to suit the task at hand. For example, with relatively rudimentary information about the case study building presented in Sect. 4, a low-fidelity prototype is developed that simply specifies scope and suggested functionality. This promotes assumptions that are simple and easy to understand for stakeholders so that they can engage and contribute effectively in the development process as opposed to bounding their participation within a predefined solution.

Contributions from the author's previous work have been used as a basis for the Building Energy Plant DSS architecture. The integrated modelling framework, first presented in Patterson et al. (2016) and generalised further in Patterson et al. (2017), has been used to formulate a flexible model that promotes participation (presented in Sect. 3). The decision support approach from Patterson et al. (2017), seen in Fig. 3, describes how the integrated model can be used for aiding short-, medium and long-term decisions.

These two contributions make up the 'model-based' analysis module at the core of the architecture, as shown in Fig. 4. This ensures that the DSS can be developed to aid

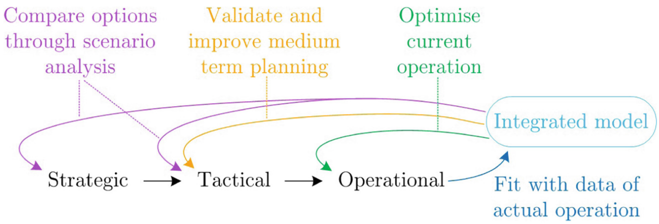


Fig. 3. General decision support approach (Patterson, et al. 2017)

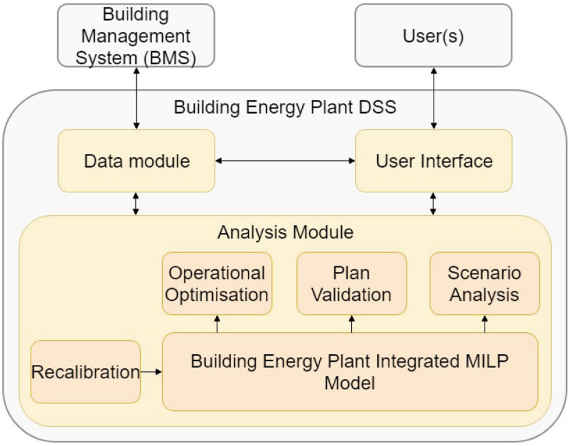


Fig. 4. Proposed building energy plant DSS architecture

a wide variety of decisions based on the requirements elicited from the participating stakeholders.

The other two modules represent the two key interfaces of the DSS. Firstly, the user interface will allow decision makers to view analysis results and edit configuration. The primary function of the data module is to collect the relevant input data from a Building Management System (BMS) and store analysis module results for users to view. It could also be used for publishing data back to the BMS, for example, if the operational optimisation module was developed to provide closed-loop set point control, the data module would be responsible for passing that control signal to the BMS.

3 Integrated Model Formulation

The modelling framework presented by Patterson et al. (2017) has been used to develop the formulation presented in this section. Language from the problem domain has been used in place of mathematics for brevity and to highlight how the modelling artefacts can be documented to be comprehendible for stakeholders without modelling expertise. This generalised formulation has been designed to be relatively simplistic so that it can apply to a variety of building energy plants with varying data availability; be more easily understood by participants that have limited modelling expertise; and be tailored to suit the specific problems that the participants wish to solve.

The framework dictates that continuous flow variables are used as the standard connections between subsystems. In this case, energy flow (MW) is used. This allows for several types of energy, for example electricity, water or gas. The forecasted energy flows between the plant and the building for each time state throughout the optimisation horizon is being proposed as the primary 'task' constraint that determines the demand on the connected subsystem instances in the model. The three types of subsystems are outlined below.

Energy Supplier Subsystem – External energy suppliers, such as grid electricity and natural gas, represent the operating costs of the plant. Contracted energy prices such as fixed monthly access charges ($/month), consumption costs ($/MWh), peak demand costs ($/maximum MW), are the key parameters of each supplier. These are used to calculate a total cost of supply over the optimisation horizon for each energy source to be included in the objective function.

Energy Junction and Storage Block Subsystem – A simple inventory formulation with multiple inputs and outputs to allow for storing, aggregating or splitting energy flows. This is parameterised by minimum, maximum and initial storage levels as well as minimum and maximum flow rates on each boundary connection.

Energy Transfer Equipment Piece Group Subsystem – A generic bank of assets that convert one type of *input* energy into an *output* energy and a *waste* stream, for example the reciprocal gas engines converting gas fuel into electricity and waste heat, or an absorption chiller converting a heat source into chilled water and waste heat. Capacity and efficiency of individual equipment pieces inside the group are parameterised to allow the model to heterogeneously reflect the actual plant.

4 Case Study

The application of these concepts on a real building energy plant and participation from stakeholders of that plant is paramount. The energy plant of Lady Cilento Children's Hospital (LCCH) in South Brisbane, Queensland, Australia has been chosen. The plant serves as a valuable proving ground due to its relatively large size, complexity, and the criticality of the services it provides.

The LCCH energy plant delivers four energy services to the hospital, electricity, chilled water, hot water and steam. The plant includes two reciprocal gas engines that feed their excess heat to absorption chillers and hot water heaters. It also has banks of electric chillers, hot water boilers, steam boilers, cooling towers and backup diesel generators. Electricity, gas and diesel are consumed by the plant to power the plant's assets. The process flow diagram presented in Fig. 5 shows the layout of the plant using the subsystems described in Sect. 3. Note, connections to and from each energy transfer equipment block and the cooling water (CW) junctions are omitted for simplicity.

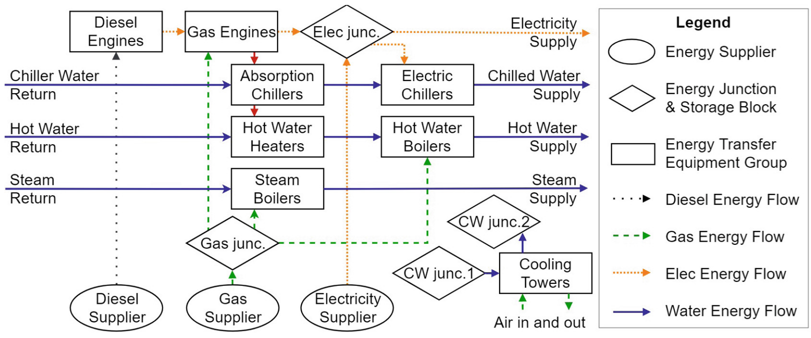


Fig. 5. LCCH model flow diagram

4.1 LCCH Energy Plant DSS Prototype

The first prototype is being developed by applying the high-level architecture shown Fig. 4 to LCCH. As explained in Sect. 2, relatively rudimentary information about the LCCH plant is being used to make a low-fidelity prototype to start the participatory process. This demonstrates the 'low barrier to entry' for using the participatory approach. A description of each module is given below.

Building Energy Plant Integrated Model Module – The formulation outlined in Sect. 3 has been applied to LCCH. A graphical representation of the system, Fig. 5, will be used alongside the problem domain language model explanations to promote engagement with participants that do not have prior modelling expertise.

Recalibration Module – The recalibration model transforms the collected data into the parameters required by the model. For example, given the energy flows provided by the BMS, the input-output method for efficiency is used to calculate equipment efficiency parameters.

Data Module – This module's functionality is primarily determined by data availability in the BMS or other related systems that are used for plant monitoring or control. In the LCCH prototype, the BACnet protocol has been specified along with a set of available input tags. A draft mapping between the available tags and the parameters required for model calibration will be used in the validation process. For example, the LCCH BMS provides return and supply water energy flow and electricity consumption of each chiller equipment piece that can be used to recalibrate that chiller's coefficient of performance.

Operational Optimisation Module – Running the model on near-real-time data and expected demand forecasts is being proposed for equipment dispatch decision support to minimise operating costs. For example, it is expected that this will be able to help the managers avoiding unnecessary demand charges from suppliers.

Plan Validation Module – Maintenance planning decisions are proposed to be supported with the model via automated analysis of the most improvement that refurbishing each equipment piece would have on the operating cost of the plant.

Scenarios Analysis Module – A scenario with an extra piece of equipment will be used to demonstrate this module. A chilled water storage tank will be added to the model to quantitatively assess the potential cost savings.

User Interface Module – Simple wireframe mock-ups that visualise the expected results of each analysis module will be used to demonstrate potential user interfaces without expensive development. These are expected to help participants understand the underlying analysis modules as well.

4.2 Proposed Methods for Participatory Decision Support System Development

As explained in Sect. 1 and shown in Fig. 2, the DSS will be iterative developed using participatory validation, based on Borenstein (Borenstein 1998). Given the access to stakeholders at LCCH, a range of methods can be employed to carry out the approach. An overview of these validation tests is given below.

Laboratory tests – The *face*, *module*, *prescriptive*, and *user validation* tests will be carried out in controlled conditions before the DSS is implemented at LCCH. They will involve scoping workshops, interviews with stakeholder participants, and include testing on actual data collected from the BMS.

Face validation – Workshops with a group of stakeholders will be used to ensure the overall prototype scope suits the specific problems faced in LCCH to inform more detailed DSS prototype development.

Module verification and validation – Example data-sets of the plant's operation will be used to construct tests that verify and validate the modules being proposed. For example, verification will be conducted to ensure the required data is available for the model module and recalibration module from the data module and that the analysis modules produce useful results that can be seen in the user interface modules. Stakeholder interviews using the results of these tests will then be conducted to validate that modules produce sensible results.

Predictive validation – Historical data-sets during known periods of interest for the analysis, such as periods of high service demand, will be selected, ingested and solved within the prototype. Interviews or independent analysis results will be used to benchmark the DSS against its expected results for the given conditions.

User validation – In addition to validating the results of the DSS, qualitative feedback and quantitative usage metrics from stakeholders using the DSS during interviews will be collected. These will be primarily used to tailor the development of the user interface module to ensure its effectiveness.

Field tests – Once the development of the DSS has converged on acceptable laboratory test results, the DSS will be implemented on the running LCCH building energy plant by installing it in their network, connecting it to the BMS and providing decision makers access to the user interface. Quantitative and qualitative methods will be used during a trail period to collect information on how accurate and suitable the DSS is for the LCCH plant stakeholder's requirements, respectively.

5 Conclusions

This paper presents the foundations of a wider research project that is applying participatory techniques for DSS development for building energy plants. A highlevel building energy plant DSS architecture is contributed for use within a participatory DSS development approach. A generic integrated model formulation is described for use as the core of the DSS's analysis module that can be altered and improved upon based on the stakeholders' requirements of the plant being studied. A description of an initial prototype and plans for the participatory development are presented for a case study on a hospital energy plant.

Aside from aims of this paper's wider research project to apply the outlined participatory approach and DSS prototype to the LCCH building energy plant, several future directions are suggested. Further work formulating and/or adding additional model subsystems is suggested to improve accuracy of the models without sacrificing generalisability. Likewise, the use of new modelling paradigms, solution techniques, or problem domains are avenues for furthering the modelling framework and DSS development approach used. As well as this, the contributions are designed for application to real-world plants and as such further case studies should prove mutually beneficial to both academia and industry.

References

Bennet, A., Bennet, D.: The decision-making process in a complex situation. In: Handbook on Decision Support Systems, vol. 1, pp. 3–20 (2008)

Borenstein, D.: Towards a practical method to validate decision support systems. Decis. Support Syst. **23**(3), 227–239 (1998)

Breuer, N., et al.: AgClimate: a case study in participatory decision support system development. Clim. Change **87**(3–4), 385–403 (2008)

Chicco, G., Mancarella, P.: Distributed multi-generation: a comprehensive view. Renew. Sustain. Energy Rev. **13**(3), 535–551 (2009)

Doukas, H., Patlitzianas, K., Iatropoulos, K., Psarras, J.: Intelligent building energy management system using rule sets. Build. Environ. **42**(10), 3562–3569 (2007)

Hung, S., Ku, Y., Liang, T., Lee, C.: Regret avoidance as a measure of DSS success: an exploratory study. Decis. Support Syst. **42**(4), 2093–2106 (2007)

Jakku, E., Thorburn, P.: A conceptual framework for guiding the participatory development of agricultural decision support systems. Agric. Syst. **103**(9), 675–682 (2010)

Lu, Y., Wang, S., Sun, Y., Yan, C.: Optimal scheduling of buildings with energy generation and thermal energy storage under dynamic electricity pricing using mixed-integer nonlinear programming. Appl. Energy Osa/vuosikerta **147**, 49–58 (2015)

Michailidis, I., et al.: Energy-efficient HVAC management using cooperative, self-trained, control agents: a real-life German building case study. Appl. Energy Osa/vuosikerta **211**, 113–125 (2018)

Patterson, S., Hyland, P., Berry, T.: Integrated modelling and decision support of continuous production systems. In: Proceedings of the 12th World Congress on Engineering Asset Management (WCEAM 2017), vol. l. Springer, Teoksessa (2017)

Patterson, S., Kozan, E., Hyland, P.: An integrated model of an open-pit coal mine: improving energy efficiency decisions. Int. J. Prod. Res. **54**(14), 4213–4227 (2016)

Valls-Donderis, P., et al.: Participatory development of decision support systems: which features of the process lead to improved uptake and better outcomes? Scand. J. For. Res. **29**(sup1), 71–83 (2014)

Van Meensel, J., et al.: Effect of a participatory approach on the successful development of agricultural decision support systems: the case of Pigs2win. Decis. Support Syst. **54**(1), 164–172 (2012)

Asset Replacement Decisions in the Context of the Mining Sector

M. H. C. Dakada and J. E. Amadi-Echendu(✉)

Department of Engineering and Technology Management, University of Pretoria, Pretoria, South Africa
{u15247555, joe.amadi-echendu}@up.ac.za

Abstract. This paper examines the application of theoretical models for practical decisions to replace engineered assets deployed for mining operations. Although return on investment, net present value, and internal rate of return are widely acknowledged theoretical approaches towards capital replacement decisions, however, the respondents mostly rely on the experience of operations and maintenance personnel in combination with adhoc service life estimates to decide when and how to replace an asset. Interestingly, the availability of funding, technical support from the asset manufacturer, supplier or vendor, and technological advancements were highlighted as having the greatest influence on replacement decisions. Noting that practical decisions to replace engineered assets deviate significantly from the theoretical models, the study proposes that asset replacement decisions should derive from a broader value-driven ethos.

1 Introduction

Globalisation, sustainability, evolutions in technology, and other imperatives are motivating increased automation and mechanisation in the capital and physical asset intensive mining sector. In order to extract and process mineral resources, a typical mining business deploys a wide range of engineered assets in the form of equipment, facilities, infrastructure and systems. With rapid evolutions in technology influencing increased automation and mechanisation (Egerton 2004; Lane and Kamp 2013; Valicek et al. 2012), there is also the need to examine the processes for acquiring, utilising and retiring engineered (i.e., the man-made) assets deployed in mining operations. When, and whether to replace an asset remains a continuing challenge confronting any business organisation that deploys engineered assets to provide goods and services to its customers and clients.

Depending on the viewpoint (e.g., designer, operator, systems integrator, vendor, etc.), the life phases or stages of a typical mining dragline may be illustrated as shown in Fig. 1.

Irrespective of the viewpoint, the duration of each life phase/stage, as well as the total life of an asset varies; one asset may have a relatively long life whereas another may have a relatively short life depending on the type and nature of the asset, its purpose, how it is utilised, the operating environment, et cetera.

There are two primordial tasks, *inter alia*, that the asset manager must perform throughout the life of an asset, they are *planning* and *decision-making*. The asset

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 358–366, 2020.
https://doi.org/10.1007/978-3-030-48021-9_40

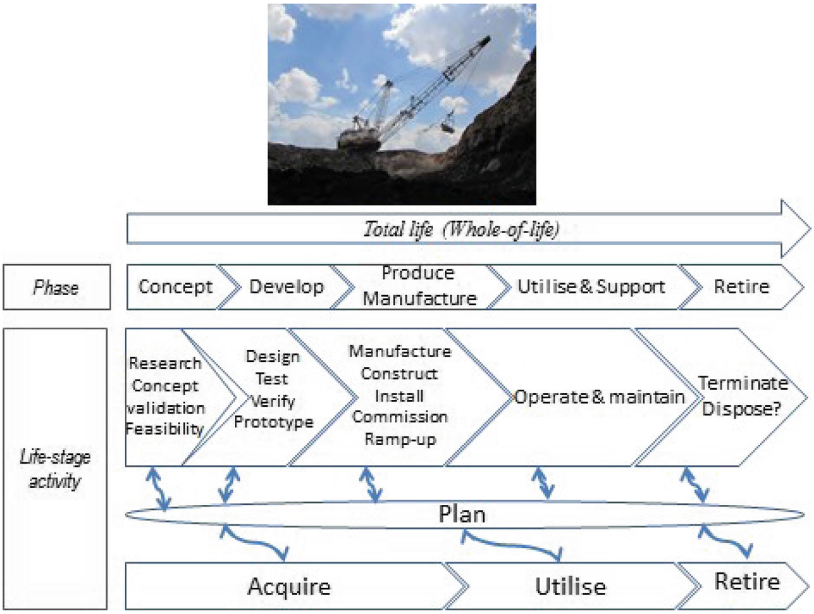


Fig. 1. Life phases or stages of an engineered asset

manager must make necessary decisions during each phase/stage, taking into consideration the respective business cycles confronting each phase/stage and the challenges encountered throughout the life of an asset. It is worthwhile remarking that decisions made during the preceding life phase(s)/stage(s) will not only have a profound influence later but also, the effects can only manifest in the latter phase(s)/stage(s) of the asset's life. One of the decisions is to replace the asset when it becomes necessary to do so. 'Replacement may be defined as the acquisition of a new asset to fulfil a particular function, together with the simultaneous scrapping or transference to another use of an old asset which has hitherto fulfilled more or less the same function' (Peters 1956). Despite the existence of formal and normative methods, this raises the question as to how decisions to replace assets are made in practice. It is in this regard that this paper examines how practical decisions are made to replace engineered assets deployed in mining operations.

In Sect. 2 of the paper, we briefly review asset life definitions and capital investment considerations. A generalised model of factors influencing asset replacement decisions is presented in Sect. 3. This is followed by a case study examination of how decisions on asset replacements have been implemented in practice. Some conclusions are highlighted in Sect. 5.

2 Asset Life and Capital Investment Considerations

In general, the life of an engineered asset can be described as depicted in Fig. 2 in terms of the following; physical life, depreciation life, useful life or service life, and economic life (Collier and Glagola 1998). The physical life represents the period of time when an asset is commissioned for use until such time that it is decommissioned from use (Alchian 1952). In accounting terms, depreciation life represents the period of time within which tax allowance is made to deduct or offset capital expended to acquire an asset. An asset will naturally degrade in use and often, the capital tax allowance depreciation gets confused with the actual loss in the capability value (i.e., degradation in capability) of an asset.

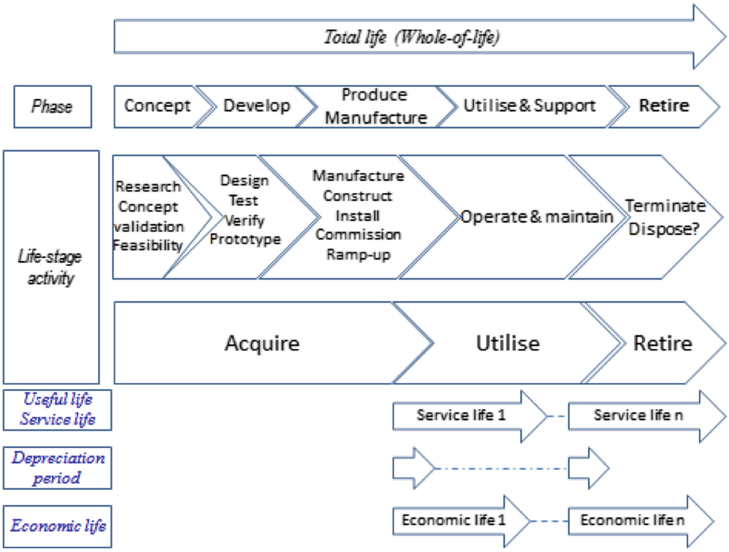


Fig. 2. Asset 'life' definitions

Useful life or service life is the period of time over which an asset is utilized for a particular purpose. An asset's economic life is sometimes also referred to as optimal life or optimal service life. Economic life can be defined as the interval within an asset's service life during which the economic value is maximum (cf: Leung and Tanchoco 1983; Meyer 1993; Regnier et al. 2004; Yatsenko and Hritonenko 2011; Hartman and Tan 2014). The economic value is realized before the termination of the physical life of an asset (Du Plessis 2010; Nurock and Porteous 2008). Jin and Kite-Powell (2000) defined the optimal life of an asset as the duration of service life until the asset is replaced for economic reasons. Similar to the asset's service life, an asset can have multiple economic lives; that is, as many times as the rights to use the asset in the same form is transferred from one legal user to another.

There are other terms, for example, remaining life which often refers to estimated time-to-failure of an asset or its potential failure interval (Si et al. 2011; Okoh et al.

2014). Remaining life is typically applied to determine or predict the lifespan of components, equipment or systems. Another term is "remaining useful life" which refers to the difference between the present time and the end of the useful/service life. An important word here is "usefulness", which determines whether or not an asset provides the means for the realisation of value.

On the one hand, much of the theory on asset replacement tends to be based on quantifiable financial economics underpinned by a cost dogma *ex post facto*, and such approach results in suboptimal asset management decisions. On the other hand, an asset is acquired as a means to provide benefits, and some aspects of the benefits are not readily quantifiable in financial terms (Amadi-Echendu 2004). Thus, in practice, the decision maker would also consider non-economic reasons in order to make the final decision to release funding to replace an asset (Scarf and Hashem 2002, Scarf et al. 2007). In any case, technological, environmental, social and political factors can influence the actual replacement decision. By definition, an asset is an investment which involves incurring present expenditure with the aim of accruing future benefits. Such investment decisions have traditionally involved net-present value (NPV), internal rate of return (IRR), cost of capital, and pay-back period considerations (Brealy et al. 2001; Lutchman 2006; Kierulff 2007; Cramer 2008; Gitman et al. 2010; Correia 2012) and latterly, economic-value added (EVA®) approaches.

3 A Model of Replacement Factors

Models for asset replacement decision-making tend to normative, thus, for brevity, we describe a model derived from the structural analysis and design technique expounded by other scholars (for example, Carole and Michael 1995; Yusuf and Smith 1996; Roh et al. 2007; Lakhoua 2013; Diviné and Le Cardinal 2014). A model of factors that may be considered when making decisions to replace an asset is illustrated in Fig. 3. The model refers to the retirement and/or replacement of an asset that is not part of a fleet, and takes into consideration accounting, financial and challenger-defender (re: Perrin 1972; Regnier et al. 2004) factors such as age-dependent physical deterioration, straight-line depreciation, opportunity costs, technological change (re: Terborgh 1956; Nair and Hopp 1992; Scarf et al. 2007; Nguyen et al. 2013; des-Bordes and Esra Büyüktahtakin 2017), and sustainability imperatives (e.g., Ralf 1992; Porter and Kramer 2006; Al-Abiyad and Handley-Schachler 2010; CGCSA 2016).

4 Case Study

In this section we summarise our investigation into practical asset replacement decision-making based on a survey of respondents selected from organisations listed in the chamber of mines whose members represent about 90% of firms involved in mining businesses in the case study country. To overcome some of the challenges with sample size and low response rate, the survey was also administered during technical meetings, seminars, and conferences. Notwithstanding, only 34 respondents fully complied with the survey requirements, and they represented 71% of mining firms listed in the chamber of mines.

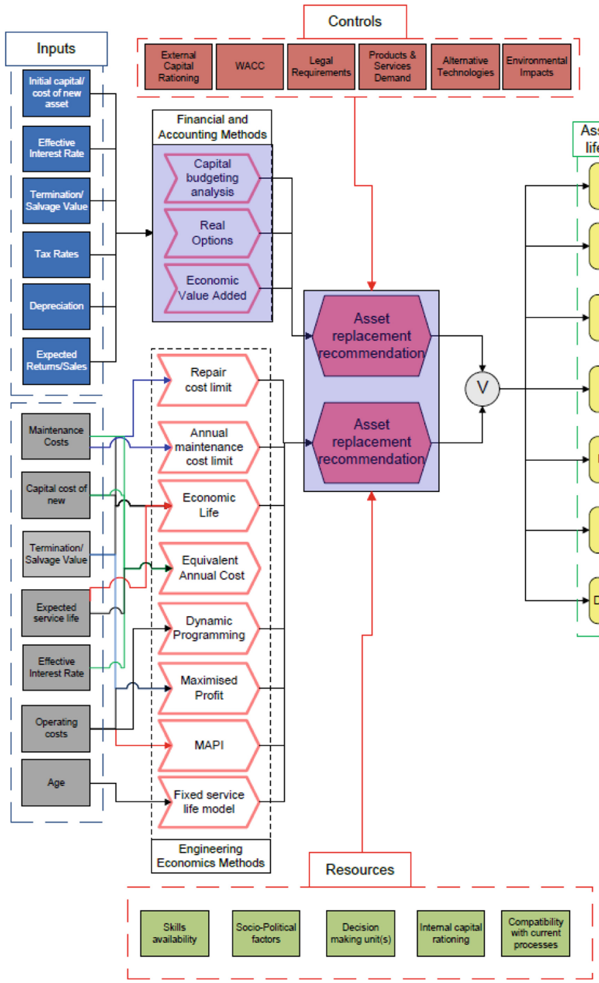


Fig. 3. A model of factors influencing asset replacement decisions

The pictures in Fig. 4 provide a summary of the some of the results obtained from the survey. Figure 4(a) shows that 24 respondents indicated that the general manager position was mostly responsible for making replacement decisions.

It is remarkable that only 4 respondents indicated that the primary users of equipment, facilities and systems deployed for mining operations were responsible for making the final decision on replacement. Figure 4(b) depicts that inadequate productivity and reduced reliability/availability of an asset were the top triggers for making the replacement decision. In fact, the two other triggers mentioned by the respondents only reiterate the *ex post facto* conventional cost dogma. Since a capital asset replacement is usually treated as a project, it is not surprising that the preferred

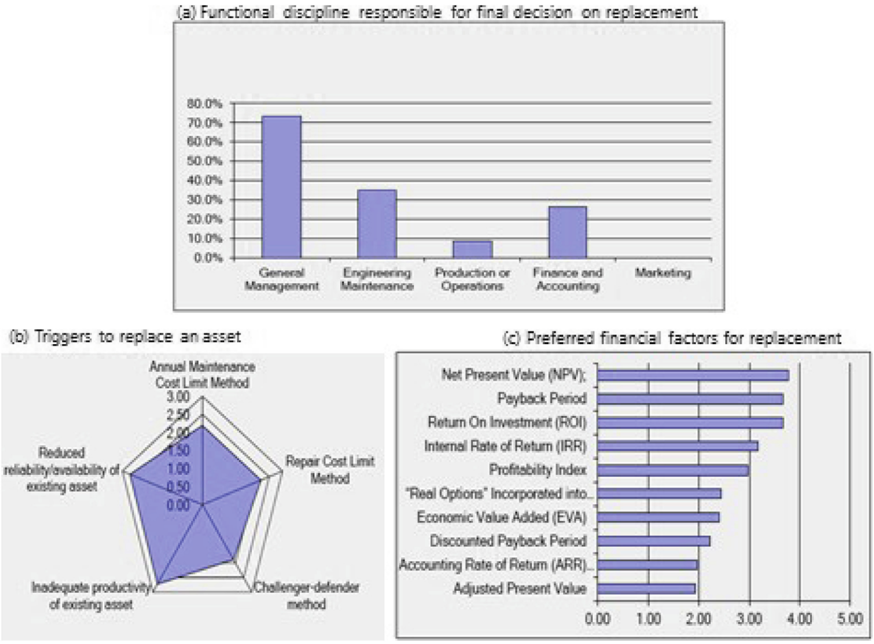


Fig. 4. Factors influencing asset replacement decisions in practice

financial approach depicted in Fig. 4(c) is dominated by the NPV factor. This is plausible because the future cash flows may be extrapolated from the past, *ceteris paribus*!

5 Conclusion

There were two asset replacement decision options investigated during study, i.e., *when* or *whether* to replace an asset. The decision as to when to replace an asset also determines the economic life or optimum service life of an asset. Based on the feedback from the respondents, we observe that the majority of mining firms tend to rely on past experience of maintenance personnel to determine the optimal service life of assets. The decision as to whether to replace an asset is commonly based on *apriori* fixed service life. Whether accurate or not, this means that the historical service live of an asset essentially determines future replacements. This highly subjective approach demands caution because it relies on the personal experiences with certain assets. On one hand, the approach of replacing an asset based on historical service lives of similar assets could also prevent mining companies from extracting maximum value from a more reliable and productive asset simply due to reference to poor performance of previous similar assets. On the other hand, non-performing assets could be kept in service for far too long while trying to match the service life to similar types of assets.

In the mining environment where production throughput is essential and safety concerns are highly important, reduced reliability and reduced productivity are the common triggers for asset replacement as these are capability and performance related. Oddly, we observed from the responses that, once funding for a replacement asset is secured, the subsequent factors considered in the decision-making process were technical support from the manufacturer/vendor/supplier, technological capabilities of the asset, as well as the cost of ownership of an asset. Socio-political and ecological footprint factors were noted but did not really influence replacement decisions. Thus, in our view, not only do practical decisions to replace assets deviate significantly from the theoretical models but also, practitioners do not adopt a broader value-driven ethos when making asset replacement decisions.

References

Al-Abiyad, S.A., Handley-Schachler, M.: Economic, social and political factors in industrial decision making: the case of property, plant and equipment replacement in Libyan manufacturing companies, World Acc. Sci. **12**(2), 189–232 (2010)

Alchian, A.A.: Economic Replacement Policy. Technical report Publication R-224. The Rand Corporation, Santa Monica, California (1952)

Amadi-Echendu, J.E.: Managing physical assets is a paradigm shift from maintenance. In: Proceedings of IEEE International Engineering Management Conference, Pan Pacific Hotel, Singapore, 18–21 October 2004, pp. 1156–1160 (2004)

Brealey, R.A., Myers, S.C., Marcus, A.J.: Fundamentals of Corporate Finance. McGraw-Hill Publishing, New York City (2001)

Carole, C., Michael, E.: How to describe your service an invitation to the structured analysis and design technique. Int. J. Serv. Ind. Manag. **6**, 6 (1995)

CGCSA: The Implementation of Carbon Tax: Consumer Goods Council of South Africa (2016)

Collier, C.A., Glagola, C.R.: Engineering Economic and Cost Analysis. Addison Wesley Longman Inc., California (1998)

Correia, C.: Capital budgeting practices in South Africa: a review. S. Afr. J. Bus. Manag. **43**, 11–29 (2012)

Cramer, P.: An analysis of cost of capital, capital structure and capital budgeting practices: a survey of South African listed companies. Meditari Acc. Res. **16**, 31–52 (2008)

Diviné, M., Le Cardinal, J.S.: How to manage virtual communities and teams using adjacencies: a process based on functional analysis and adaptive structuration theory. Int. J. e-Collab. (IJeC) **10**, 35–56 (2014)

Du Plessis, C.: Replacement of earthmoving equipment at surface coal mining operations in South Africa. MBA dissertation, University of Pretoria, Pretoria (2010). http://hdl.handle.net/2263/23398

Egerton, F.: The mechanization of UG2 mining in the Bushveld Complex. J. S. Afr. Inst. Min. Metall. **104**, 439–455 (2004)

des-Bordes, E., Esra Büyüktahtakın, I.: Optimizing capital investments under technological change and deterioration: a case study on MRI machine replacement. Eng. Econ. **62**(2), 105–131 (2017)

Gitman, L.J., Juchau, R., Flanagan, J.: Principles of Managerial Finance, Pearson Higher Education (2010)

Hartman, J.C., Tan, C.H.: Equipment replacement analysis: a literature review and directions for future research. Eng. Econ. **59**, 136–153 (2014)
Jin, D., Kite-Powell, H.L.: Optimal fleet utilization and replacement. Transp. Res. Part E: Logist. Transp. Rev. **36**, 3–20 (2000)
Kieruff, H.E.: The replacement decision: getting it right. Bus. Horiz. **50**, 231–237 (2007)
Lakhoua, M.N.: Systemic analysis of an industrial system: case study of a grain silo. Arab. J. Sci. Eng. **38**, 1243–1254 (2013)
Lane, A., Kamp, R.: Navigating above-the-ground risk in the platinum sector. J. South Afr. Inst. Min. Metall. **113**, 00 (2013)
Leung, L.C., Tanchoco, J.M.: Replacement decision based on productivity analysis—an alternative to the MAPI method. J. Manuf. Syst. **2**, 175–187 (1983)
Lutchman, R.: Sustainable asset management: linking assets, people, and processes for results, DEStech Publications, Inc. (2006)
Meyer, B.C.: Market obsolescence and strategic replacement models. Eng. Econ. **38**, 209–221 (1993)
Nair, S.K., Hopp, W.J.: A model for equipment replacement due to technological obsolescence. Eur. J. Oper. Res. **63**, 207–221 (1992)
Nguyen, T.K., Yeung, T.G., Castanier, B.: Optimal maintenance and replacement decisions under technological change with consideration of spare parts inventories. Int. J. Prod. Econ. **143**, 472–477 (2013)
Nurock, D., Porteous, C.: Methodology to determine the optimal re- placement age of mobile mining machines. In: Third International Platinum Conference on ‘Platinum Transformation’, The Southern African Institute of Mining and Metallurgy, 2008, pp. 297–306 (2008)
Okoh, C., Roy, R., Mehnen, J., Redding, L.: Overview of remaining useful life prediction techniques in through-life engineering services. Procedia CIRP **16**, 158–163 (2014)
Perrin, R.K.: Asset replacement principles. Am. J. Agric. Econ. **54**(1), 60–67 (1972). https://doi.org/10.2307/1237734
Peters, W.: Notes on the theory of replacement. Manchester Sch. **24**, 270–288 (1956). https://doi.org/10.1111/j.1467-9957.1956.tb00987.x
Porter, M., Kramer, M.: Strategy and society: the link between corporate social responsibility and competitive advantage. Harv. Bus. Rev. **84**, 78–92 (2006)
Ralf, B.: Sustainability in Mining. Environ. Manag. Health **3** (1992)
Regnier, E., Sharp, G., Tovey, C.: Replacement under ongoing technological progress. IIE Trans. **36**, 497–508 (2004)
Roh, H.-S., Lalwani, S.C., Naim, M.M.: Modelling a port logistics process using the structured analysis and design technique. Int. J. Logist. Res. Appl. **10**, 283–302 (2007)
Scarf, P.A., Hashem, M.: Characterization of optimal policy for capital replacement models. IMA J. Manag. Math. **13**, 261–271 (2002)
Scarf, P.A., Dwight, R., McCusker, A., Chan, A.: Asset replacement for an urban railway using a modified two-cycle replacement model. J. Oper. Res. Soc. **58**(9), 1123–1137 (2007)
Si, X.-S., Wang, W., Hu, C.-H., Zhou, D.-H.: Remaining useful life estimation–a review on the statistical data driven approaches. Eur. J. Oper. Res. **213**, 1–14 (2011)
Terborgh, G.: Some comments on the Dean-Smith article on the MAPI formula. J. Bus. **29**, 138–140 (1956)
Valicek, P., Fourie, F., Krafft, G., Sevenoaks, J.: ‘Optimization of mechanized mining layout within Anglo American Platinum’. In: Fifth International Platinum Conference, Sun City, South Africa, 17–21 September 2012 (2012)

Yatsenko, Y., Hritonenko, N.: Economic life replacement under improving technology. Int. J. Prod. Econ. **133**, 596–602 (2011)

Yusuf, K.O., Smith, N.J.: Modelling business processes in steel fabrication. Int. J. Proj. Manag. **14**, 367–371 (1996)

Predictive Life Cycle Forecasting: Innovative Decision-Making for Complex Asset Management in the Naval Environment

Tobias Lemerande(✉)

BMT, Hackney, SA, Australia
tlemerande@bmtdt.com.au

Abstract. Australia's warships and submarines are collectively the most complex, critical and expensive warfighting assets within Defence's inventory. Asset managers make decisions where beneficial short-term effects may cause unforeseen long-term repercussions leading to increased life cycle costs, decreased (or lost) capability and reduced operational availability that affect the operations and maintenance profile across each usage and upkeep cycle. Predictive life cycle forecasting provides an objective and empirical method to quantify budgetary requirements based on estimated future effects to operational readiness and seaworthiness. The life cycle forecast is a key component of each vessel's asset management plan and records the operations and maintenance profile across the asset's service life by establishing requirements for products and services needed to support the vessel within the prescribed asset management system. Predictive life cycle forecasting initially begins with establishing a baseline life cycle model that amalgamates contiguous operational running periods and scheduled maintenance activities across multiple usage and upkeep cycles to provide a time-phased representation that projects expected costs, operational availability and capability baselines from commissioning to disposal. Variable phases, states & modes provide the means to adjust model parameters to probabilistically characterise options available to asset managers when evaluating and assessing various scenario outcomes. An interactive model can provide asset managers with immediate feedback based on options explored within the model. Using each vessel's life cycle model, predictive life cycle forecasting can provide a consistent and logical method for systematically updating asset management plans. Robust and comprehensive predictive life cycle forecasting supports asset management decision-making to more accurately optimise warships' and submarines' availability, capability and affordability across the life cycle. As a fully scalable method, it can be applied to a single vessel, class of assets or to the collective fleet as a fundamental technique to support Fleet Life Cycle Management.

1 Introduction

Australia's national naval enterprise is a virtual organisation comprised of government personnel from the Royal Australian Navy (RAN) and Capability and Acquisition Sustainment Group (CASG) and commercial industry in the defence maritime sector.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 367–375, 2020.
https://doi.org/10.1007/978-3-030-48021-9_41

As a group, it designs, builds, sustains, operates and disposes of RAN vessels. Asset management has been identified as a core enterprise function. Department of Defence (Defence) sustainment policy mandates alignment to methods, practices and principles contained in ISO 55000, ISO 55001 and ISO 550002 (CASG 2017). The Fleet Life Cycle Management (FLCM) concept applies asset management to Australia's naval fleet through a framework that includes high-level Fleet Life Cycle Objectives (FLCOs) (Lemerande 2018). Enterprise decisions that affect individual RAN assets or the collective fleet should be made with the intention of meeting these high level strategic objectives, supported by lower-level asset management objectives, that reside at the forefront of any decision-making process. FLCM optimisation can be achieved through maximising the concurrent achievement of availability, capability and affordability across the life cycle of the collective naval fleet (Lemerande 2017). To optimise FLCM, the naval enterprise needs an objective and quantifiable method to improve decision-making capability uniformly throughout the enterprise. This paper describes how predictive life cycle forecasting (PLCF) can support enterprise stakeholders to make better asset management decisions during a vessel's service life to improve optimisation of availability, capability and affordability of the entire Australian fleet. Section 2 provides a brief overview of a ship's life cycle and highlights the symbiotic nature of operations and maintenance during its service life and discusses major considerations and key variables pertinent to PLCF. Section 3 discusses how modelling & simulation (M&S) techniques can be used to deliver PLCF to the naval enterprise. The conclusion summarises the benefits and opportunities PLCF can deliver.

2 Life Cycle of a Naval Vessel

Figure 1 graphically depicts the multiple phases of a ship's life cycle. During its service life, a naval vessel is either in a Maintenance Availability (MA) or an Operational Running Period (ORP) when assigned tasking and activities are conducted between consecutive MAs. A ship's service life schedule (SLS) is the time-phased plan of alternating ORPs and MAs between commissioning and decommissioning. MAs are conducted to ensure the ship can meet successive ORPs' availability and capability requirements. The SLS provides the time-phased constraints on which plans for predicting availability, capability and affordability can be based.

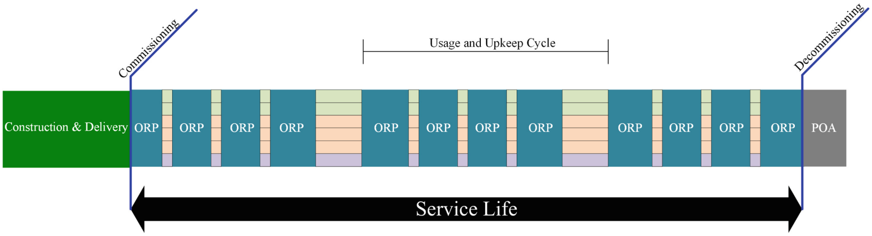


Fig. 1. Life cycle for a royal Australian navy ship

A Maintenance Availability Work Package (MAWP) contains all the work scheduled for accomplishment during a specific MA. There are six generic types of tasks contained in any MAWP, represented in Fig. 2. A correctly scoped, planned and executed MAWP should deliver a vessel at the end of the MA that can meet its inherent reliability and performance characteristics as intended throughout the next ORP.

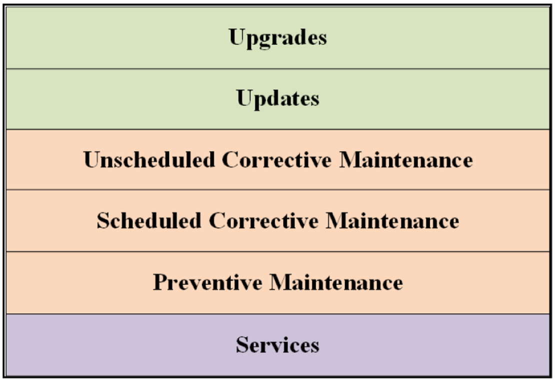


Fig. 2. Types of tasks within a maintenance availability work package

Upgrades and updates are two types of modernisation activities. Updates replace older equipment or parts with newer versions or components to improve reliability or prevent future logistical problems due to obsolescence or parts unavailability. Upgrades are modifications to ships or ship systems that will increase or improve military capability or functionality. Maintenance, or upkeep, can be either preventive or corrective and seeks to restore equipment to its intended operating condition. Preventive maintenance is predetermined work that is either time-based or respondent to conditional factors and aims to keep equipment and systems operating at or above designed performance levels. Corrective maintenance, which can be categorised as either scheduled or unscheduled, seeks to rectify deficiencies and restore equipment, components and systems back to minimum acceptable levels of performance in accordance with technical specifications. Scheduled corrective maintenance are known deficiencies that can be properly scoped and planned prior to the MA; unscheduled items are new deficiencies that were previously unknown or identified as growth from poorly scoped known deficiencies. Service tasks consist of support provided to the ship's crew and maintenance service providers that are necessary to accomplish work during a MA. Some common examples of services provided during a MA include: temporary electrical power; ventilation; compressed air; air conditioning; potable water; sewage and waste removal; rigging and lifting & handling; scaffolding; and activities required for dry docking.

During an ORP, a ship should be able to conduct one of four different types of activities: trials & material certification; military exercises; unit specific training or training within a larger task group or force; and missions to achieve operational objectives. A commissioned ship exists in one of several different states of material readiness. Figure 3 identifies generic states & modes a vessel will occupy across its

S tates I - Available, Capable, Deployable II - Available, Capable, Not Deployable III - Available, Not Capable, Not Deployable IV - Not Available, Not Capable, Not Deployable		At Sea		Alongside			Dry Dock	
		Op Area	In Transit	Home Port	Australian Port	Foreign Port	Home Port	Other
Operations	Mission	I	I	I	I	I		
	Exercise	II	II	II	II	II		
	Training	II	II	II	II			
	Trials/Certification	III	III	III	III			
Maint	Operational Maintenance Period			III	III	II		
	Scheduled Maintenance Availability			IV	IV		IV	IV
	Emergent Maintenance Period			IV	IV	IV	IV	IV

Fig. 3. States & modes of an in-service naval vessel

service life.

States I through IV have been identified through various combinations of a vessel being available, capable or deployable and paired with the appropriate geographic location and activity a ship must undertake. "Available" means a ship has the ability to safely put to sea and get underway under its own power. Not available means a ship is unsafe to go to sea or requires assistance to get underway and transit to another location. "Capable" means the ship's material condition enables it to safely conduct its assigned mission(s) or tasking. Not capable means the ship cannot fulfil these functions. "Deployable" means the operational commander has designated the ship for deployment and the vessel can execute the assigned mission. Thus, a naturally escalating hierarchy exists that must be preserved: a capable ship must be available; a deployable ship must be capable and thus is also deemed available. (A ship that is capable but not designated for deployment will not be deployable even though the material condition may support being in the deployable state.)

Modes are best characterised by the ship's generic location, employment and activity. Geography is the identified type of physical location; a ship is either at sea, moored alongside a pier, or docked ashore. A ship at sea is either in the designated operational area or transiting to or from it. A moored ship is berthed either overseas or in Australia where it can be in its home port or some other Australian port. The same applies to dry dock periods, depending on the location of the maintenance service provider. In each of these modes, a ship will exist in a single state at any given time. However a ship can be in more than one state for certain modes but can never be in two states simultaneously, just as it cannot be in two modes simultaneously.

Each combination of states & modes represents a unique condition for the given ship. States & modes are critical to PLCF because they correlate directly to "state changes" inside the model discussed in the next section. These unique combinations make up discrete conditions. Moving from one condition to another is signified by a change of states or "state change" that will be used in M&S software for PLCF.

3 Predictive Life Cycle Forecasting

Planning for an economically managed fleet includes forecasting costs associated with operating, maintaining and modernising shipboard equipment across the life cycle. Forecasts must be time-phased and specifically identify discrete activities within each ORP and MA. These activities are dynamic and, thus, are variables that when adjusted, will alter any future predictive plans. Forecasting should be based on the time-phased activities expected to occur in alternating ORPs and MAs. When considering optimisation of FLCM, these forecasts must evaluate the effects on availability, capability and affordability.

Forecasting techniques are classified as two general types: qualitative and quantitative. Quantitative forecasting techniques are objective and based on mathematical and statistical methods that relies on data (Al-Fares and Duffuaa 2009) and analytics. Qualitative forecasting methods rely on experts to apply expertise to make judgments based on intuition and expertise to produce informed estimates about the future (Goetschalckx et al. 2011). The FLCM concept requires an approach to utilising quantitative forecasting as much as possible to remove the subjectivity from the approach to life cycle management. Forecasting should cover the fleet's perpetual life cycle and be predictive in nature. Life cycle activities can be evaluated in three distinct categories – capability, availability, affordability – by amalgamating individual ship's data into a larger repository that assesses the entire fleet. This information can be represented in separate forecasted plans to which actual performance can be compared. The SLS establishes the time-phased plan for ORPs and MAs and tags specific states & modes to discrete periods within those activities. It provides the baseline upon which service life plans can be made.

3.1 Service Life Model

FLCM requires an enterprise architecture that accounts for every ship in the fleet and other support systems to provide a holistic and inclusive management environment (Lemerande 2018c). Within this architecture, each ship's service life model (SLM) would contain all pertinent information for ORPs and MAs and should be built using Systems Modeling Language (SysML) because the model can be partitioned into four distinct categories: structure, behaviour, requirements and parametric relationships (Grobhstein et al. 2007). SysML will enable a tailored method for focussing limited time, effort, money and human resources (Lane and Bohn 2012) on the most important and applicable aspects the SLM. SysML's functionality supports a hierarchical structure and facilitates the scalability and commonality needed to consolidate multiple SLMs into the fleet-wide model.

The model structure can establish a common ontology and define the basic structure and process architecture as it specifically applies to PLCF. Key parameters within the model that can be varied should include: dates and durations of each MA; the allocation of states & modes within each ORP; and the contents of each MAWP across the service life. These parameters within the SLM's structure will produce three separate plans that forecast availability, capability and affordability across the service life. All SLMs should contain the same basic structure as a way to facilitate scalability and easy amalgamation of all SLMs into the complete FLCM portfolio. This will enable forecasted plans to be combined into a composite representation of the entire fleet.

3.2 Forecasted Plans

RAN ships must be routinely modernised by installation of updates and upgrades throughout its service life. These updates and upgrades should be allocated to scheduled MAs and included in MAWPs. Figure 4 shows incremental capability increases resulting from updates and upgrades being installed during scheduled MAs. Capability increases can be quantified through linkages to a Functional Performance Specification (FPS) or other capability measurement function within the FLCM architecture. Just as updates and upgrades are assigned to MAWPs, all types of maintenance in MAWPs must be allocated to specific MAs to support availability predictions and forecasts. Cumulative availability increases during ORPs but flatlines during MAs, shown in Fig. 5, or whenever the required material readiness state cannot be met. Future availability is heavily dependent on the makeup and composition of each MAWP because it is highly reliant on the appropriate maintenance being conducted within each MA. Availability forecasts, collated from individual ship models, provide the data for incorporation into the consolidated fleetwide model. Cost estimates for discrete ORP and scheduled MAs can be quantified based on the expected tasking and length of scheduled sustainment periods. Maintenance, modernisation (i.e. updates and upgrades) and MA services are crucial elements within life cycle planning; they undoubtedly affect financial programming and budgeting efforts regardless of the expected service life. ORPs are often prorated to account for fuel, ammunition, victuals and other recurring actions required to support maritime operations. For MAs, each MAWP can be estimated based on the known services, expected maintenance and planned updates and upgrades. Estimates will yield overall summations that are based on labour and materiel. Total material costs and labour rates applied to expected (or allocated) efforts will produce a cost profile across a ship's service life similar to that depicted in Fig. 6. Service life plans' data will provide useful, pertinent and necessary information for the PLCF concept.

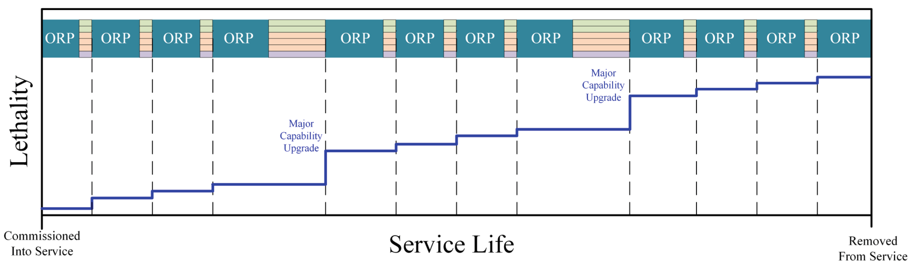


Fig. 4. Forecasted service life capability plan

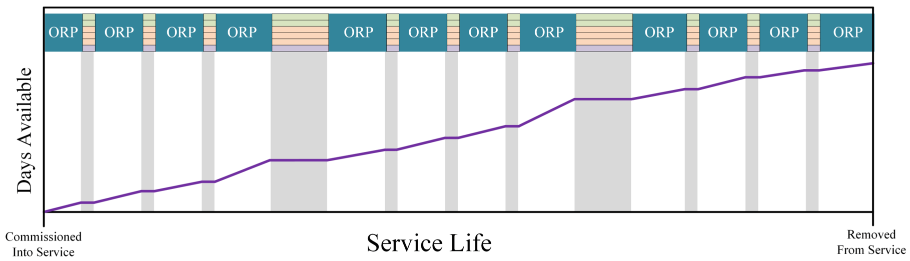


Fig. 5. Forecasted service life availability plan

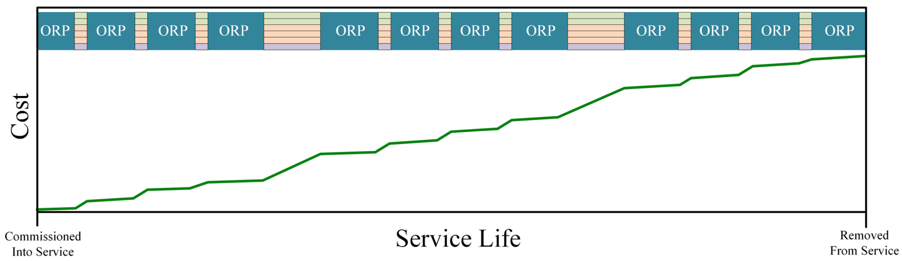


Fig. 6. Forecasted service life affordability plan

3.3 Decision Support from Discrete Event Simulation

Decision support through algorithms or software tools help stakeholders and decision makers consider implications and consequences of options and courses of action. Simulation is an effective technical means for system(s) optimisation. Decision support tools are designed to improve decision-making processes for complex systems and are widely used to assist decision-makers who must consider wide ranging areas. Discrete Event Simulation (DES) simulates operations by stepping through time and skipping periods where no changes occur. This method is desirable in that implementation is easy, execution time is relatively short and the environment is flexible (Griendling and Mavris 2011). DES works well when states & modes are clearly defined because these become "state changes" and can easily be modelled. For a dynamic system characterised by complexity and uncertainty, as would be encountered by dozens of naval ships across 30 or more years, DES offers a powerful way to gain insight and knowledge about the system. In DES, abstract system models use a continuous but bounded time base where only a finite number of relevant events occur. These events cause state changes within the system which are then evaluated within the model to determine the effects on the overall system. DES is the method by which stakeholders and modellers can collaborate to explore different courses of action through "optioneering" by changing key variables within SLM(s) and executing simulation on the consolidated fleet model and observing the results for each forecasted plan.

3.4 Optimisation Through Optioneering

MAs are key drivers for parts, material and services which account for a significant portion of the Life Cycle Cost (LCC) attributed to a ship's the service life. Accurately forecasting MAs and all the elements that contribute to effective maintenance periods are critical to FLCM. MAWPs are susceptible to significant change and can be the source of significant disruption to a forecasted plan. ORPs can also be extremely dynamic. This aspect of the SLM has many states & modes that can be varied within simulations. Moreover, unexpected state changes due to emergent operational requirements or unplanned repairs can wreak havoc on a forecasted plan. Costs predicted for each MA and ORP, when tallied across a ship's service life, can produce a cumulative summary. However, when contents of MAWPs change or operational schedules change, the associated costs will also be affected. The dynamic nature of both ORPs and MAs necessitates the SLM be flexible, simple in its design and robust to handle significant changes within the consolidated fleet model. It must also give modellers the ability to change the environment to allow exploratory simulation that produces objective and quantitative results that can be easily compared.

In PLCF, optioneering is the method by which this dynamic nature can be managed. Optioneering allows stakeholders and decision-makers to stimulate imagination, visualise possibilities and quantify the magnitude of change while revealing alternatives and trade-offs associated with different options. It seeks to use highend computing capability to reduce latency, increase integration amongst contributing factors and deliver quantifiable evaluation of various alternatives (Gerber and Flager 2011). Furthermore, it also allows stakeholders to explore various and more complex solution possibilities through simulations and optimisation methods like DES without relegating this work solely to engineers or designers (Gerber et al. 2012). Using the tools and models available in the PLCF concept, optioneering allows stakeholders and decision-makers to use M&S to explore multiple scenarios and observe how the results affect the fleet's overall availability, capability and affordability in the short-, medium- and long-terms.

4 Conclusion

PLCF, as described in this paper, can provide a scalable decision support mechanism that gives naval enterprise stakeholders an ability to explore numerous scenarios through dynamic M&S techniques in order to validate or refute different potential options for a single ship and observe the effect at the consolidated fleet level. PLCF, if developed and implemented using appropriate M&S, will provide naval enterprise stakeholders with improved decision-making capabilities that can better support optimisation of availability, capability and affordability throughout the naval fleet in the coming decades.

References

Al-Fares, H.K., Duffuaa, S.O.: Maintenance forecasting and capacity planning. In: Handbook of Maintenance Management and Engineering, pp. 157–190. Springer, London (2009)

Buede, D.M., Miller, W.D.: The Engineering Design of Systems: Models and Methods. Wiley, Hoboken (2016)

Capability and Acquisition Sustainment Group: Functional Policy (PM) 001: Sustainment Management in Capability and Acquisition Sustainment Group Version 2.0., Commonwealth of Australia, Canberra (2017). Accessed 11 Aug 2017

Gerber, D.J., Flager, F.: Teaching design optioneering: a method for multidisciplinary design optimization. In: Computing in Civil Engineering 2011, pp. 883–890 (2011)

Gerber, D.J., Lin, S.H.E., Pan, B.P., Solmaz, A.S.: Design optioneering: multi-disciplinary design optimization through parameterization, domain integration and automation of a genetic algorithm. In: Proceedings of the 2012 Symposium on Simulation for Architecture and Urban Design, p. 11. Society for Computer Simulation International, March 2012

Goetschalckx, M.: Supply Chain Engineering, vol. 161. Springer (2011)

Griendling, K., Mavris, D.N.: Development of a Dodaf-based executable architecting approach to analyze system-of-systems alternatives. In: 2011 IEEE Aerospace Conference, pp. 1–15. IEEE, March 2011

Grobshtein, Y., Perelman, V., Safra, E., Dori, D.: Systems modeling languages: OPM versus SysML. In: International Conference on Systems Engineering and Modeling, 2007. ICSEM 2007, pp. 102–109. IEEE, March 2007

Lane, J.A., Bohn, T.: Using SysML modeling to understand and evolve systems of systems. Syst. Eng. **16**(1), 87–98 (2012)

Lemerande, T.J.: Optimising availability, capability and affordability across the fleet: a total life cycle management approach to improving seaworthiness. In: Proceedings of 12th Annual World Congress on Engineering Asset Management, Brisbane, 2–4 August 2017 (2017)

Lemerande, T.J.: Leading the naval asset management effort with a framework for fleet life cycle management. In: E-proceedings of AMPEAK 2018, Hobart, Tasmania, 16–18 April 2018 (2018)

Lemerande, T.J.: A system of systems approach to developing a digital asset management environment for the naval enterprise. In: Systems Engineering Test and Evaluation Conference 2018, International Council on Systems Engineering (INCOSE), Sydney, 30 April–2 May 2018 (2018)

Condition and Performance Monitoring of Emergency Shutdown Systems: Data Visualization and Analysis for Decision Support

Pengyu Zhu[1](✉), Jayantha P. Liyanage[1], and Simon Jeeves[2]

[1] University of Stavanger, Stavanger, Norway
{pengyu.zhu,j.p.liyanage}@uis.no
[2] MRC Global, Bergan, Norway
simon.jeeves@mrcglobal.com

Abstract. The critical role of emergency shutdown (ESD) systems is widely acknowledged in the oil and gas industry. Monitoring the condition and performance of ESD systems is known to have potential benefits in further improving cost-effective production, while ensuring safety availability. This paper takes a special perspective on data visualization and analysis for decision support regarding ESD systems. The paper identifies the needs and expectations of core stakeholders and explores how to improve the user interface of a condition and performance monitoring system. The proposed solution has a focus on obtaining the most critical data and information in a visualized and integrated interface and on accessing automatic trending and failure analyses to assist decision-making.

1 Introduction

Emergency shutdown (ESD) systems are widely used in the process industry, which involves large volumes of flow during operations. ESD systems are mainly designed to shut down production flow in the case of emergencies and to limit the escalation of hazardous events before and when they occur (NORSOK 2008). A failure of ESD systems on demand can be a great threat to production safety and availability.

In the oil & gas (O&G) industry on the Norwegian continental shelf (NCS), ESD systems have been managed, using corrective maintenance strategies in more traditional terms. However, several critical ESD systems, such as downhole safety valve systems (DHSV), have recorded an unsatisfactory average failure frequency over the last 15 years (Petroleum Safety Authority Norway 2017). The application of condition and performance monitoring systems to ESD systems has been discussed as being helpful in early failure/degradation detection, automatic data collection, increased safety failure fraction, decreased probability of failure on demand (PFD), and so on (Greenlees and Hale 2012; Juvik et al. 2002; Lundteigen and Rausand 2007; Zhu et al. 2019).

However, in general, there has not been a major focus on the advanced use of condition and performance monitoring applications on ESD systems, as well as on

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 376–386, 2020.
https://doi.org/10.1007/978-3-030-48021-9_42

other asset-critical on-demand systems, due to specific operational practices. There is a need to raise the awareness, among both service providers and asset owners, of the exploration of new practices, especially focusing on collecting and analyzing relevant data and improving decision support regarding ESD systems.

2 Research Methodology

This paper is a part of a research study, conducted in collaboration with the oil & gas industry, in relation to operation and maintenance decision support of ESD systems. Data were collected based on 21 surveys conducted with both current and potential users of the system, technical service providers, researchers, and safety authorities during a period of over two years. Data and materials were also collected through interviews with experts, who had extensive experience and knowledge of the studied system. A specific condition monitoring application technology for ESD systems, called ValveWatchTM, provided the basis for this paper to further explore the potential use of a visualization and analytical interface to improve the current practice.

3 ValveWatch^{TM} Technology to Monitor Condition and Performance of ESD Systems

ValveWatchTM is an automated online condition and performance monitoring system for critical valve systems, including ESD systems (Valvewatch 2017). The original development of the monitoring system dated back to 1997, as a response to the Piper Alpha disaster (Juvik et al. 2002). The ValveWatch^{TM} system was the first of its kind for monitoring critical valve systems in the O&G industry. The system collects, visualizes, and analyzes real-time condition data from multiple sources, to support decision-making. It allows users from different disciplines and geographical locations to collaborate seamlessly and simultaneously.

When entering the ValveWatch^{TM} interface, users are shown a visualized real- time condition and performance overview of all monitored ESD systems, as depicted in Fig. 1. The conditions of systems are clearly marked and presented with different colors, representing the health conditions of systems, giving users an insight into the overall picture.

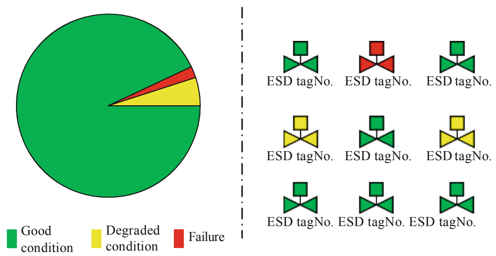


Fig. 1. ValveWatch^{TM} interface for monitoring ESD systems

A data directory page is also provided, if users want to look into details of a specific system's condition. Graph previews of real-time sensor signature curves, such as actuator pressure readings, are plotted and displayed for diagnosis and prognosis purposes, as shown in Fig. 2. In practice, pattern recognition analysis and trending analysis are used to carry out diagnostic and prognostic analyses (Hale 2003; Juvik et al. 2002; Zhu et al. 2019). The implementation of such analyses, however, largely depends on experts' ability to interpret the data. Extensive knowledge and experience of the mechanical and electronic principles of the ESD system, data analysis, and operation and maintenance basics of the system are greatly needed for such tasks.

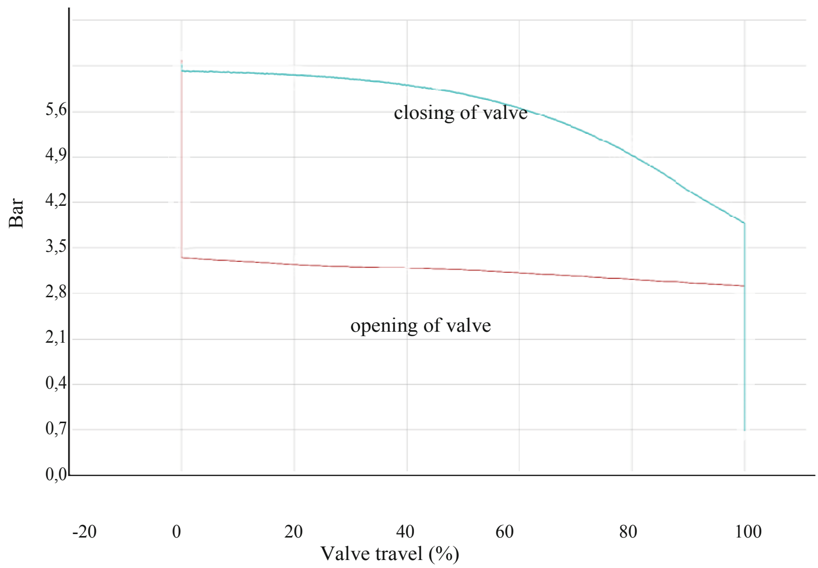


Fig. 2. Actuator pressure sensor reading during ESD system operation

4 Identification of Users' Needs of ValveWatch™

In practice, ValveWatch™ users were divided into three main groups: offshore users (asset owner), onshore users (asset owner), and onshore users (monitoring service provider). It is important that each group of users has access to the right data and information, in order to make informed decisions. In addition to condition data, equipment data, operation & maintenance data, failure history, reliability data (failure rate), and failure mechanism data are also critical contextual data for all users, regarding ESD systems (ISO 14224 2016; OREDA 2015). Different users' roles, responsibilities, and requirements for data & information were identified, based on a series of surveys and discussions, as shown in Table 1.

It was noted that, as a common practice, data were stored on different information management platforms and communicated in inconsistent forms. In addition to this,

Table 1. Roles, responsibilities and data needs of user groups

Users	Onshore user (service provider)	Onshore user (asset owner)	Offshore user (asset owner)
Role	- Collection & analysis of condition data - On-site service - Periodic valve performance report	- Registration of maintenance notification - Maintenance planning and optimization	- Registration of maintenance notification - Implementation of inspection, repair and replacement
Responsibility	- Failures are reported upon detection - Identification of failure causes	- Prioritization of work orders - Ensure safety level is fulfilled	- Failures are reported upon inspection - Work orders are carried out
Data needs	- Equipment data - Condition data - Failure data - Reliability data - Failure mechanism - Key performance indicators (KPI)	- Equipment data - Condition data - O&M data - Failure data - Reliability data - Failure mechanism - System KPIs	- Equipment data - Operation & maintenance (O&M) data - Failure data - System KPIs

users often found it time-consuming and challenging to perform holistic analysis of available data. There was a need for an improved, insightful data management and visualization solution, which can provide the relevant basis for in- depth analyses, to support informed decision-making. Thus, in the next section, a suitable data visualization and analysis solution is proposed, to enhance the decision support capability.

5 Developing a Data Visualization and Analysis Solution

Developing an effective data visualization and analysis solution for ValveWatch was a key deliverable of the research study. With the extended availability of contextual data from online condition monitoring programs, this study explored a number of technical aspects for improved decision support regarding ESD systems.

5.1 Data Modeling

One major purpose of having online condition and performance monitoring systems is to make early detection of failure/degradations possible and practical, and to eventually improve the safety levels of ESD systems. The demands of different users for data, as summarized in Sect. 4, need to be fulfilled. The study investigated the need for such a logical data model to holistically describe the scope and data interfaces between the technical system, ValveWatch™, SAP or computational maintenance management systems (CMMS), and users. A brief logical data model, built using crow's foot notation, is shown in Fig. 3.

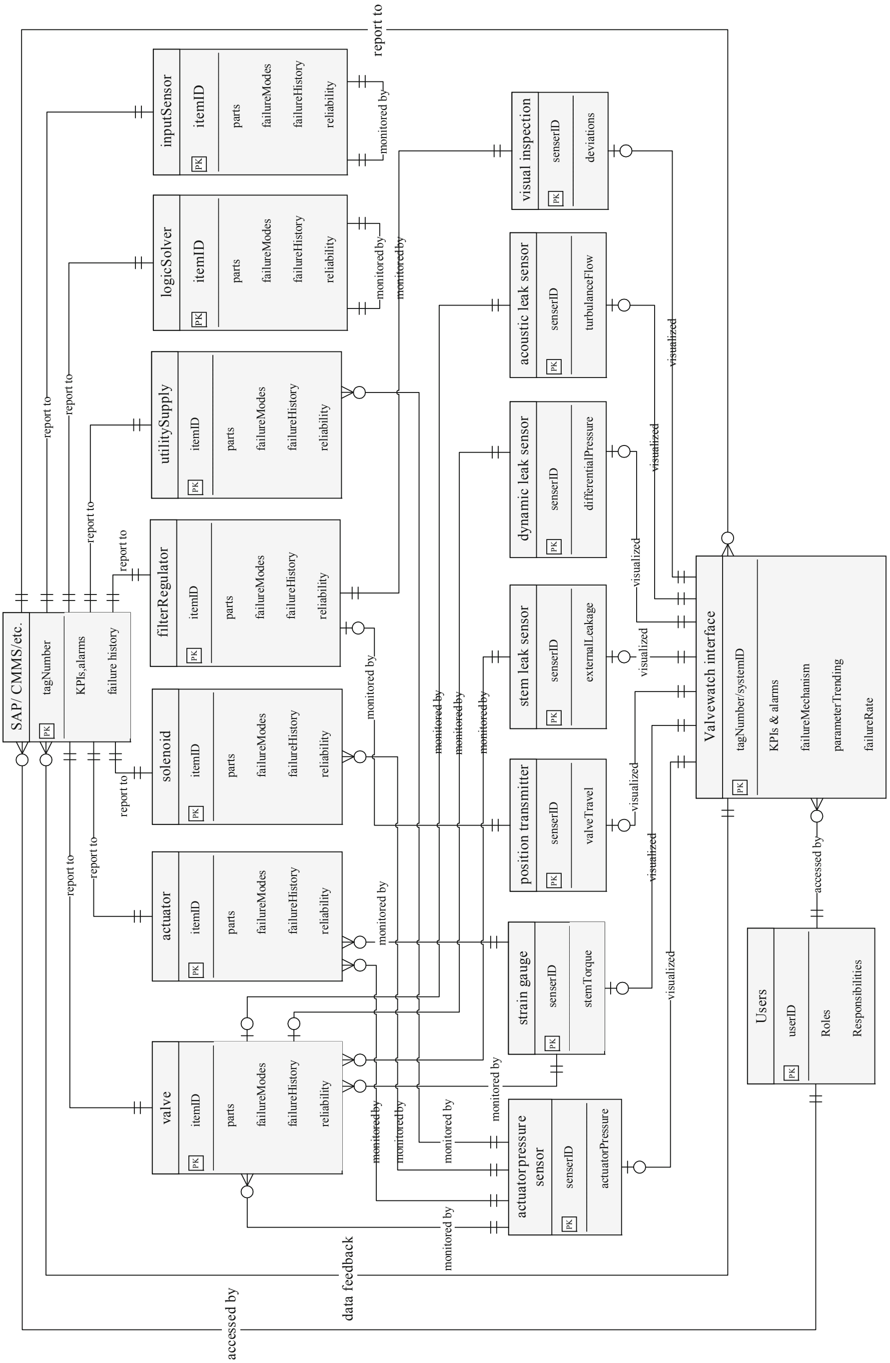


Fig. 3. Data model of online monitoring system using crow's foot notation

5.2 A Holistic Overview of ESD Systems in the Operating Phase

In practice, the results of most analyses are communicated through technical reports or offshore visits. The level of understanding of the diagnostic and prognostic analyses and the complexities of analyses' reports can easily make it difficult for them to be integrated and described in the current interface. Based on this study, an intuitive interface is suggested, as shown in Fig. 4, that embeds concise but holistic data & information and provides a practical visualization and analysis basis for decision support.

In practice, various ESD systems are managed, tested and evaluated in groups, such as riser ESD system and downhole safety valve system (DHSV). It is natural for all monitored ESD systems to also be displayed in groups in the overview interface. Hauge and Lundteigen (2008) explained the need for and importance of updating the PFD and safety integrity level (SIL) of different systems in the operating phase and explained how function test intervals can be determined, based on failure data and operation experiences. In the new interface, these key maintenance parameters are displayed when users click on a certain group of ESD systems. The quantitative values are calculated using embedded analytic engines, based on failure data that are automatically extracted from SAP or CMMS. Failure histories related to a specific group of ESD systems are displayed by activating the 'Failure history' function. Users can also click on a specific ESD system node to call up the system synopsis window, which shows brief information and the condition of the system, as depicted in Fig. 4. Some advanced analyses, such as safety factor trending and failure analysis, can be accessed by pressing the respective button shown in the window.

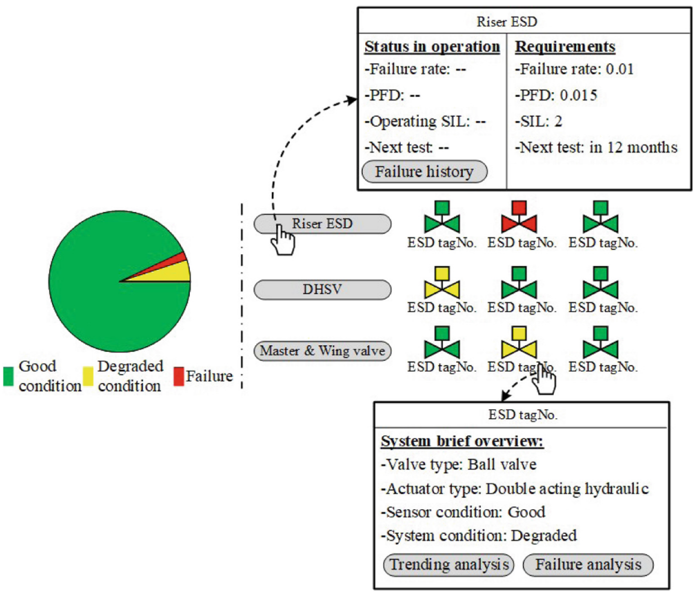


Fig. 4. Enhanced overview of ESD system groups

5.3 Trending of System Performance

System performance is measured continuously with ValveWatch^{TM}. However, for most users, data collected at key valve positions, such as valve break-to-close point, are of the greatest interest. The sampling process is achieved with the use of a position transmitter. The reduction in data that need to be processed and analyzed is helpful in reducing both workload and competence requirements. There are several ways to visualize the data sets, to reveal system condition and performance.

The monitoring of safety factors or function margins of actuated valve systems has been applied in the nuclear industry since the 1980s (U.S. Nuclear Regulatory Commission 1989; Hale 2003). On the NCS, ValveWatch^{TM} introduced the concept of measuring safety factor degradations regarding critical valve systems in the O&G industry (Juvik et al. 2002). For an ESD system, operability is a major concern. The ratio of actuator force to friction is thus defined as a key safety factor. Readers should bear in mind that this safety factor needs to be higher than 1 in order to operate an ESD valve. Instead of plotting detailed sensor readings, as shown in Fig. 2, the trending of safety factors at the most critical valve positions gives a clear picture of system operability and performance and is suggested, as depicted in Fig. 5.

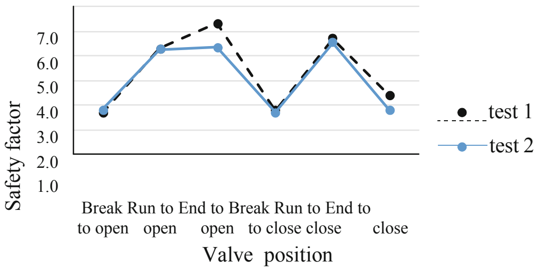


Fig. 5. Trending of safety factor of ESD system, an example

A decrease in a safety factor, if still above 1, reveals a developing failure of the system. It raises the need for further investigation, when the degree of degradation becomes unacceptable, based on predefined criteria. However, not all degradations can be revealed by simply trending the safety factor. Increased friction, due, for example, to debris, may be overcome by an oversized actuator. Detailed performance indicators are sometimes needed to reveal root causes.

The trending of key performance indicators (KPI) of the ESD systems is suggested, as it provides extra insight into the detection and analysis of early failures/degradations. In traditional terms, KPIs regarding ESD systems were often defined to measure the overall system performance, such as valve travel time and internal valve leakage rate at closed position (Petroleum Safety Authority Norway et al. 2016; Norwegian Oil & Gas Association 2017). With the increased detectability, more detailed KPIs can be and should be defined, such as breakout torque to move a valve. Historical measurements of these KPIs are trended, and degradations can be revealed that are otherwise hidden by simply trending the safety factor. At the 'break-to-close' valve position, for example, the safety factor of the system remains stable in two tests, according to Fig. 5, but increases in friction and actuator force are revealed by trending maintenance KPIs, as shown in Fig. 6. The monitoring and trending of safety factors and KPIs of ESD systems is a major step forward to predictive maintenance.

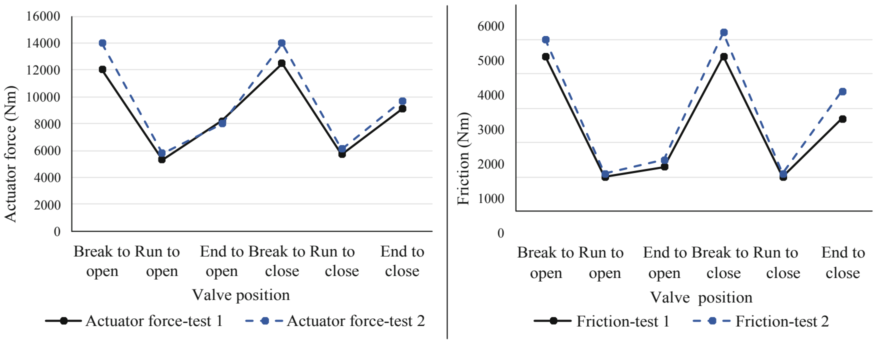


Fig. 6. Trending of maintenance KPIs of an ESD system, an example

5.4 Embedding Failure Analysis Ability

Failure analyses are normally carried out by dedicated diagnosis experts. They pick out slight variations on sensor signature curves that have been caused by different failure mechanisms. Various tools such as fault tree analysis (FTA) and failure modes and effect analysis (FMEA) are also used to assist analyses. The authors believe that such a process can be largely captured and modeled into a digital interface. The key is to establish the links between sensor detectability and failure mechanism. The detection of early degradations/failures of ESD systems has been explained and discussed from a failure progression perspective (Hale 2003; Juvik et al. 2002; Zhu et al. 2019). A typical mapping approach is illustrated in Fig. 7.

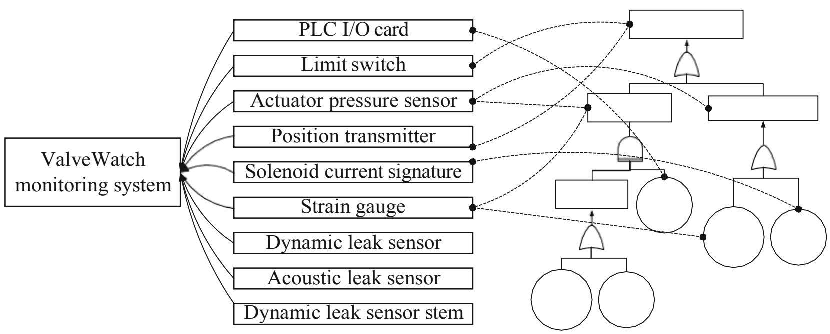


Fig. 7. Inter-relationships between sensor readings and failure logic

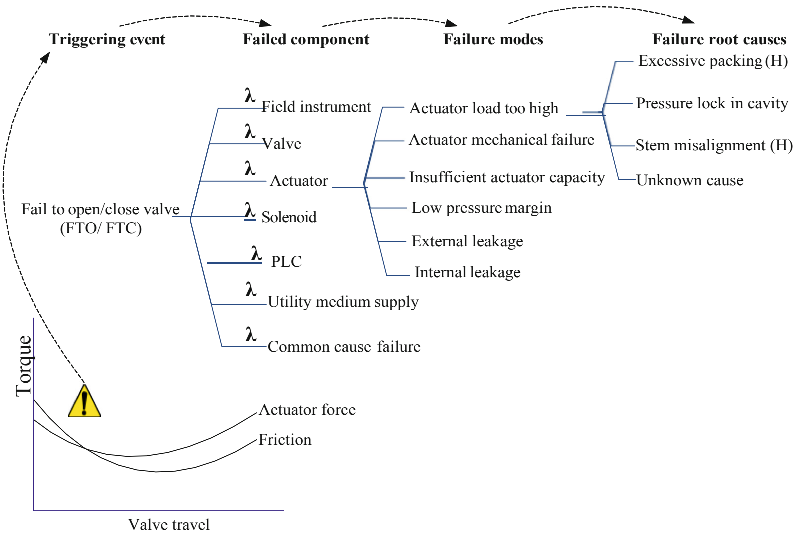


Fig. 8. Assistance for failure analysis

Expert judgements are valuable during failure analyses, but the authors aim to introduce an approach that allows anyone to use the system in an intuitive way. Based on mapping results, an elimination process is followed so that users can cross off impossible failure causes and consequences and identify the mostly likely failures, based on reliability data and failure logic, as shown in Fig. 8, where human-related failures are marked with '(H)'.

The approach aims to find a balance between uncertainty and efficiency. In this case, uncertainty refers to the identification of the exact failure cause. The suggested approach provides an extended list instead of a certain failure cause. The list is recommended according to the failure propagation logic and prioritized based on the likelihood of occurrence, which is measured by continuously updated failure rate λ, with failure data from SAP/CMMS systems. The list can be used as a checklist for the maintenance team. The process provides users with a structured guide to localize failure and identify failure causes. During the process, a 'real' reliability database, based on the corporation's own practices, will also be established in the long term, which will improve the precision of prediction, from a statistical perspective.

6 Conclusion

Current practices related to condition and performance monitoring of ESD systems have been developed, based on operational experience in the O&G industry and nuclear industry practices. Not all the needs of important users can be fulfilled by current practices; hence, the study explored how the situation can be improved, through data, analysis, and visualization.

The proposed solution leverages the value of data, especially real-time condition data, in assisting decision-making processes. The paper tries to avoid suggesting sophisticated diagnosis and prognosis guidelines or processes that can only be used by limited user groups. Efforts were made to keep the suggested approach simple and logical, to view the overall condition and performance of all monitored systems and access various analyses' results from the integrated interface for decision support. The proposed solution can be used to implement routine checks, to carry out retrospective failure analysis, and to plan preventive and proactive maintenance regarding ESD systems.

In practical terms, it is not always the case that the most statistically likely failure cause is the exact failure cause in low-demand systems. In many cases, the use of expert judgements combined with the suggested solution can be more efficient. A future study may explore how to build a machine-learning engine that can complement experts' judgements.

References

Greenlees, R., Hale, S.: Optimizing valve maintenance and testing utilizing acoustic emissions (AE) technology. In: 2012 20th International Conference on Nuclear Engineering and the ASME 2012 Power Conference (2012)

Hale, S.: ICONE11-36068 Improvements in valve reliability due to implementation of effective condition monitoring programs. Paper presented at the Proceedings of the International Conference on Nuclear Engineering (2003)

Hauge, S., Lundteigen, M.A.: Guidelines for follow-up of Safety Instrumented Systems (SIS) in the operating phase. SINTEF, Trondheim (2008)

ISO 14224: ISO/TC 67, ISO 14224:2016, Petroleum, petrochemical and natural gas industries – Collection and exchange of reliability and maintenance data for equipment. 3rd edn., International Organization for Standardization (2016)

Juvik, T., Hermansen, T., Carr, R., Hale, S.: Online valve monitoring systems used on off-shore platforms in the North Sea. Paper presented at the ASME 2002 21st International Conference on Offshore Mechanics and Arctic Engineering (2002)

Lundteigen, M.A., Rausand, M.: The effect of partial stroke testing on the reliability of safety valves. In: ESREL 2007 (2007)

NORSOK: S-001. Technical safety. Stavanger (2008)

Norwegian Oil & Gas Association: 070-Application of IEC 61508 and IEC 61511 in the Norwegian Petroleum Industry (2017)

OREDA: OREDA Handbook. Offshore and onshore reliability data, Topside equipment, vol. 1, SINTEF, NTNU (2015)

Petroleum Safety Authority Norway: Trends in risk level in the petroleum activity (RNNP), Main report 2017, Stavanger (2017)

Petroleum Safety Authority Norway, Norwegian Environment Agency, Norwegian Directorate of Health, Norwegian Food Safety Authority, Guidelines regarding the activities regulations (2016)

U.S. Nuclear Regulatory Commission: Safety-related motor-operated valve testing and surveillance. Generic Lett. **89**(10), 89 (1989)

Valvewatch: Sensors of valve monitoring system (2017). http://www.mrcglobal.com/Global-Region/Products/ValveWatch/Sensors

Zhu, P., Liyanage, J.P., Jeeves, S.: Data-driven failure analysis of emergency shutdown systems in oil and gas industry: Evaluation of detectability from failure progression perspective. J. Qual. Maint. Eng. **26**(1) (2019)

Hybrid Modelling for Lifetime Prediction

Fikri Hafid[1,3](✉), Maxime Gueguin[2], Vincent Laurent[2], Mathilde Mougeot[4], Nicolas Vayatis[3], Christine Yang[1], and Jean-Michel Ghidaglia[3]

[1] R&D Department, Réseaux de Transport d'Electricité (RTE), Paris, France
fikri.hafid@rte-france.com
[2] Scientific Computing Branch, Eurobios, Cachan, France
[3] Centre de Mathématiques et de leurs Applications (CMLA), ENS Paris-Saclay, Cachan, France
[4] Laboratoire de Probabilités, Statistiques et Modélisations (LPSM), Paris Diderot University, Paris, France

Abstract. Asset management is a major challenge for RTE (Réseau de Transport d'Electricité), the French TSO (Transmission System Operator) since electrical grid component maintenance and renewal represent important economic issues. Thus, improved assessment of their lifetime could lead RTE to substantial savings while assuring a high quality of service. To tackle this issue, RTE has gathered large amounts of data (monitoring, material failures, asset database …) and at the same time has tried to understand the phenomena of equipment ageing that could eventually lead to failure. RTE has chosen to take advantage of these two information sources when developing a hybrid model. This model combines data analysis of the gathered data sets and numerical simulation applied to the physical modelling of the phenomena involved in the accelerated ageing of equipment. To illustrate this approach, we are going to present a complete run of the hybrid computation to estimate lifespan of overhead line conductors due to aeolian vibrations. Aeolian vibrations, which are a type of vortex induced vibration, are a well-known phenomenon in the overhead lines (OHL) community. These vibrations are caused by the interaction between conductors and low-speed winds (velocities ranging from 1 m/s to 7 m/s). They induce one of the main causes of conductors damage: fretting fatigue in the clamp/conductor system. In this hybrid computation, damage due to fretting fatigue is evaluated using a mechanical multiscale approach where representative loading and loading scenarios are obtained using machine learning techniques.

1 Introduction

RTE is charged with maintaining the transmission of high and very high voltage electricity in France. Overhead line conductors represent an important part of the technologies employed by RTE, since they total more than 100,000 kilometres in length. The sound management of these assets, which is mainly the replacement of the electrical conductors that make up the lines, is therefore a major issue for the company.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 387–395, 2020.
https://doi.org/10.1007/978-3-030-48021-9_43

At this point in time, RTE has programmed the replacement of these conductors when they reach 85 years of service, regardless of their position in France and therefore also their condition. In order to replace this subjective threshold, RTE has launched several R&D projects aimed at better understanding the phenomena related to the ageing of overhead lines.

Conductors are heterogeneous solids, since they are made of stranded steel, aluminium or aluminium alloy wires. These conductors initially undergo a mechanical tension, then they are fixed to pylons by means of suspension systems. They are then electrically energized and are subject to complex external stresses (wind, snow, solar radiation) (Fig. 1).

Fig. 1. Overhead line conductor (Adto Group)

In order to better understand the conditions that the overhead lines have been subject to, it is necessary to know their context in the French power system. This information is available from RTE in the form of several databases that can be grouped into different categories:

- asset data (e.g. type of conductor, distance between pylons),
- operating data (i.e. electric power transmitted),
- meteorological data (e.g. ambient temperature, average wind speed),
- event data (e.g. cut wires, lightning),
- maintenance data (i.e. repair/replacement operations),
- experimental testing data (e.g. mechanical characterization of wires).

2 Hybrid Modelling

2.1 Description of the Approach

A maintenance method that is widely used in all industrial areas consists of automatically and continuously monitoring equipment in operation in order to optimize strategies that prevent breakdowns and incidents. However, in the case of overhead lines the means of making these type of measurements are limited. More-over, to the best of our knowledge, no solution dedicated to the evaluation of conductor damage exists.

In view of the maturity of existing mechanical models dedicated to overhead line conductors and the available data describing the environmental conditions and use of the overhead lines, it is possible to propose an original approach based on a so-called hybrid model.

The hybridity of this model is based on the use of two distinct components, namely statistical analysis & mechanical modelling that interact in order to take advantage of the advances of each of these components (Fig. 2).

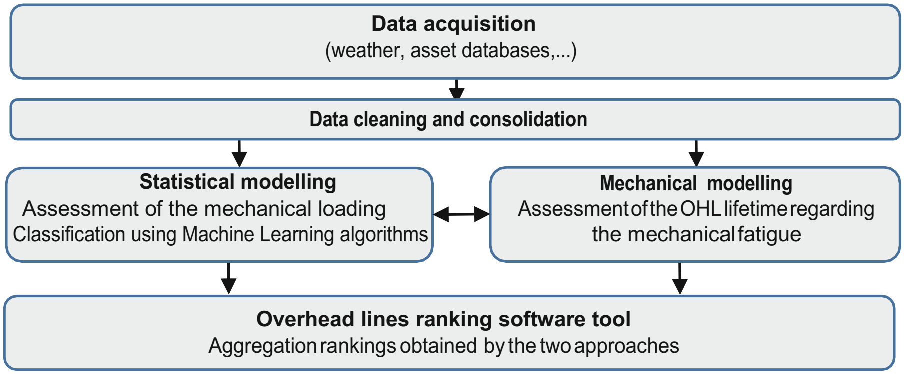


Fig. 2. Hybrid approach

2.2 Mechanical Modelling

Aeolian vibrations [1, 2] caused by low speed winds are one of the phenomena that lead to premature ageing of conductors by inducing fretting fatigue [3, 4]. The damage caused by this fatigue phenomenon is located in the contact areas between the conductor wires. Thus, it is necessary to evaluate the effect of the wind on these contacts. This last fact and the high heterogeneity of overhead line conductors have led us to propose the use of a multi-scale approach to analyse the physical phenomena associated with this problem. In this context, three modelling scales naturally appear, namely:

- the overhead line span (distance between pylons, hundreds of meters): account for external solicitations,
- a short section of the conductor (tens of centimetres): including all interactions between wires,
- contact area between wires (millimetres): zone where the damage usually appears.

Each scale can be associated with one or more numerical models, adapted to the studied physics (Fig. 3).

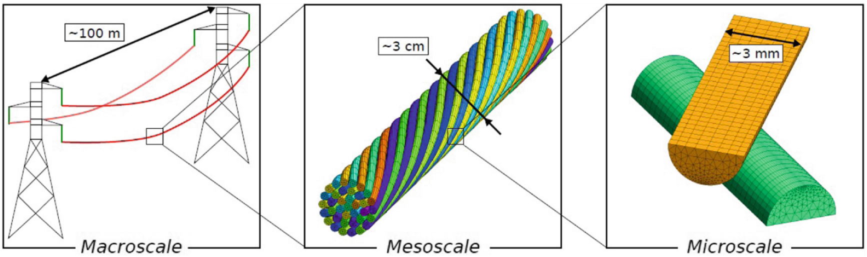


Fig. 3. Multi-scale mechanical model

2.2.1 Macroscale Model

The geometric characteristics of the span and the mechanical characteristics of the conductors induce large displacements (e.g. sag of several metres under dead weight).

The interaction phenomena between the wind and conductor takes several minutes to fully develop, but vibrations are produced that can be critical in the long term. The chosen macroscale model is therefore based on beam-type elements for large displacements, where a degree of simplification is made to reduce simulation times. Since the interactions between the internal wires are not directly taken into account, the simplified models must integrate a homogenized behaviour that is as close as possible to experimental observations (Fig. 4).

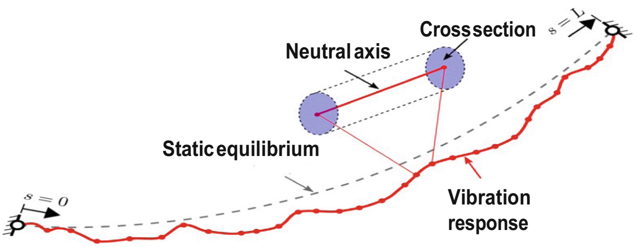


Fig. 4. Macroscale beam elements based model

2.2.2 Mesoscale Model

An overhead line conductor is made up of several tens of stranded wires in several layers, meaning that there are contacts between the wires. The contacts can be divided into two categories:

- those within the same layer: line contact,
- those between wires belonging to two successive layers: punctiform contact, elliptical in shape.

As an illustration, for a conductor composed of 37 wires arranged on 4 layers, a 22 centimetre portion of the conductor has 36 linear contacts and 294 punctiform contacts (Fig. 5).

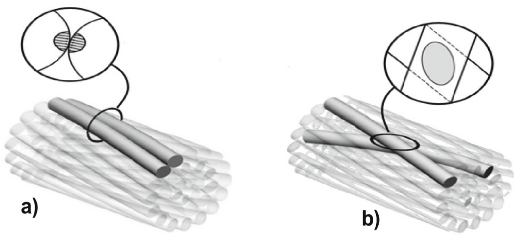


Fig. 5. a) Linear contact, b) punctiform contact (elliptic shape)

With the current numerical methods, it is not conceivable to deal with complex solicitations using a fully 3D finite element model that can directly consider all the contacts. So we built a simplified wired model, where:

- each conductor wire is modelled by a set of beam elements,
- the interactions between wires are represented by discrete contact elements, arranged at analytically calculated positions.

The simplified wire model induces an improvement in computational performance (reduction of the number of degrees of freedom, faster convergence), but generates a loss of local information (e.g. wire's internal stress field) (Fig. 6).

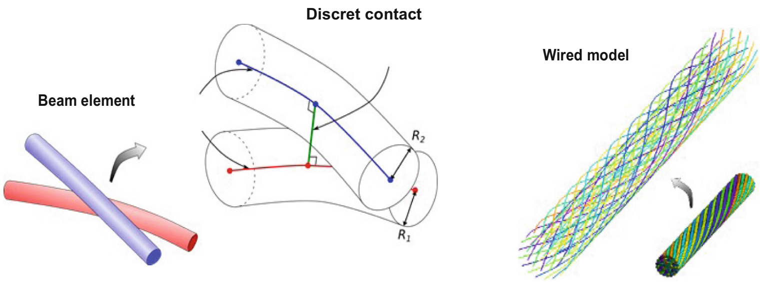


Fig. 6. Mesoscale simplified model

2.2.3 Microscale Model

A 3D model is necessary to capture the phenomena appearing at the scale of a contact between two wires. It is indeed at this scale that crack initiation can occur and lead to breakage of the conductor when a sufficient number of wires are broken. The evaluation of such fatigue damage, which appears after millions of loading cycles, requires the knowledge of mechanical quantities at a very local level. However, the computations for such a model can be very time consuming, so it is thus necessary to use a strategy to evaluate the damage after a large number of loading cycles. A good strategy is to use an adequate fatigue criterion applied to a calculated stress field [5, 6]. An example of this strategy applied to our specific case is completely described in [7] (Fig. 7).

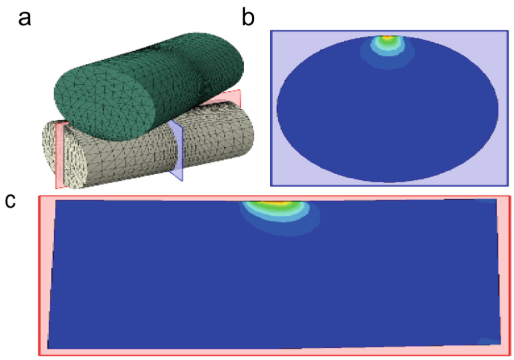


Fig. 7. a) Punctiform contact, b) and c) contact pressure field respectively in the cross and transverse section

2.2.4 Interaction Between the Different Scales

In order to be able to deal with a real problem, it is essential that the different scales exchange information in order to form a complete computation chain. This communication between models requires bi-lateral dialogues that ensure, as far as possible, the representativeness of calculations at different scales. Although the models used in the multi-scale approach involve different physical quantities (e.g. generalized stress vs local stress), it is necessary to compare results from both the linked models. It is this comparison that makes it possible to verify that the calculations carried out lead to a converged solution (Fig. 8).

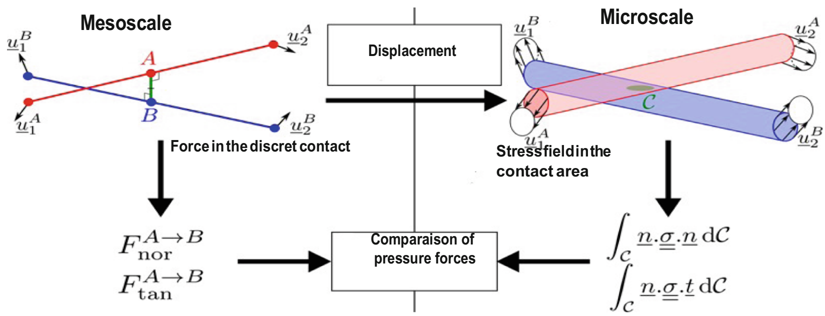


Fig. 8. Comparison of the pressure forces computed by the mesoscale and microscale models

2.3 Statistical Modelling

Statistical modelling is based on data, so it is essential to ensure the relevance of the latter. A first step which is very time-consuming is to clean and consolidate the databases to be used for statistical modelling. This work was carried out for the databases listed in the first part of this article.

We then built a training set composed of independent variables (temperature, wind speed,…) and one dependent binary variable (presence or not of a cut wire in the span).

Several Machine Learning algorithms were then applied to the training set. These included logistic regression, Naïve Bayes methods [8, 9], which are all parametric, and the Random Forest method [8, 10], which is non-parametric. Finally, cross validation methods were used to evaluate the performance of the different methods.

A Receiver Operating Characteristics (ROC, [11]) curve was used as a technique for visualizing the performance of a classification algorithm. ROC curves represent the trade-off between true positives and false positives. The y-axis represents the true positive rate, and the x-axis represents the false-positive rate. In order to be able to compare the performance of the different methods, a criterion named Area Under Curve (AUC) is used. This criteria is evaluated by calculating the area under the ROC curve for each method. In the figure below there is a comparison between several algorithms. Non-parametric algorithms seem to be more accurate for our example (Fig. 9).

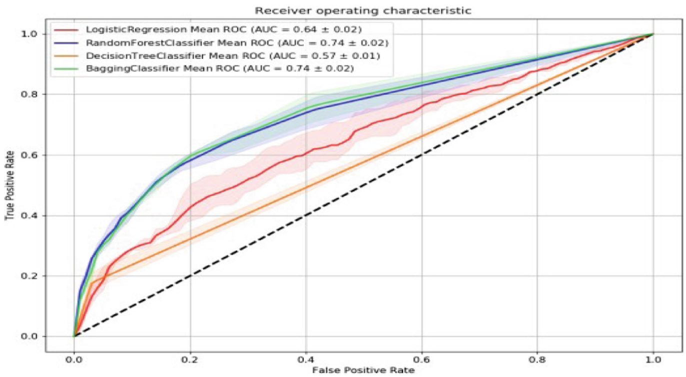


Fig. 9. ROC curve comparison for different algorithms

3 First Results

This tool is made up of two parts. The first part includes the databases, the machine learning algorithms and the mechanical simulations. The second part consists of a visualisation tool that allows visualization of the results of the hybrid modelling on a map of France.

In the below figure we present the first results regarding the presence or not of cut wires for a group of spans. The map on the left represents real observations (low number in light purple, high number in dark purple) and the plot on the right represents the risk of cut wires for a span (low risk in green, high risk in orange) obtained using the hybrid model. The results of the hybrid model and the observations seem to be in good accordance, but more validation is required (Fig. 10).

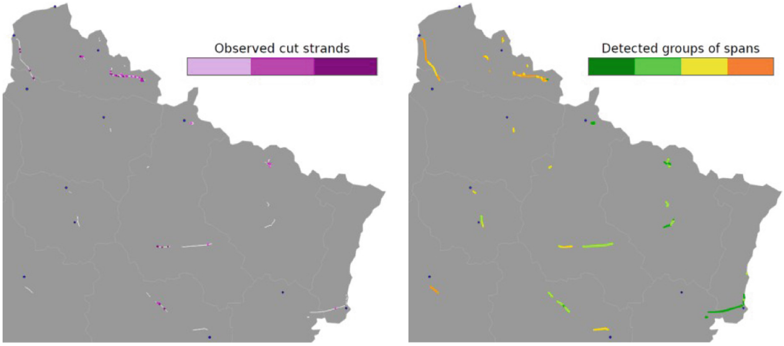


Fig. 10. Observed and detected cut wires by the decision-making tool

4 Conclusion and Perspectives

A decision-making tool has been developed in order to help RTE's engineers optimize their asset management policies. This tool is based on hybrid modelling and it is still in its beta version. The first results are promising but they need improvement. Future research perspectives that will lead to improvement of the results are the following:

- Statistical modelling: the data set needs to be enriched, notably by including new meteorological data,
- Mechanical modelling: the multi-scale model of an overhead line span is mature, but we need to improve the evaluation of the wind loading to have a better assessment of the stress field in the contact area,
- Decision-making tool: a validation process will be set up by in field verification of the predictions.

References

1. Cloutier, L., Goudreau, S., Cardou, A.: EPRI Transmission Line Reference Book: Wind-Induced Conductor Motion (2006). Orange book Chapter 1 to 3
2. Leblond, A., Hardy, C.: A unified way for presenting aeolian vibration response of single undamped over-head conductors. In: Proceedings of the Fifth International Symposium on Cable Dynamics (2003)
3. Zhou, Z.R., Cardou, A., Goudreau, S., Fiset, M.: Fundamental investigations of electrical conductor fretting fatigue. Tribol. Int. **29**(3), 221–232 (1996)
4. Hobbs, R.E., Raoof, M.: Mechanism of fretting fatigue in steel cables. Int. J. Fatigue **16**(4), 273–280 (1994)
5. Maouche, N.: Modeling of damage phenomena due to small amplitude of displacement in contacts, Ph.D. thesis, École Nationale des Ponts et Chaussées (1997)

6. Wang, D., Zhang, D., Ge, S.: Fretting-fatigue behavior of steel wires in low cycle fatigue. Mater. Des. **32**(10), 4986–4993 (2011)
7. Redford, J., Lieurade, H-P., Gueguin, M., Hafid, F., Yang, C., Ghidaglia, J-M.: Modélisation numérique du phénomène de fretting-fatigue intervenant dans le vieil- lissement des conducteurs de lignes aériennes, Materiaux & Techniques, under publication
8. Friedman, J., Hastie, T., Tibshirani, R.: The Elements of Statistical Learning. Series in Statistics. Springer, New York (2013)
9. Peng, C.-Y.J., Lee, K., Ingersoll, G.M.: An introduction to logistic regression analysis and reporting. J. Educ. Res. **96**(1), 3–14 (2002)
10. Louppe, G., Wehenkel, L., Sutera, A., Geurts, P.: Understanding variable importances in forests of randomized trees. In: Advances in Neural Information Processing Systems (2013)
11. Fawcett, T.: An introduction to ROC analysis. Pattern Recogn. Lett. **27**(8), 861–874 (2006)

Intelligent Decision Support for Maintenance: A New Role for Audit Trails

Chris J. Turner[1(✉)], Christos Emmanouilidis[2], Tetsuo Tomiyama[2], Ashutosh Tiwari[3], and Rajkumar Roy[2]

[1] Surrey Business School, University of Surrey, Rik Medlik Building, Guildford, Surrey GU2 7XH, UK
christopher.turner@surrey.ac.uk
[2] Manufacturing Department, Cranfield University, Building 50, Bedford, Bedfordshire MK43 0AL, UK
[3] Department of Automatic Control and Systems Engineering, University of Sheffield, Amy Johnson Building, Portobello Street, Sheffield S1 3JD, UK

Abstract. The changing nature of manufacturing, in recent years, is evident in industries willingness to adopt network connected intelligent machines in their factory development plans. While advances in sensors and sensor fusion techniques have been significant in recent years, the possibilities brought by Internet of Things create new challenges in the scale of data and its analysis. The development of audit trail style practice for the collection of data and the provision of comprehensive framework for its processing, analysis and use should be an important goal in addressing the new data analytics challenges for maintenance created by internet connected devices. This paper proposes that further research should be conducted into audit trail collection of maintenance data and the provision of a comprehensive framework for its processing analysis and use. The concept of 'Human in the loop' is also reinforced with the use of audit trails, allowing streamlined access to decision making and providing the ability to mine decisions.

1 Introduction

Increasingly manufacturing industry is adopting network connected intelligent machines in their factory development plans. This movement to incorporate new technology incorporating advances in Artificial Intelligence is described and encouraged by a number of international government/industry initiatives. The Industry 4.0 movement is one such initiative, between the German government and national industries, with a role to envisage and promote the use of new technologies and organizational methods for manufacturing (German Federal Government 2016). Cyber Physical Systems (CPS) are a core theme of Industry 4.0 encompassing the further integration between machines and computing resources. According to Lee and Bagheri (2015) CPS as the integration of physical assets with intelligent software systems that will enable a new generation of maintenance practice. In addition, the enhanced information processing and analysis opportunities provided by the ubiquity of sensor use in modern machinery to provide data streams and resulting Big Data sets is seen to

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 396–403, 2020.
https://doi.org/10.1007/978-3-030-48021-9_44

create potential for new products and new types of manufacturing models. A key challenge for maintenance practice is how the opportunities brought by this expansion of data can be best realised. The issue of data provenance is key in the formation of meaningful analytics and, when coupled with an audit trail framework, offers a valuable methodology for reliable and secure data driven decision making in the implementation of maintenance practice.

2 Literature Review

The quality and provenance of data are important factors when engaging in any form of analytics. Lin et al. (2007) conducted a survey into data quality relating to asset management information. The survey found that processes and software for asset related data quality management were missing in a majority of organisations interviewed; in addition organisations did not have a strategy in place regarding data quality (Lin et al. 2007). Haider (2015) propose a framework for asset lifecycle management data governance, stating that organisations need policies to ensure data quality is inherent in their operation. Woodall et al. (2015) define seven information quality dimensions for organisations to audit their operations by to establish an actual level of data quality and potentially identify areas for further improvement. The OPC UA (OPC Unified Architecture) standard for industrial system inter-communication while comprehensive in its specification can be complex and expensive for an organisation to implement. The work of Henßen and Schleipen (2014) examines the role that the AutomationML mark-up language can play in simplifying the use of OPC UA models with existing data sets and streams expressed in XML. According to Henßen and Schleipen (2014) use of OPC UA directly is a complex task, utilising AutomationML mapping to OPC UA opens up the opportunity of streamlined connectivity with OPC UA compliant systems and manufacturing systems. Liyanage et al. (2009) detail the semantic web, ontology and use of XML metadata description use for information exchange in e-maintenance. Karray et al. (2009) make the point that semantic interoperability between systems is key to the successful operation of e-maintenance. Grangel-Gonzalez et al. (2016) take the semantic communication notion a step further by producing a metadata software shell for Industry 4.0 components. The approach is based on RDF (Resource Description Framework) and OWL (Web Ontology Language) and aims to allow for new functionality, described by ontological elements, to be integrated into the communication framework with minimum disruption (Grangel-Gonzalez et al. 2016). In combination with machine intelligence such a framework could acts as an enabling protocol for automation efforts in maintenance activities and factory operations alike.

Vaughn et al. (2005) examine the possibility for automated cyber vulnerability recognition where sensor data is used to trigger security warnings. The aim of automated cyber security is also sought by Abreu et al. (2015) with the use of audit trail data. With this work Abreu et al. (2015) employ machine learning techniques to derive patterns and insights to, in principle, enable automated actions and decisions to be made. Duncan and Whittington (2016) advise on the regular analysis of audit trails in the effective securing of cloud based systems. While useful in countering intrusions into maintenance systems it is also the case that such approaches provide much of the rigor and data management practise required to ensure quality and enforce standards within an organisation and its supply chain and linked parties. The use of such audit trail techniques in manufacturing has been much less evident though its use with IoT has in outline been explored by Lomotey et al. (2018) in research exploring the need for visualisation of internet connected devices. In addition Lomotey et al. (2018) propose a provenance methodology to allow for improved traceability and identification of routes through a network that specific data points may take. Efforts towards a unified metadata syntax and model for provenance are embodied in the work of Moreau et al. (2011) who put forward the Open Provenance Model (OPM), enabling the unified and secure exchange of such data between networked systems and entities.

3 Maintenance and Retaining 'Humans in the Loop'

With the use of such audit trail based intelligent data mining there arises the potential need to explain the reasoning behind automated decisions to humans for the purposes of evaluating/ensuring provenance of maintenance data. Duncan and Whittington (2016) make a number of recommendations on how the audit trail for cloud computing could be improved; the following are an adaptation of a subset of those recommendations with relevance to the maintenance field:

- A strict regime of data log migration to data storage is required
- Further understanding on information flow within manufacturing is required
- Enhanced data security is required to safeguard collected audit trail data and digital entry points to manufacturing systems from cyber attackers

It is the case that a 'human in the loop' is required as their expert knowledge and overview capability can be leveraged, in particular, to help ensure data and processes security. A vital step along the road to automation is the inclusion of human expertise along with standards such as the MIMOSA open system architectures for CBM and EAI (Enterprise Application Integration) (MIMOSA 2017), which potentially provide a wider underlying structure for the concept of maintenance audit trails. Furthermore, "human in the loop" – generated events can be viewed as a crowdsourced timeline of maintenance linked knowledge, contributing to maintenance and asset management data quality, if adequately mined (Pistofidis et al. (2016).

4 The Audit Trail for Maintenance

When considering the implementation of audit trail practice within an industrial setting it is important to consider the data flows currently available. Many enterprise systems in organisations, such as ERP (Enterprise Resource Planning), possess event logging capabilities. Such event logs may be mined in order to reconstruct a chain of activities that have taken place within the organisation and administrated by the system (Turner et al. 2012). Events, even when not mined as complete process chains, may still be further augmented by semantic descriptions (attached as metadata tags) to establish provenance. Sensors in both production machines and located within the production line/maintenance environment can provide a wealth of data when systematically connected to data repositories. This data may even be available in the form of a continuous stream, informing of the live status of the asset in question. Real time or near to real time evaluation of machine condition data can provide a new level of insight when combined with prognostic techniques within maintenance decision support and planning systems. Use has been made of audit type data in industrial applications.

A sensor fusion approach has been used by Payan et al. (2016) in the development of proactive safety metrics for helicopters. In this research Payan et al. (2016) fuse the outputs of flight data monitoring to form the basis for predictive safety measures, with the potential to advise preventative measures. Such an approach may also inform the development of audit trail compilation and use to enhance the scheduling and performance of maintenance activities. One way of presenting data provenance information in a non-technical fashion is through the provision of a text based audit trail (written in plain language) describing the individual steps taken to arrive at a generated recommendation (including any calculated weightings or percentages used). In addition the user should be made aware of the sources of data used in the generation of decisions and a suitable provenance should be available for each individual source for human based cross checking and evaluation. The concept of 'Human in the loop' is reinforced within this framework through streamlined access to decision making and the ability to mine audit trails of decisions (and the reasoning behind decisions) and activities that have occurred within the Internet connected production line. There will be a necessity to capture and store data streams from the production line and audit trails of decision making within the semantic sandwich layer and monitored activities within the system. A big data repository will be required along with connectivity to other data stores and streams within the organisation.

Figure 1 illustrates the concept of the audit trail with an example drawn from railway maintenance activities. In Fig. 1 it can be seen that for a section of track there are a range of maintenance activities that may involve: maintenance workers, feeding back reports via mobile devices; rail maintenance vehicles with sensors; passenger trains fitted with track and infrastructure monitoring sensors. In addition a number of trackside sensors may also stream back data to a control centre concerning a range of environment specific parameters. The scenario depicted in Fig. 1 relates to the possibility that sensors have registered faults with a Balise (track based transponder forming part of an Automatic Train Protection (ATP) system) and trackside signals in a period of time after the section of track has been tamped (where the ballast bed of the track is adjusted). In addition a bankside sensor has noted some occasional subsidence in the past. All these data streams are recorded at a central control centre. The use of data mining may establish a causal link between these events taking into account the outlier measurement from the bankside sensor leading to the root cause of the fault.

New forms of data capture are proposed for use in the rail industry that include RFID for the location of trains, lower cost (and more compact) condition monitoring sensors built into train components, and sensors on trains monitoring the track and overhead power lines (Kans et al. 2015); in essence the use of IoT for data point and stream collection is proposed allowing physical assets to communicate their real time status and health. Kans et al. (2015) outline an intelligent maintenance approach

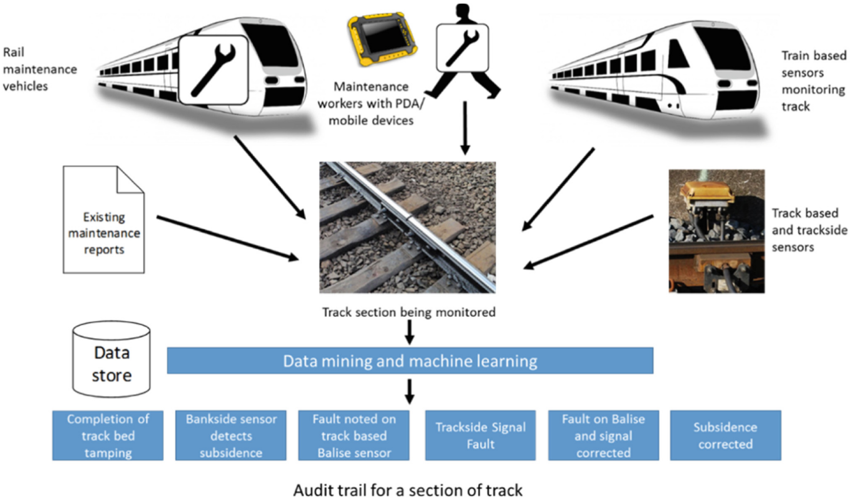


Fig. 1. An audit trail drawn from rail maintenance activities and sensor streams for a section of track

capable of interaction with Rail Traffic Management systems for the dynamic mitigation and rerouting of services on identification of infrastructure and or rolling stock breakdowns. A focus within the work of Kans et al. (2015) is the use of forecasting and prediction of faults through range of techniques including machine learning. Decision making within intelligent maintenance systems can be complex due to the sheer variety of data sources. Through the use of metadata descriptions and timestamps it is possible to establish a process for the analysis of collected data. The audit trail establishes the order of events via timestamps and the output from data mining/machine learning. The metadata descriptions can be used to describe the data collected and its potential use (along with numerical values) and combined with event logs generated by ERP systems; such descriptions can be used within the data analysis process and the text displayed in the form of an annotated process to human operators and managers (and potentially track side workers). Such audit trails can help reveal the path of the automated decision making within an intelligent system. Vergidis et al. (2015) propose the use of an annotated process to describe the intelligent optimised composition of web services; an extension of such a technique could be employed to compose metadata trails within analysed data through text extraction and the additional use of semantics, context based computing and machine learning based data mining approaches.

5 Discussion

It is clear that initiatives such as Industry 4.0 are changing attitudes towards digital connectivity and automation in manufacturing, though it is the case that there is a need for a holistic understanding of data being collected and analysed. It is also true that industry is still missing an overall framework for digital maintenance. Advances in sensors and senor fusion techniques have run-ahead of suitable processes and systems capable of fully harnessing their outputs. The quality and provenance of data are important factors in data management. With maintenance rapidly adopting key Industry 4.0 technologies, such issues attain increased importance in the delivery of successful applications and services. Product and asset lifecycle data are increasingly acknowledged as valuable assets (Kubler et al., 2015). Therefore their own lifecycle needs to be appropriately managed and this could become a key factor in establishing a credible audit trail for maintenance activities and data. Imran et al. (2017) provide the following outline points to take account of when collecting and describing the provenance (perhaps through metadata tagging) of data products:

- From where a data product was acquired?
- By whom and when the data product was created?
- Who are the authorized stakeholders of the concerned data product?
- In what transformations and computations has it been used?
- What were the inputs for the generated output data item?
- Which criteria were applied in the creation of the data product?

Work in the area of event logging based audit trails that have been utilised in the field of cyber threat detection within networked software systems, are of direct complimentary use in ensuring that data flows are not compromised by intruders. The implementation of audit trail methodology could therefore bring inherent additional security benefits to any implementing organisation.

6 Conclusion

Further work remains to be completed on understanding information flows within manufacturing and how digitisation of systems and information will impact maintenance activities. Though it is put by the authors that clear processes to support audit trail style collection of maintenance data and the provision of a comprehensive framework for its processing, analysis and use should be important goals for the research that must be completed in the near future for full enablement of digital maintenance practice. The concept of ‘Human in the loop’ is also reinforced with the use of audit trails, allowing streamlined access to decision making and the ability to mine decisions (and the reasoning behind decisions for machine assisted workers and managers). Security concerns inherent in the connection of live production line assets to networked systems may well be allayed by the introduction of an audit trail methodology. Developments in the security arena and their use of audit trail must also be acknowledged in future research relating this practice to maintenance activities. The ability to provide procedural structure to data for reuse and communication within an Industry 4.0 maintenance system will be vital for any future move towards semi or fully autonomous maintenance activities.

References

Abreu, R., Bobrow, D.G., Eldardiry, H., Feldman, A., Hanley, J., Honda, T., de Kleer, J., Perez, A., Archer, D., Burke, D.: Diagnosing advanced persistent threats: a position paper. In: DX@ Safeprocess, pp. 193–200 (2015)

Duncan, R.A.K., Whittington, M.: Enhancing cloud security and privacy: the power and the weakness of the audit trail. In: Westphall, C.B., Lee, Y.W., Rass, S. (eds.) Cloud Computing 2016: The Seventh International Conference on Cloud Computing, GRIDs, and Virtualization, IARIA, p. 137 (2016)

German Federal Government: The new high-tech strategy innovations for Germany (2016). https://www.bmbf.de/pub/HTS_Broschuere_eng.pdf. Accessed 16 Mar 2018

Grangel-González, I., Halilaj, L., Auer, S., Lohmann, S., Lange, C., Collarana, D.: An RDF-based approach for implementing Industry 4.0 components with administration shells. In: 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1–8. IEEE (2016)

Haider, A.: Asset lifecycle data governance framework. In: Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012), pp. 287–296. Springer, Cham (2015)

Henßen, R., Schleipen, M.: Interoperability between OPC UA and AutomationML. Procedia CIRP **25**, 297–304 (2014)

Imran, M., Hlavacs, H., Haq, I.U., Jan, B., Khan, F.A., Ahmad, A.: Provenance based data integrity checking and verification in cloud environments. PLoS ONE **12**(5), e0177576 (2017)

Kubler, S., Yoo, M.-J., Cassagnes, C., Framling, K., Kiritsis, D., Skilton, M.: Opportunity to leverage information-as-an-asset in the IoT – the road ahead. In: 3rd International Conference on Future Internet of Things and Cloud, pp. 64–71 (2015)

Lee, J., Bagheri, B.: Cyber-physical systems in future maintenance. In: 9th WCEAM Research Papers, pp. 299–305. Springer, Cham (2015)

Lin, S., Gao, J., Koronios, A., Chanana, V.: Developing a data quality framework for asset management in engineering organisations. Int. J. Inf. Qual. **1**(1), 100–126 (2007)

Liyanage, J.P., Lee, J., Emmanouilidis, C., Ni, J.: Integrated e- Maintenance and intelligent maintenance systems. In: Handbook of Maintenance Management and Engineering, pp. 499–539 (2009)

Lomotey, R.K., Pry, J.C., Chai, C.: Traceability and visual analytics for the Internet-of-Things (IoT) architecture. World Wide Web **21**(1), 7–32 (2018)

MIMOSA Machinery Information Management Open Systems Alliance (2017). http://www.mimosa.org/. Accessed 16 Mar 2018

Kans, M., Galar, D., Thaduri, A.: Maintenance 4.0 in railway transportation industry. In: Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015), pp. 317–331. Springer, Cham (2015)

Karray, M.H., Morello, B.C., Zerhouni, N.: Towards a maintenance semantic architecture. In: 4th World Congress of Engineering Asset Management, WCEAM09, Athens, Greece (2009)

Moreau, L., Clifford, B., Freire, J., Futrelle, J., Gil, Y., Groth, P., Kwasnikowska, N., Miles, S., Missier, P., Myers, J., Plale, B.: The open provenance model core specification (v1.1). Fut. Gener. Comput. Syst. **27**(6), 743–756 (2011)

Payan, A.P., Gavrilovski, A., Jimenez, H., Mavris, D.N.: Review of proactive safety metrics for rotorcraft operations and improvements using model-based parameter synthesis and data fusion. In: AIAA Infotech@ Aerospace, San Diego, California, USA, 4–8 January 2016, p. 2133 (2016)

Pistofidis, P., Emmanouilidis, C., Papadopoulos, A., Botsaris, P.N.: Management of linked knowledge in industrial maintenance. Ind. Manag. Data Syst. **116**(8), 1741–1758 (2016)

Turner, C.J., Tiwari, A., Olaiya, R., Xu, Y.: Business process mining: from theory to practice. Bus. Process Manag. J. **18**(3), 493–512 (2012)

Vaughn, R.B., Farrell, J., Henning, R., Knepper, M., Fox, K.: Sensor fusion and automatic vulnerability analysis. In: Proceedings of the 4th International Symposium on Information and Communication Technologies, Trinity College Dublin, pp. 230–235 (2005)

Vergidis, K., Turner, C., Alechnovic, A., Tiwari, A.: An automated optimisation framework for the development of re-configurable business processes: a web services approach. Int. J. Comput. Integr. Manuf. **28**(1), 41–58 (2015)

Woodall, P., Gao, J., Parlikad, A., Koronios, A.: Classifying data quality problems in asset management. In: Engineering Asset Management-Systems, Professional Practices and Certification, pp. 321–334. Springer, Cham (2015)

Part V: Asset Safety and Occupational Health

Toxicity Limit States in Tunnel Fire Safety Designs

Lene Østrem[1(✉)] and Ove Njå[2]

[1] Aibel AS, Jens Risøens gate 72, 5527 Haugesund, Norway
Lene.Oestrem@aibel.com
[2] University of Stavanger, Ullandhaug, 4036 Stavanger, Norway
ove.njaa@uis.no

Abstract. The Mont Blanc tunnel fire March 1999 killed 39 persons, of which most died within 15 min due to intoxication. In Norway there have been several fires the recent seven years. No single road-user has died from intoxication in those fires, in spite of being engulfed with smoke for more than 1.5 h. The tunnel safety discourse amongst tunnel owners and researchers turns towards questioning whether current longitudinal ventilation strategies can be used to design the tunnel system to meet the self-rescue principle. Smoke control would then be the design criterion. The Norwegian Public Roads Administration could in this perspective re-duce its effort to invest in safety measures ensuring safe havens for road users trapped in smoke and other fire preventive measures. We are very critical to such a development of tunnel fire safety. This paper raises questions about predictability of smoke dispersions in case of tunnel fires as well as human tolerability of toxic gases from fires. We conclude with issuing designs of research studies to reduce the gaps of knowledge revealed in the literature.

1 Introduction

In Norway there is approximately 1100 road tunnels, and 33 of them are below sea level. Several new tunnels are being planned and many are under construction. In the last 10 years there has been an increasing focus in the Norwegian community on fire safety related to tunnels. This is due to an increase in the number of tunnel fires in Europe and Norway, but also because the society in general has an increased focus in risk analyses and safety. Designs are getting more and more complex within several industries, including the transportation sector, increasing the need for credibility of acceptance criteria based on human tenability when the possibility to escape is restricted. This paper is a first step to raise questions about current engineering practices when it comes to human tenability limits.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 407–415, 2020.
https://doi.org/10.1007/978-3-030-48021-9_45

2 Discussing the Norwegian Statistics

A review of Norwegian tunnel fires in the period 2008–2015 (Nævestad et al. 2016) reveals that larger vehicles are overrepresented in tunnel fires in Norway. Tunnels with steep incline is overrepresented; 5% of the tunnels in Norway represented 42% of the fires. 40% of these fires involved heavy goods vehicles, which is a lot taken into account that they only measure 14% of the traffic load on Norwegian roads (Nævestad et al. 2016). Tunnel design in Norway has an increasing degree of complexity, resulting in a need for a performance-based design.

The NPRA has an ambition to develop a model for fire risk that can predict probability and severity of fires in the Norwegian tunnels. Such modelling work will hardly reflect all major factors and system characteristics and thus, the use could be counterproductive. "Though the primary goal of stochastic modelling is to provide insights and not numbers, numerical answers are often indispensable for gaining system knowledge" (Tijms 1994). Tijms' basic text book concludes in its beginning on what is important with quantitative analyses. The interpretation of the results of stochastic models is of no value unless we know the models and data material supporting them. Nævestad et al. (2016) shows that the average occurrence of fires in the entire Norwegian population of tunnels are 24 per year (min 17 - max 34), which of rather few occurred in heavy vehicles (>3,5 t). Distributed over the NPRA-regions the fires that somehow included heavy vehicles were:

Region East – 17 fires in 8 years, of which 5 in the Opera tunnel, 5 in the Oslofjord tunnel and 2 in the Tåsen tunnel, while the last 5 occurred in five different tunnels.
Region South – 4 fires in 8 years, all in different tunnels.
Region West – 24 fires in 8 years, of which 3 in the Mastrafjord tunnel, 2 in the Bømlafjord tunnel, 2 in the Gudvanga tunnel, and 17 fires in different tunnels.
Region Mid Norway – 11 fires in 8 years, of which 4 in the Hitra tunnel, 2 in the Stavsjøfjell tunnel, 2 in the Eiksund tunnel, and 3 in different tunnels.
Region North – 7 in 8 years all in different tunnels.

Major fire loads have been reported in the two fires in the Gudvanga tunnel, the Brattli tunnel, the Follo tunnel, the Skatestraum tunnel and in two fires in the Oslofjord tunnel. The fire in the Follo tunnel killed the HGV-driver. No other fires have killed road-users due to intoxication, but in the same period 5 road-users have been killed in tunnel-accidents which have also included fires. We scrutinized the major accidents (major injuries to people) and found:

- In 2009 a person was killed in a head on collision between a private car and a HGV in the Stavsjø tunnel (mid night before Saturday – the private car came over into the HGV's lane). The young male driver (in the 20-ies) was killed. A fire in the private car was put out immediately.
- In 2010 a head on collision in the Hordvik tunnel also between a HGV and a private car implied death of the driver of the private car. Both vehicles caught fires but were immediately extinguished.

- In 2010 a Lithuanian driver of a HGV died due to mechanical injuries and smoke inhalation from the collision with the Follo tunnel portal and the tunnel wall in the entrance zone.
- In 2011 a fire occurred in a Polish HGV in the Oslofjord tunnel, which implied major smoke inhalation injuries for several road-users (Njå and Kuran 2015). The data material from Statistics Norway and the study made by Institute of Transport Economics (Nævestad et al. 2016) described the consequences as minor, which is an error.
- In 2011 two persons were killed in a head on collision between a bus and a private car. The accident occurred outside the Vassenda tunnel. Smoke was seen from the private car, but no fire occurred in any of the vehicles. The event is part of the official tunnel fire statistics, but this could be questioned.
- In 2012 a fire started in the rear tires of a HGV-trailer in the Mastrafjord tunnel. Two persons were reported with small injuries from smoke inhalation (ref. a local newspaper), but the information from Statistics Norway described it as serious injuries.
- In 2013 a head on collision between a private car and a HGV in the Storesand tunnel killed the driver of the private car. The HGV caught fire, but was extinguished rapidly.
- In 2013 a motorcycle driver was killed in a collision with a lorry in the Naustdal tunnel. A minor fire was immediately extinguished.
- The fire in the Gudvanga tunnel included many intoxicated patients, of which many were seriously injured.
- In 2015 the second Gudvanga fire included a bus. It implied five injured persons from smoke inhalation.

The largest HRRs were observed in two other fires that luckily included no persons still in those tunnels when the fires developed (Skatestraum and Brattli tunnels). The fact that no road-users were in the tunnels were not subjected to any system safety measures, but “pure luck”. In the work with tunnel safety, it is acknowledged that there are major uncertainties about the consequences of exposing people to fire smoke over a longer period. Research shows that design of the tunnel will affect the fire growth and development (Ingason et al. 2014), hence the possibility to ensure safe evacuation.

3 Fire Toxicity in Norwegian Tunnels

The main cause of injury and death in fires is exposure to toxic fire smoke and gases (Stec and Hull 2010). In the event of a fire, fire safety depend upon the outcome of two parallel timelines: the time from ignition of the fire to the development of incapacitating conditions (ASET) and the time required for occupants to reach a place of safety (RSET) (Hurley 2016, chapter 63, p. 2308–2428). When occupants become immersed in smoke, behavioural, sensory and physiological effects occur. Toxic fires effluents are responsible for the majority of fire deaths and an increasingly large majority of fire injuries (Stec and

Hull 2010). Fire safety in general has often focused on preventing ignition and reducing flame spread, and not so much focus have been given to the fire toxicity.

Since no one have been killed by the smoke in a tunnel fires in Norway, a perception seems to emerge amongst tunnel owners that the smoke has been non-toxic. As of today there is no acceptance criteria or design requirement regarding smoke obscuration in Norwegian tunnels (SVV 2016). This silently supports the choice of ventilation strategy, which is longitudinal instead of transverse ventilation (SVV 2016). Smoke exposure can however delay or prevent escape for an extended period, during which fire conditions may become life threatening (Stec and Hull 2010). Major accidents worldwide have had devastating consequences. On 24 March 1999 a Belgian truck with a refrigerated trailer carrying margarine and flour caught fire in the Mount Blanc tunnel and resulted in 39 deaths, and major complexities in the fire and rescue work. Those who tried to escape managed to make only 100 $\sim$ 500 m before collapsing due to lethal smoke compositions (Duffé and Marec 1999). In the St Gotthard Tunnel, 24 October 2001 two HGVs collided and a fire broke out. The fire spread rapidly, and even though the fire brigade managed to enter the tunnel in less than 7 min, the fire burned for approximately 24 h. After the fire was brought under control, the bodies of 11 people were found to the north of the incident location within a distance of approximately 1250 m. Some were inside their vehicles, other were on the road way. Ten died as a result of smoke inhalation (Carvel and Beard 2005). In Kaprun November 2000 155 tourist were killed in a ski train blaze. Several passengers ascending on foot, as well as the train conductor, were asphyxiated by the smoke and then burned by the fire (Sempio 2013).

Taken into account the tragedy and lesson learned from these accidents, we question the design approach used in Norwegian tunnels. Understanding the fire dynamics in tunnels and how the fire interacts with its surroundings is important factors when evaluating fire safety design in tunnels (Ingason et al 2014). The components present in the fire smoke is a result of the goods carried by the HGVs, and how different components interact with each other when exposed to extreme heat load. Currently there are very few restrictions. New type of vehicles, technologies and fuels are integrated in fast pace in the transport systems. The tunnel regulator uses performance based rules, but we questions the design practices. These shortcomings have been evident in the investigations and national audits. There is a need for design requirements to keep up with the technological development, resulting in a more functional and performance based legislation that challenge existing practices.

There is a need for informed expert decision making for professional management regarding safety in tunnels, taking into account the uncertainties regarding the fire smoke toxicity. It is important to understand that the yields present will be scenario based dependent on the contexts, materials and ventilation conditions. The time-concentration curves of the toxic products depend on the mass burning of the fuel (kg/s), dispersal volume (to give kg/m3) and the yields of each toxic product (kg/kg). The yields at different stages will depend upon fuel substances, fuel/air equivalence ratio, temperature and oxygen concentration in the flame zone (Stec and Hull 2010). The main dangers presented by smoke are obscurity (lack of visibility prevents people from fleeing), toxicity (which incapacitates) and temperature (which also incapacitates) (Hurley 2016, chapter 63, p. 2308–2428).

3.1 Regulation

The EU Road Tunnel directive (Directive 2004/54/EC) is prevailing in Norway. The effect of the Directive is to constitute what is regarded as a minimum European level of safety in road tunnels. However, the obligation of engineers to exercise professional care remains even if there are directives in place. The EU directive and the N500 tunnel code require risk assessment to demonstrate acceptable safety levels during evacuation in special circumstances. We raise the question; what is acceptable safety level during evacuation? Why is it possible for a tunnel owner to direct the self-evacuation principle, when road-users do not know what it mean?

Further, a comparative analysis of safety standards for road and rail tunnels performed by Arnold Dix (2004) assessed the regulatory frameworks in Germany, Austria and Switzerland against the EU Road Tunnel directive, and furthermore Japan's approach. The comparison demonstrates a vast range of designs and operational conditions when it comes to underground transportation safety. Dix concluded that the great variation in key safety parameters such as ventilation, lightning, emergency evacuation, control systems and pedestrian ways, require expert engineering in design and operation. In the US the NFPA (National Fire Protection Association) standards are highly recognised and used for fire safety. NFPA 502 is the standard for road tunnels, bridges, and other limited access highways. When it comes to means of egress from a road tunnel it is stated that a tenability level shall be provided in the means of egress during the evacuation phase in accordance with the emergency response plan. A criteria for tenability and time of tenability should also be established (NFPA 2017, ch. 7.16.2). Further reference is given to NFPA 101, Life Safety Code, which is the most widely used source for strategies to protect people from the hazardous exposure from fires (NFPA 2018). The NFPA codes used in fire safety design are moving away from being prescriptive to become more performance-based or scenario-based. For instance in Hong Kong, the fire safety strategies optimize fire protection and fire prevention measures to attain specified fire-safety objectives. Fire safety systems must be defined clearly and include at least three parts: detection and alarm system, fire control system and air and smoke control system (Miclea et al. 2007). Keeping the thermal and toxic effects under acceptable and tenable limits are considered extremely important and tenability limits are stated, amongst other on CO_2 concentration levels.

Effective dose and concentration levels are more commonly used to provide an indication of lethality and incapacitation, from the cumulative effect of the most noxious fire effluents, expressed as fractional effective dose or concentration (FED or FEC) (Hurley 2016, chapter 63, p. 2308–2428). ISO 13344:2015 states that pyrolysis or combustion of every combustible material produces a fire effluent atmosphere, which, in sufficiently high concentration, is toxic (ISO 2015). The standard provides means for estimating the lethal toxic potency of fire effluent produced during a fire. The lethal toxic potential are related to the fire model selected, the exposure scenario and the material evaluated. Lethal toxicity values associated with 30-min exposures of rats are predicted, using calculations. The intended use of fire safety engineering calculations is for life-safety prediction for people and is most frequently for time intervals somewhat shorter than 30 min. It must be kept in mind that the importance of considered engineering decision making and evaluation of risk and uncertainties underlies all standards and guidelines.

3.2 Fire Toxicity

From the 1970s until the early 1990s fire toxicity was recognised as a serious problem and some high quality research was undertaken (Stec and Hull 2010). It was then discovered that real fires had a much higher level of toxicity than small-scale laboratory tests. The difficulties of replicating real fires on a bench scale could be one of the reason for diminution of research into fire toxicity. The focus of fire safety research changed towards reducing peak heat release rates. Recent years there have been a resurgence of interest in fire toxicity, mainly due to performance-based design approaches to fire safety engineering in several industries.

In forensic investigations fire toxicity has played an important role and blood samples are routinely analysed for carbon monoxide to ensure whether or not the victim was alive after the fire started (Stec and Hull 2010). This has however led to the assumption that because the carbon monoxide levels were easily quantified in the blood, this is the only important toxicant, which is not the case. Most of the fire models used today is based on 30–50 years old experiments (Stec and Hull 2010). Studies of fire effluent toxicity is a multidisciplinary area where both fuel chemistry and conditions of the complex process of fire have significant influences (Stec and Hull 2010). It requires understanding of the stages of fire growth – from ignition to ventilation controlled burning, the behaviour of a fire in different scales combined with the effect of the interactions with the surrounding environment (air supply, walls, ceiling etc.), the product formation from flaming polymer pyrolysates, the behaviour of the aerosol particulates and the response from the human body to the components present, the chemical quantification of those fire effluents and the toxicity of these (Stec and Hull 2010). Proper investigations of fires and victims involved have also been scarce, mostly directed towards liability investigations.

Assessing the fire safety in a road tunnel, thus, requires application of detailed knowledge of fire development and smoke toxicity combined with the understanding of risk management. When introducing the concept of risk assessment to decide upon acceptable levels of risk in a tunnel, the situation may easily arise where the analyst do not see the full scope of the choices that are made. We questions the outcomes of risk management strategies, especially when we take into account the limitations in the knowledge regarding tenability limits and the large variance of human behaviour in fires. The effects of fire on occupant's can be divided into three phases (Hurley 2016, chapter 63, p. 2308–2428):

- Phase 1: The fire is growing but the occupants are not affected by heat or smoke.
- Phase 2: Occupants are exposed to smoke, heat and toxic products. At this stage irritancy and asphyxiation will affect their escape capability. At this point in time, factors such as the toxicity of the fire smoke and the dynamics of their production become critically important when trying to escape.
- Phase 3: This phase is the terminal phase of victims as a result of the fire.

The toxic effects of the fire product are important in the second and third phase. However, most studies of fire toxicity have been regarding lethality, for instance the Strathclyde study in the United Kingdom. The lethality, in terms of the LC_{50} on laboratory animals, have been focused on individual fire products such as carbon monoxide (CO) or hydrogen chloride (HCL) or a mixture of thermal decomposition products from materials (Hurley 2016, chapter 63, p. 2308–2428, chapter 63, p. 2308–2428). The phase with incapacitation in fires can be studied either by animal experimentation or by investigations of the circumstances surrounding real fire casualties, particularly survivors of serious smoke exposure. This crucial area of toxicity has been largely neglected. Research performed by TNO (The Netherlands Organisation of Applied Scientific Research) published in the coloured books; *Methods for the determination of possible damage to people and objects resulting from release of hazardous materials* (CPR 1992), states that the combustion products that theoretically can appear in a fire are mainly determined by the chemical composition of the substance. If for instance hetro-atoms are present, such as chlorine and sulphur, in addition to carbon and hydrogen, then next to CO, CO_2 and H_20 also CL_2, HCL, $COCL_2$, $S0_2$ and COS will appear. This will typically be called primary combustion products. But in addition to this also secondary combustion products will be generated (CPR 1992), as a result of mutual reactions between the combustion products that are formed. Generally there are very little data available with regard to secondary combustion products. However there are some exceptions, for instance combustion of polychlorinated aromats polychlorodibenzo-p-dioxins (PCDD), which are products of incomplete combustion of organic materials.

3.3 Modelling Fire Toxicity

TNO suggests methods to help making it clearer in which manner the formation of combustion products can be defined (CPR 1992), but the guidelines should be considered as an indications containing a relatively high degree of uncertainty. The "green book" establishes a methodology, for a number of substances, acute toxicity data, which are applicable for the inhalation by human beings. Data available (mostly from animals), with the help of an extrapolation model, a 30 min LC_{50} value for human being was derived. An LC value is the concentration at which a given percentage of exposed population will die. In this case 50%. The calculation of LC_{50} for human is based on the known LC_{50} for animals. The latter are converted to values corresponding to a 30 min exposure duration. Thereafter extrapolation is made by the help of an extrapolation factor. The methodology contains several uncertain factors. The parameters used in the method are also only valid for lethal injury. Proper values of the parameter for other types of injures, for instance lung damage, respiratory system and alimentary canal disturbance, does not seem to available. TNO recommends for future research on acute toxicity to try and obtain better definitions for these type of injuries that could arise as a consequence when exposed to smoke from a fire. They also states that the probit constants for human beings represent no more than an indication, and that a lot of research is required to arrive at really reliable dose-effect relationships (CPR 1992). Newer standards and guidelines are often based on the coloured books, which are based on old studies and experiments.

The models present in the "green book" is for use in quantitative risk analysis. The uncertainty in the models must be considered within the framework of, sometimes, other relatively large uncertainties, that could be effect models, probability models, population data, etc. Still, when assessing risk analysis for road tunnels, a risk analysis containing results from these methods are used when making decisions regarding fire safety. In order to evaluate toxic conditions for people in fires, one needs to determine physiological and pathological effects of exposure to toxic smoke and how they impair escape, cause incapacitation and death. When evaluating the risk of fire toxicity present in a road tunnel it is necessary to identify the main toxic species responsible for these effects amongst the hundreds of chemical species known to occur in fire effluents and combinations that might occur. This is to say at least a rather complex task. In tunnels transportation of a large variety of non-hazardous and hazardous goods will be present.

4 Recommendation and Conclusions

The design phase for complex tunnel structures with limitations in escape routes introduces gaps in how to work with risk factors. When looking at the uncertainties present in models and knowledge regarding tenability of humans exposed to smoke, it is a big surprise to us that analyses of toxins and human responses are often neglected in tunnel designs. The fire safety engineering practise moving towards a performance based approach to fire safety design has not improved this situation.

Incapacitation in tunnel fires, and how these products affect the capability of escape during fire, is a crucial area containing several uncertainties that has been neglected. Some of these uncertainties regarding fire risk in tunnels have been reflected on in this paper. Regulatory variations globally demonstrates the importance of professional expert engineering decision in the design phase. Lesson learnt from previous tunnel fires tragedies requires attention. Stronger attention should be given to smoke toxicity in the design phase. Some questions that needs to be further explored are:

Design phase:

1. What are the experiments and knowledge base for the empirical models used in fire safety today when it comes to fire smoke toxicity?
2. How manage risk and uncertainties in the design phase regarding exposure to smoke during evacuating, to continuously strive for an inherently safer tunnel design?
3. To which degree is the uncertainty regarding fire smoke toxicity reflected when modelling fire in risk analysis for road tunnels?

Human aspect:

1. Is it possible to develop methods to determine different level of incapacitation when exposed to fire smoke?
2. Is the use of human tenability limits beneficial, considering that what is adequate from a safety perspective often is a political and/or industry question?

All questions above (but not limited to), are important aspects to consider in fire safety engineering. There is a need to increase the understanding of the experiments and empirical models used, when introducing human tenability as a risk acceptance criteria. Empirical models are often used in fire safety engineering, but there is a gap of knowledge when it comes to incapacitation and long terms effect of being exposed to fire smoke during evacuation. Especial long tunnels with very limited possibility for escape introduces challenges. Using human tenability limit in design, knowledge regarding combustion chemistry and human biological and psychological effects on people are just as important as the fire development itself. Modelling fires and allowing development of fire scenarios, introduces a need for tenability limits and risk acceptance criteria`s to make decisions. But there is a need to demonstrate that human tenability limits and risk acceptance criteria's regarding fire toxicity are beneficial, or if it only introduces a perceived level of safety for the asset owner making decisions regarding fire safety in public transportation systems.

References

Carvel, R., Beard, A.: The Handbook of Tunnel Fire Safety. Thomas Telford, London (2005)

CPR: Methods for the determination of possible damage to people and objects resulting from releases of hazardous materials "Green book", Hague, the Netherlands, Directorate -General of Labour of the Ministry of Social Affairs and Employment (1992)

Dix, A.: Safety standards for road and rail tunnels - a comparative analysis. In: 2nd International Conference Tunnel Safety and Ventilation, Graz (2004)

Duffé, P., Marec, M.: Task force for technical investigation of the 24 March 1999 fire in the Mont Blanc vehicular tunnel, Minister of the Interior - Ministry of Equipment, Transportation and Housing (1999)

Hurley, M.J.: SFPE Handbook of Fire Protection Engineering. Springer, New York (2016)

Ingason, H., Li, Y.Z., Lönnermark, A.: Tunnel Fire Dynamics. Springer, New York (2014)

ISO: Estimation of the lethal toxic potency of fire effluents. The International Organization for Standardization, Geneva (2015)

Miclea, P.C., Chow, W.K., Chien, S.W., Li, J., Kashef, A.H., Kai, K.: International tunnel fire-safety design practices. ASHRAE J. **49**, 50–60 (2007)

NFPA, NFPA 502: Standard for road tunnels, bridges, and other limited access highways. National Fire Protection Association, Quincy, Mass (2017)

NFPA, NFPA 101: Life Safety Code. National Fire Protection Association, Quincy, Mass (2018)

Njå, O., Kuran, C.: Erfaringer fra redningsarbeidet og selvredningen ved brannen i Oslofjord-tunnelen 23. juni 2011 [Experiences from the rescue work and the self rescue in the Oslofjord tunnel fire 23 June 2011]. International Research Institute of Stavanger, Stavanger (2015)

Nævestad, T.-O., Ranestad, K., Elvebakk, B., Meyer, S.F.: Kartlegging avkjøretøybranner i norske vegtunneler 2008-2015 [Vehicle fires in Norwegian road tunnels 2008-2015]. Institute of Transport Econoimics, Oslo (2016)

Seconds from disaster - S01E09 - fire on the ski slope, Directed by SEMPIO, T. Youtube (2013)

Stec, A.A., Hull, R. (eds.): Fire Toxicity. CRC Press, Oxford (2010)

SVV: Håndbok N500 Vegtunneler [Manual N500 Road Tunnels]. Statens vegvesen, Oslo (2016)

Tijms, H.C.: Stochastic Models: An Algorithmic Approach. Wiley, Chichester (1994)

Construction Sites as Shared Workplaces – An Occupational Safety and Health Profile Based on Workplace Inspection Reports

Päivi Kekkonen[1(✉)], Arto Reiman[1], Seppo Väyrynen[1], and Hanna-Kaisa Rajala[2]

[1] Industrial Engineering and Management, University of Oulu, Oulu, Finland
{paivi.kekkonen, arto.reiman, seppo.vayrynen}@oulu.fi
[2] Occupational Safety and Health Section, Regional State Administrative Agency of Northern Finland, Oulu, Finland
hanna-kaisa.rajala@avi.fi

Abstract. A *shared workplace*, according to Finnish legislation, is a workplace in which one employer exercises the main authority while additional employers or self-employed workers operate simultaneously or successively in such a way that their work may affect the safety or health of other employees. This study aimed to form a holistic view of the occupational safety and health (OSH) challenges facing shared workplaces within the construction industry. The material consisted of randomised OSH inspection reports (N = 79) from the Regional State Administrative Agency. Reports were analysed to gain information about the observed deficiencies. In the analysis, the reports were categorised based on the holistic work system model (ISO 6385, 2016). The analysis carried out in this study resulted in the recognition of common challenges at shared workplaces within the construction industry. Using this method, an individual observation profile for each industry branch can be formulated. Such profiles can be used in the planning of industry-specific inspection checklists for the supervision of OSH as well as in developing the OSH management at shared workplaces.

1 Introduction

A *shared workplace* is defined in Finnish legislation as a workplace in which one employer exercises the main authority while additional employers or self-employed workers operate simultaneously or successively in such a way that the work may affect the safety or health of other employees (Occupational Safety and Health Act 738/2002). The European Union (EU) directives on safety and health at work address a situation similar to that of a shared workplace, but they do not use a specific term for it, and they present less detail in the OSH Framework directive (Directive 89/391/EEC). Both the EU and Finnish legislation require that employers and self-employed workers at such workplaces ensure in adequate cooperation that their activities do not endanger employees' safety and health.

Shared workplaces occur in several industries but are most common in the construction, manufacturing and transportation and storage sectors. In the construction

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 416–422, 2020.
https://doi.org/10.1007/978-3-030-48021-9_46

industry, almost all workplaces can be considered shared workplaces due to the fact that, at large construction sites, there are always employees of several different employers working on their specific areas of expertise (Häkkinen and Niemelä 2015). In process and manufacturing industries, it is also very common that certain tasks, such as those related to maintenance, are carried out by contractors and their employees. Within the transportation and storage sector, shared workplaces often occur in the hubs of transportation routes, such as ports, airports and terminals and storages of ground transport (Turunen et al. 2015; Teperi 2012; Reiman et al. 2015).

Despite the widespread occurrence of shared workplaces and increased outsourcing of services and other tasks, which has become more and more common in many industries over the last few decades, the concept of a *shared workplace* is not very well known internationally, and research on the safety of shared workplaces is sparse (Rantanen et al. 2007; Nenonen 2012). However, the challenges related to safety management in situations in which cooperation and collaboration among several employers and their employees are needed are similar in all industrialised countries (Rantanen et al. 2007). The challenges at shared workplaces often include service provider selection, safety management resource availability, hazard identification, communication, working culture, and employee competence and training (Nenonen 2012). Special needs related to shared workplaces in process industry companies' sites was a main contributor to the development of the Health, Safety, Environment, Quality Assessment Procedure (HSEQ AP), which is now widely applied in Finland (Väyrynen et al. 2012).

The work system model, which can be used to examine work and its elements, consists, in its basic form, of the following elements: the person or employee, his or her work task, the tools and technologies he or she uses to accomplish the task, the work environment and the organisation in which the work takes place (Smith and Sainfort 1989; Carayon and Smith 2000; Carayon 2009). Depending on the input into the work system and the interplay of its elements, the work that is carried out in various processes leads to positive and/or negative outcomes. Examples of positive outcomes include the productivity, health and wellbeing of the employees, while negative outcomes include stress, accidents, discomfort, absence from work and loss of time. Employees can be seen as important assets to an organisation, meaning their safety and health is essential to the performance of the whole organisation.

In this study, occupational safety and health (OSH) inspection reports were analysed to gain a holistic view of the OSH situation at shared workplaces within the construction industry. The challenges, special characteristics, situations and phenomena related to OSH, viewed from the standpoints of different actors within the shared workplace, were the targets of interest in this study. The work system was used as a framework to compartmentalise the observations under analysis. The aim was to answer the following research questions:

1. What kind of OSH challenges exist at shared workplaces within the construction industry?
2. What special characteristics, situations and phenomena exist at shared workplaces within the construction industry?

2 Study Design and Methods

The materials analysed within this study were OSH inspection reports from the Regional State Administrative Agency of Finland. These inspections were carried out from 2012 to 2016. The analysed OSH inspection reports (N = 79) were chosen randomly, but selection was done in such a way that all the inspections were targeted at workplaces in the construction industry and each inspection report included at least one observed deficiency that led to a written advice or improvement notice by the agency. The cases were regarded as construction industry cases if the inspection took place at a construction site, even if the employer who was targeted in the inspection represented another branch of industry. This decision was made because the branch of the worksite being inspected was considered a main contributor to the possible hazards and deficiencies observed during the inspection.

Regional State Administrative Agencies' OSH inspections concentrate on safety risks, the management of overload and the minimum conditions of employment, according to the supervision guidelines issued by the Ministry of Social Affairs and Health. The inspections also aim to help workplaces develop their OSH functions and work conditions. The situation at a workplace is observed on the basis of both discussion and documents, as well a visit to the workplace. In cases where the inspector observes matters that are contrary to OSH legislation, he or she issues written advice and improvement notices that are recorded in the inspection report.

In this study, observations within OSH inspection reports that led to written advice or improvement notices were analysed according to different thematic categories. The categories were formed in a larger analysis spanning 200 inspection reports from various branches of industries, which resulted in 61 separate observation categories. An open coding approach was used to form themes arising from the material (Flick 2009). The categories were further divided according to the elements of the work system model, namely, organisation, employee, task, tools and technology, and work environment.

The categorisation was carried out by one researcher (PK), and the final set of categories was decided upon by the researcher and the expert from the Regional State Administrative Agency (H-KR). NVivo 11 Pro software, which was designed for analysing qualitative data, was used in the analysis. Each of the observations could belong to more than one category, and one inspection report could have more than one observation that belonged to a certain category. On the basis of this analysis, it was possible to recognise the categories on which the observations at shared workplaces within the construction industry were focused.

3 Results

The analysis carried out on the OSH inspection reports aimed to answer the following research questions: "What kind of OSH challenges exist at shared workplaces within the construction industry?" and "What special characteristics, situations and phenomena exist at shared workplaces within the construction industry?" The results are presented in the following sections.

3.1 Challenges at Shared Workplaces Within the Construction Industry

The observation categories that contain the most observations can be regarded as the main challenges experienced at shared workplaces within the construction industry. These are presented as a percentage of the total number of observations in Table 1. Four categories (i.e., deficiencies in planning, use of personal protective equipment, fall hazard and access ways) could be clearly separated from the rest of the categories based on their prominence in the analysed material.

Table 1. Categories with most observations, presented as a percentage of the total number of observations.

Observation category	%
Deficiencies in planning	12.0
Use of personal protective equipment	11.1
Fall hazard	10.5
Access ways	9.4

3.2 The Observation Profile

The observation profile that was based on the OSH inspection reports about construction worksites is presented in Fig. 1. The profile presents the percentage rates of all the observations in each category. The observation categories contained anywhere from 0 to 55 observations from the workplace inspection reports. For clarity, the observation categories are divided under the elements of the work system (i.e., organisation, employee, task, tools and technology, and work environment). Due to the large number of categories under the organisation element, these categories are further divided under the themes of human resources and documentation, safety management, occupational health and safety, occupational health service, and general practices.

The observation profile offers visual insight, both into the categories in which there are a lot of observations and into the categories in which there are none or only a small proportion of observations. To use the information provided by the gaps in the observation profile, background information is needed about the original checklists used in the workplace inspections that formed this material. However, observation profiles, such as the one presented in Fig. 1, can be used to develop checklists that target the special challenges of the industry branch in question (in this case, the construction industry).

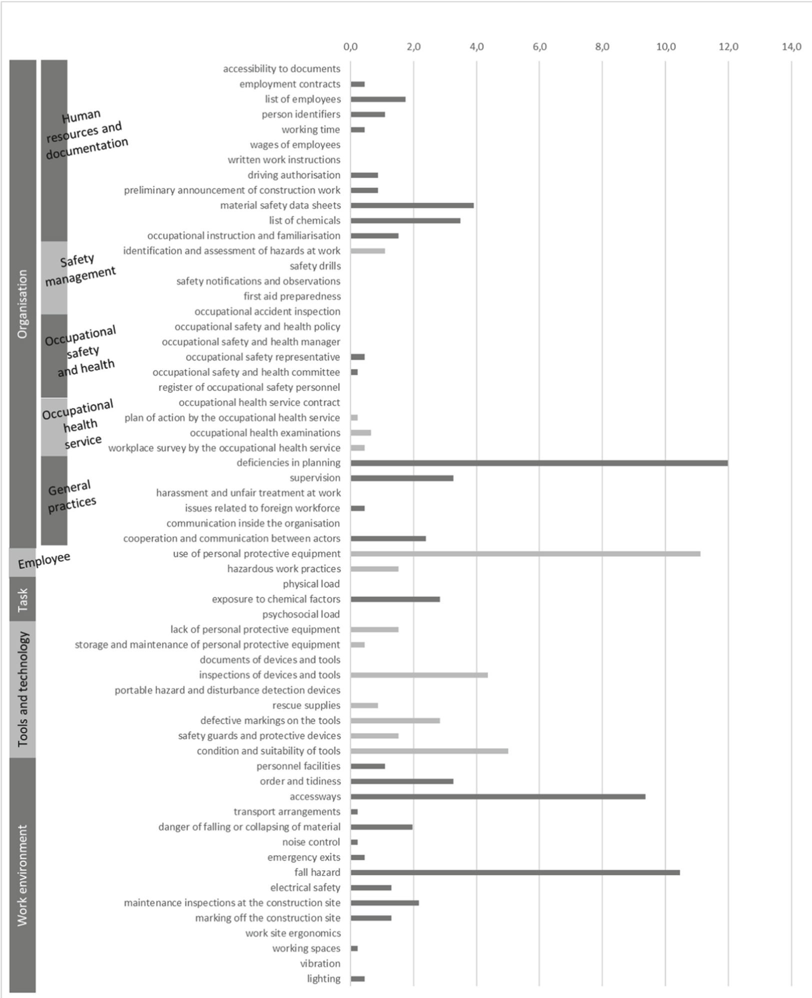


Fig. 1. Observation profile presenting the percentage rates of observations in each category. For clarity, the categories are divided under the elements of the work system, with further classifications of various themes under the organisation element.

4 Discussion

Looking at the results on categories related to organisation, a clear peak in deficiencies in planning can be seen. This area of safety in the construction industry has previously been identified as a challenge and a common factor contributing to accidents (Häkkinen and Niemelä, 2015; Rantanen et al., 2007; Lind-Kohvakka 2015). Observations related to access ways and fall hazards were also strongly present in the material. These are also common factors contributing to accidents in the construction industry (Häkkinen and Niemelä 2015).

The material used in this study—namely the workplace inspection reports—contains observations that have been detected as matters contrary to Finnish OSH legislation. Thus, the view represented in these results is based on the minimum level set by the legislation, and possible challenges and good practices that could be identified at similar worksites using other methods are not necessarily present in these results. The inspection reports were written by individual OSH supervisors, causing some variance in the level of detail in the reports; this variance may also be reflected in the analysis carried out by the researcher.

The material is also comprised solely of reports of workplace inspections targeted at shared workplaces. In the case of the construction industry, however, the situation is different than it is in many other branches of industry, as almost all construction sites are shared workplaces in which there are employees of several employers working simultaneously or successively in such a way that their work may affect other employees' safety or health. This said, the results can be seen as comparable to the results of earlier research related to OSH in construction worksites.

5 Conclusions

This study aimed to form a holistic view of the OSH challenges at shared workplaces within the construction industry. The concept of a *shared workplace* is not very well known or widespread internationally, although the situation of having several employers' employees working at the same worksite is identified in the EU legislation. Deeper, human-centred insight into the OSH challenges at shared workplaces could offer ways in which to tackle OSH issues in today's complex organisational environment, where shared workplaces and situations similar to these have become more and more widespread due to, for example, to outsourcing and networking.

The method presented in this study offers a way to formulate an individual observation profile for each industry branch. Such profiles can be useful in the planning of industry-specific inspection checklists for the supervision of OSH. In addition to OSH authorities, this information could also be utilised by experts of industries in which shared workplaces occur. In the construction industry - and other industries as well - identifying the special characteristics and challenges related to shared workplaces could also be beneficial to the companies themselves as they develop their OSH management.

References

Carayon, P.: The balance theory and the work system model … twenty years later. Int. J. Hum. – Comput. Interact. **25**(5), 313–327 (2009). https://doi.org/10.1080/10447310902864928

Carayon, P., Smith, M.J.: Work organization and ergonomics. Appl. Ergon. **31**(6), 649–662 (2000)

Council of the European Communities: Council directive of 12 June 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work

(89/391/EEC). OJL, vol. 183, pp. 1–14 (1989). http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:01989L0391-20081211&from=EN
Flick, U.: An Introduction to Qualitative Research, 4th edn. Sage Publications Ltd, London (2009)
Häkkinen, K., Niemelä, V.: Accident sources and prevention in the construction industry - Some recent developments in Finland. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management. Solutions and Industrial Cases. Springer, Switzerland (2015)
ISO 6385: Ergonomics principles in the design of work systems. International Organization for Standardization (2016)
Lind-Kohvakka, S.: Application of accident information to safety promotion - Case industrial maintenance. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management. Solutions and Industrial Cases. Springer, Switzerland (2015)
Nenonen, S.: Implementation of Safety Management in Outsourced Services in the Manufacturing Industry. Tampere University of Technology, Tampere (Finland), Publication 1023 (2012)
Ministry of Social Affairs and Health: Special situations of organising work. In: Occupational Safety and Health Act, 738/2002. Unofficial translation (2002). http://www.finlex.fi/en/laki/kaannokset/2002/en20020738.pdf
Rantanen, E., Mäkelä, T., Sauni, S., Lappalainen, J., Piispanen, P.: Yh-teisten työpaikkojen työturvallisuus. TOT-raporttien analyysi. (Occupational safety of shared workplaces. Analysis of TOT reports). VTT, Tampere, Finland (2007)
Reiman, A., Väyrynen, S., Putkonen, A.: Truck drivers' work systems in environments other than the cab - A macro ergonomics development approach. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management. Solutions and Industrial Cases. Springer, Switzerland (2015)
Smith, M.J., Sainfort, P.C.: A balance theory of job design for stress reduction. Int. J. Ind. Ergon. **4**(1), 67–79 (1989)
Teperi, A.-M.: Improving the mastery of human factors in safety critical ATM organisation. Faculty of Behavioural Sciences, University of Helsinki. Yliopistopaino Unigrafia, Helsinki, Finland (2012)
Turunen, H., Väyrynen, S., Lehtinen, U.: Introducing a scenario of a seaport's HSEQ framework: review and a case in Northern Finland. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management. Solutions and Industrial Cases. Springer, Switzerland (2015)
Väyrynen, S., Koivupalo, M., Latva-Ranta, J.: A 15-year development path of actions towards an integrated management system: description, evaluation and safety effects within the process industry network in Finland. Int. J. Strateg. Eng. Asset Manag. **1**(1), 3–32 (2012)

Common HSEQ Performance Improvement Areas Among Industrial Suppliers: A Study of Audits in a Finnish Industrial Cluster Network

Henri Jounila(✉), Arto Reiman, Osmo Kauppila, and Tero Savola

Industrial Engineering and Management, University of Oulu, Oulu, Finland
{henri.jounila, arto.reiman, osmo.kauppila}@oulu.fi,
tero.savola@student.oulu.fi

Abstract. Supplier management is a key issue for many large industrial companies from both performance and sustainability viewpoints. In this study, a supplier audit system in a Finnish industrial cluster is examined. The cluster has created a supplier assessment procedure, and roughly 200 supplier HSEQ audits have taken place. The aim of this study is to analyse a sample of these audits to identify common improvement areas among supplier companies. Improvement suggestions taken from the audits are classified into five themes and 39 categories. The results paint a picture of the most common HSEQ improvement areas for suppliers.

1 Introduction

Large industrial companies face the challenging task of managing a vast network of suppliers. For example, nurturing a good safety culture for every individual and organisation working on site requires a great deal of attention on the part of industrial companies. Adopting an even broader perspective, the challenge is to ensure that supplier companies act in a responsible way, particularly in the areas of Health and Safety but also in terms of Environment and Quality (HSEQ). This implies that socially responsible industrial purchasing companies must extend their management system to take stock of their suppliers' HSEQ performance (Kauppila et al. 2015; Koivupalo et al. 2015; Väyrynen et al. 2012). One practical reason for mutual interest in HSEQ performance is that both the purchaser and the supplier work together in shared workplaces. Secondly, industrial companies prefer operating with suppliers who handle their social-, economic- and environmental affairs in a sustainable manner. Stakeholders also often require this (Hofmann et al. 2014; Meixell and Luoma 2015; Zink 2014). Purchaser-supplier HSEQ activity has been shown to create long-term positive results for both parties (Väyrynen et al. 2012). A well-developed safety management system has a positive effect on safety, as well as competitiveness (Fernández-Muñiz et al. 2009), especially in certified organisations (Vinodkumar and Bhasi 2011). Thus, HSEQ management is justified and profitable. Generally, management systems are often intended to be simple, well-guided and efficient (Dujim et al. 2008). In this article, one

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 423–430, 2020.
https://doi.org/10.1007/978-3-030-48021-9_47

simple though challenging management tool is introduced, and on the basis of this tool, the common HSEQ performance improvement areas of industrial suppliers are analysed.

A group of twelve large Finnish industrial purchasing companies have created a HSEQ cluster network over a developmental process spanning over a decade. This HSEQ cluster is open to all interested purchasing companies who meet certain criteria and accept the terms of membership. The HSEQ assessment procedure (HSEQ AP) is intended for all suppliers who want to be assessed based on their own initiative or whose assessment is requested on the purchasers' initiative. The aim of HSEQ AP is to provide feedback and thus improve the HSEQ issues and productivity of the companies. In addition, through the HSEQ AP, companies can develop their HSEQ operations and management systematically (anon 2018). As a concrete tool, an assessment procedure for auditing supplier HSEQ capabilities has been created. In the HSEQ AP, the HSEQ performance of suppliers is assessed, audited and scored with specified criteria covering 41 topics. The assessment criteria are divided into subject areas following the EFQM Excellence Model structure. To this day, a total of around 200 HSEQ AP audits have taken place.

2 Methods

In this study, the database of HSEQ audits was utilised. In total, 48 audits from 2015 to 2017 were analysed in order to answer the following research question: *"What kind of common improvement areas can be identified from HSEQ AP supplier audit results?"* The research question was answered by collecting, analysing and classifying improvement suggestions and deviations from the audit database.

3 Results

Altogether, 456 improvement suggestions and deviations (hereafter referred to as simply improvement suggestions) were gathered from the audit documents. The categorisation resulted in five main themes and 39 categories. The main themes were as follows: 1. Operations planning and management (n = 155); 2. Occupational safety management (n = 113); 3. Customers, stakeholders and suppliers management (n = 83); 4. Human resource management (n = 64); and 5. Environmental and chemical safety (n = 41).

3.1 Operations Planning and Management

The "Operations planning and management" theme included the most suggestions (Table 1). Within it, the largest category was the use of indicators. This included a large number of shortcomings regarding indicators and follow-up, such as the number and frequency of incident reports, sickness absences, environmental indicators, delivery reliability, reclamations and generally proactive indicators: *"There was no evidence of comprehensive follow-up of indicators." "There were no incident report frequency,* *no LTA, no LTIF indicators."* The management should better monitor and use the indicators to control their work and communicate the results: *"Proactive safety indicators, e.g., safety observations, should be used and followed up on. The company should activate staff to make observations, and create processes for recording and follow-up."*

Table 1. The categories within "Operations planning and management"

Category	Number of improvement suggestions
Use of indicators	45
Describing the processes	23
Principles and policies	22
Business plan and long-term planning	15
Meetings and issues to be dealt with, management activities	15
Ethics and responsibility	10
Certification and standards	9
Defining responsibilities	8
Organisation chart	5
Planning products and services	3

Many kinds of shortcomings related to processes emerged in the "Describing the processes" category. High-level process mapping and more focus on key processes was required: *"A description of the order-to-deliver process would be necessary." "It is recommended to draw up a process map showing key (core) processes."* To sum up the "Principles and policies" category, more attention to creating and communicating these was requested, particularly related to safety, environment and quality principles. In other categories, the improvement suggestions concerned, for example, long-term planning and more systematic management review practices: *"The business plan could not be verified… With the help of the business plan, there is a good chance of crystallising guidelines that have been jointly defined." "It is recommended that the company's key persons have a shared meeting practice in which they regularly go through the issues in accordance with the standard agenda that they agreed to together (personnel, customers, economy, development of operations, etc.)… Memos should be drawn up from the meetings… The communication of key decisions to staff should also be agreed upon."* In addition, issues such as developing operations for certification (mainly quality), attention to ethical values, defining the acquisition of information and HSEQ responsibilities and creating an organisation chart were also pointed out: *"It is important to draw up an organisation chart in which the most important HSEQ responsibilities are also presented."*

3.2 Occupational Safety Management

The "Occupational safety management" theme consisted of 113 improvement suggestions divided into ten categories (Table 2). In the category regarding improvements to the recording and handling of accidents and incidents and communicating improvement measures, seventeen suggestions were proposed. In some cases, the related procedures were completely missing, for example: *"They do not have their own procedure for investigating and handling accidents and incidents…"* In the orientation category, shortcomings in the documentation and the content of the orientation, as well as in renewing the orientation, emerged: *"The documentation of the orientation ➜ the recording of the orientation into the ERP." "Based on a new person's orientation, it is necessary to draw up a signed document…"*

Table 2. The categories in "Occupational safety management"

Category	Number of improvement suggestions
Procedure for incidents and accidents	17
Induction	17
Tools and facilities	17
Occupational safety training	13
HSEQ observations	11
Development of safety	10
HSEQ risk management	9
Occupational safety and health	7
Occupational safety responsibilities	7
Tidiness and order	5

The tools and facilities category included various suggestions relating to work equipment and facilities: *"The individualised instruction signs must be affixed to the machines, including those regarding the protective equipment." "The company must draw up a list of all assets requiring an annual inspection…" "One must ensure the functionality of the emergency stop devices in equipment."* Occupational safety training needs were recognised, particularly for superiors and foremen, as well as for staff: *"Superiors and workers have not received safety training, except for occupational safety card training." "Superiors must be aware of their responsibilities and powers. That's why one must go to a course where these things are gone over."* The improvement suggestions concerning HSEQ observations applied to the entire process, from making observations to utilising the results. Above all, the daily but also the broader identification of work hazards and risk assessment occurred within the categories related to developing safety and HSEQ risk management: *"There was no evidence of the proactive identification of work hazards." "A risk mapping should be carried out to assess environmental risks and hazards."* The shortcomings in occupational health and safety policy and in defining the responsibilities of superiors were emphasised in their respective categories: *"Evidence of the occupational health and* *safety policy is requested." "It is also recommended to ensure that all safety responsibilities are individualised within the job descriptions of superiors."*

3.3 Customer, Stakeholder and Supplier Management

The categories within the "Customer, stakeholder, and supplier management" theme are shown in Table 3. The most suggestions were found within the customer satisfaction and feedback category. These suggestions stated that customer satisfaction must be measured and followed up on and that satisfaction discussions should be organised with the customer. Customer feedback must be collected, documented, processed, and analysed; e.g., *"There is no systematic way of measuring customer satisfaction." "Customer feedback is not collected systematically, and feedback is not documented or analysed."* Related to the previous category, suggestions regarding reclamations were recorded as well.

Table 3. The categories within "Customer, stakeholder and supplier management"

Category	Number of improvement suggestions
Customer satisfaction and feedback	33
Supplier selection	15
Ensuring the responsibility of suppliers	11
Stakeholder management	7
Supplier safety	5
Stakeholder communication	5
Reclamations	5
Monitoring the development of legislation	2

On the basis of the these categories, more systematic supplier selection and management was required: *"Supplier management is memory-dependent and based on subjective experience. The company must set up a list of suppliers and principles for supplier selection…" "Suppliers' acceptance criteria or assessment criteria could not be verified…" "A procedure should be created for the assessment of suppliers' environmental responsibility." "A procedure must be specified to ensure supplier safety."* In the categories related to stakeholders, suggestions regarding mapping stakeholders and their needs, as well as stakeholder communication, arose: *"In addition to stakeholder analysis, a systematic procedure for communications and feedback collection must be set up."*

3.4 Human Resource Management

Categories in the "Human resource management" theme are shown in Table 4. The category that attracted the most improvement suggestions was registry development: suppliers must create a training register and record workers' qualifications and orientations as well. The well-being at work category included several mentions of job satisfaction measurement: *"There is no evidence of the systematic measurement of job satisfaction."* The incentives and rewards category, as well as the initiatives, development proposals, and development discussions category, included encouragement for initiatives and the creation of an initiative system. In addition, suppliers must communicate reward principles *"It is necessary to develop the initiative system and the handling process. Consider rewards."* The importance of writing down decisions and communicating them, as well developing a staff backup system, emerged from four last categories.

Table 4. The categories in "Human resource management"

Category	Number of improvement suggestions
Registry development	13
Well-being at work	10
Incentives and rewards	9
Initiatives, development proposals and development discussions	9
Knowledge and data management	8
Internal communication	6
Monitoring and management of working time	5
Training and development	4

3.5 Environmental and Chemical Safety

The "Environmental and chemical safety" theme drew 41 improvement suggestions divided into three categories (Table 5). The environmental matters category contained various management issues, the most common being impact mapping*: "There was no evidence of environmental management."* In the Chemicals category, three areas were highlighted: chemical storage issues, shortcomings in the list of chemicals and shortcomings regarding safety data sheets, for example: *"There was no person in charge appointed. There was no list of chemicals or list of safety data sheets."* The waste sorting category included various shortcomings, such as a lack of instructions regarding waste and statistics regarding the volume of waste.

Table 5. The categories within "Environmental and chemical safety"

Category	Number of improvement suggestions
Environmental matters	20
Chemicals	14
Waste sorting	7

4 Discussion and Conclusion

In this article, the most common improvement suggestions from HSEQ audits were considered and categorised. By analysing individual categories in more detail, more exact targets for development were determined. On the basis of this study, the following targets can be recommended and they should be considered for the development of supplier companies. *Theme 1*: developing and following up on indicators and utilising them as a management tool; drawing up of a process map describing the most important processes, especially the order-to-deliver process; describing and communicating principles, especially safety, environment and quality principles; drawing up a long-term plan and a business plan; organising meetings of the management and writing agendas and memos; developing operations for certification; drawing up or assessing ethical values; defining the acquisition of information and HSEQ responsibilities and drawing up the organisation chart. *Theme 2:* creating or improving the recording, handling and utilising accidents, incidents and HSEQ observations and communicating the appropriate measures to take in response; identifying and assessing work hazards; reviewing shortcomings in the documentation of the orientation and the content of the orientation, as well as the renewal of the orientation; enforcing occupational safety training, especially to the superiors, and defining and increasing awareness of safety responsibilities. *Theme 3:* measuring and following up on customer satisfaction; organising satisfaction discussions with the customer; defining, collecting, documenting, processing and examining customer feedback; paying attention to supplier selection criteria and how they affect selection; mapping stakeholders and their needs and developing stakeholder communication. *Theme 4*: creating a training register; creating systematic job satisfaction measurements; creating an initiative system and handling processes; improving incentives and rewards processes and communicating decision-making. *Theme 5:* performing environmental impact mapping; developing chemical storage and recognising shortcomings in the list of chemicals and safety data sheets.

It should be noted that in this classification system, many improvement suggestions could have been placed within several categories, but each of them was classified within only one category. The purpose of the research was to identify common improvement areas, and this was accomplished. The improvement suggestions that were the most commonly mentioned in the audit reports were represented in this research. These can be utilised in the management and control of the suppliers and supplier audits.

References

anon: HSEQ Assessment Procedure (HSEQ AP) (2018). https://www.hseq.fi/index.php?p=BrieflyinEnglish

Duijm, N.J., Fiévez, C., Gerbec, M., Hauptmanns, U., Konstandinidou, M.: Management of health, safety and environment in process industry. Saf. Sci. **46**(6), 908–920 (2008)

Fernández-Muñiz, B., Montes-Peón, J.M., Vázquez-Ordás, C.J.: Relation between occupational safety management and firm performance. Saf. Sci. **47**(7), 980–991 (2009)

Hofmann, H., Busse, C., Bode, C., Henke, M.: Sustainability-related supply chain risks: conceptualization and management. Bus. Strategy Environ. **23**(3), 160–172 (2014)

Kauppila, O., Härkönen, J., Väyrynen, S.: Integrated HSEQ management systems: developments and trends. Int. J. Qual. Res. **9**(2), 231–242 (2015)

Koivupalo, M., Junno, H., Väyrynen, S.: Integrated management within a Finnish industrial network: steel mill case of HSEQ assessment procedure. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management - Solutions and Industrial Cases, Production & Process Engineering, pp. 41–67. Springer, Heidelberg (2015)

Meixell, M.J., Luoma, P.: Stakeholder pressure in sustainable supply chain management: a systematic review. Int. J. Phys. Distrib. Logist. Manag. **45**(1/2), 69–89 (2015)

Vinodkumar, M.N., Bhasi, M.: A study on the impact of management system certification on safety management. Saf. Sci. **49**(3), 498–507 (2011)

Väyrynen, S., Koivupalo, M., Latva-Ranta, J.: A 15-year development path of actions towards an integrated management system: description, evaluation and safety effects within the process industry network in Finland. Int. J. Strateg. Eng. Asset Manag. **1**(1), 3–32 (2012)

Zink, K.J.: Designing sustainable work systems: the need for a systems approach. Appl. Ergon. **45**(1), 126–132 (2014)

Occupational Health and Safety in the Trucking Industry – Current Trends and Future Challenges

Arto Reiman(✉) and Seppo Väyrynen

Industrial Engineering and Management, University of Oulu, Oulu, Finland
{arto.reiman, seppo.vayrynen}@oulu.fi

Abstract. Professional truck drivers face various kinds of challenges during their workday. A high prevalence of work-related musculoskeletal disorders and high accident rates are associated with the trucking industry. In addition, various differ- ent psychosocial stressors affect truck drivers' work ability. Accidents, disorders, and stressors all affect working careers. There is both a constant need for a skilled new workforce, but also a need for prolonging working careers. Some characteristics can be identified on truck drivers' work; the drivers mainly work alone and their work contains static work postures while sitting and physical activities while working outside the cab. While working, the driver often faces opportunities for unethical and unsafe actions to ease the workload. This poses challenges to the occupational health and safety (OHS) management. This article provides a scoping review of the risks and hazards that the professional truck drivers face while working. Both driving and non-driving work activities are covered. Special attention is paid to selected new modes of transportation and a discussion is held on the possible OHS challenges that they may bring along. High capacity transports (HCT) are discussed as an emerging mode of road transportation that enables larger loads to be freighted and intermodal transportations (IMT) are used increase the efficiency of the transportations by combining different transportation modes. However, very little attention is being paid in the current OHS literature on the possible adverse OHS effects concerning the driver.

1 Introduction

Occupational health and safety (OHS) is an ambiguous concept, that can be associated to various different determinants, depending on the definition. In this article, we use a holistic definition that includes not only the aspects that are traditionally associated to OHS, i.e. occupational hygiene, accidents and injuries, and ergonomics and human factors, but also the concept of well-being at work. All these elements include objective and subjective aspects and can be considered interlinked and partly overlapping to each other (e.g. Schulte and Vainio 2010).

OHS can be identified and discussed as an asset and as a strategic element, affecting companies' activities and economic performance (Dul et al. 2012; Zweetslot et al. 2013). Improper OHS practices and processes lead to negative consequences like decreased work performance and quality and accidents and incidents. The costs related

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 431–438, 2020.
https://doi.org/10.1007/978-3-030-48021-9_48

to these can be substantial, reaching from the employee and employer levels to the society level (Rikhardsson 2004). Well-being at work on the other hand is often associated to more complex subjective determinants and adverse effects may include aspects such as stress, job dissatisfaction, and non-stable work force.

1.1 OHS Challenges at Road Transportations

OHS in road transportations has been studied by various different authors (van der Beek 2012). Truck drivers have been identified as lone workers who cannot be fully monitored by their employers (Huang et al. 2013). For that reason, they often face opportunities for unethical and unsafe actions to ease and quicken their work. Recently for example Reiman et al. (2015), Reiman et al. (2018), Murphy et al. (2018) and Anderson et al. (2017) have discussed on OHS challenges related to truck drivers' work outside the cab at short haul (SH) delivery transportations and Chandler et al. (2017) have reached that discussion to long haul (LH) transportations. Research literature covers traditional LH and SH freight transportation modes, however very little attention is given on the new modes of transportation.

In this article, we raise two modes of road transportation as examples that have been given increasingly more attention in the Nordic countries and in Europe. Both of them are associated with higher efficiency and being a step towards more sustainable transportations (Ye et al. 2014; Bergqvist and Monios 2016). Firstly, we raise up High Capacity Transportations (HCT). HCTs – as understood in this context – are performed by longer and heavier high capacity vehicles (or mega trucks) whose weights and/or dimensions are outside the permissions and regulations. The EU standard for heavy vehicles is 18.75 metres and 40 tonnes. In Sweden and in Finland, for instance, vehicles that reach the maximum length of 25.25 metres and weigh up to 60 tonnes are in a test use. (Ye et al. 2014; Bergqvist and Monios 2016; Sandin 2016). Secondly, in intermodal transportation (IMT) more than two modes of transportation are used to freight transportations. For instance, road transportations may be connected to rail, air or sea transportations to reach higher efficiency. There are some indications that the tendency for severe road crashes rises during the wintertime for HCTs (Sandin 2016), however very little – almost none – research has been conducted about OHS at HCTs and IMTs.

1.2 Objectives

Our aim here is to conduct a scoping review (Grant and Booth 2009) in order to facilitate future discussions related to OHS and well-being at work at driving and non-driving tasks for certain modes of road transportation. Especially we pay attention to discussing these challenges from the point of view of Nordic work environmental conditions and social sustainability. We attempt to build new knowledge that reviews and synthesizes the existing literature on the subject. The objective behind this is to summarize findings and to determine research gaps and targets for empirical research in the future.

2 Methodology

2.1 Analysis Approach

To conduct our survey, we searched for literature in the following research areas. As search words describing the OHS challenges we used: “occupational safety”, “occupational health”, “well-being at work”, “ergonomics” and “human factors”. As search words describing the transportation modes we used: “long haul”, “short haul”, “high capacity transportation” and “intermodal transportation”. Search words were used as search terms to find scientific literature via the database of Scopus, a large abstract and citation database of peer-reviewed literature. As we aimed to focus on current and future trends we concentrated only in research literature that was published within last ten years, i.e. 2009–2018. The search was confined to the documents published in English. Additionally, relevant sources found in the previous studies carried out by the authors and in the reference lists of the existing sources were employed as found appropriate. The relevance of the literature was assessed based on the subject areas, the titles and the abstracts. Finally the articles that were deemed the most important were assessed based on the whole text.

2.2 Search Results

As the areas to be covered were diverse in nature, the search involved several independent searches, each of which concentrated on a specific area of interest in this paper. Table 1 describes the search word combinations and quantitative results of the searches.

Table 1. Search word combinations and the amount of documents found. OS = Occupational safety, OH = Occupational health, W-BW = Well-Being at Work, HF = Human factors, LH = Long haul, SH = Short haul, HCT = High capacity transportation, IMT = Intermodal transportation

	OS	OH	W-BW	HF	Ergonomics
LH	11	22	0	8	16
SH	2	4	0	2	3
HCT	0	0	0	0	0
IMT	0	0	0	0	0

As the results show, the amount of publications in the fields of OS, OH, W-BW, HF and Ergonomics at road transportations is rather limited. The publications focus on traditional long haul and short haul transportation modes, and no research literature was found concerning the two modes of transportation that are in our interest: HCTs and IMTs. The searches did not produce any results when W-BW was used as a search word. In the latter part of the analyses, we combine the categories OS and OH as OHS and HF and Ergonomics as HFE.

3 Results and Discussion

3.1 OHS & HFE Challenges Related to Drivers Work Tasks

OHS & HFE challenges related to truck drivers work tasks at LH and SH transportations are described briefly in Table 2. In general, driving related challenges were emphasised in LH transportations, whereas in SH transportations the challenges varied from driving to different actions that are performed outside the cab, i.e. to loading and unloading tasks at customers' premises and courtyards and to actions performed at home terminals, cargo spaces and other parts of the truck bodies.

Table 2. OHS & HFE challenges related to driving tasks at LH and SH transportation modes.

OHS & HFE challenges	
LH	• Long working hours
	• Non-compliance with hours-of-service rules
	• Tight schedules
	• Inadequate sleep duration and quality
	• Sleep fragmentation and sleepiness
	• Bad weather conditions
	• Heavy traffic and possibilities to road accidents/crashes
	• Frustration on other road users actions
	• Lack of social connections during working hours
	• Prolonged sitting and whole body vibration
	• Uncomfortable driving postures
SH	• Long working hours
	• Non-compliance with hours-of-service rules
	• Tight delivery schedules
	• Accident risks and ergonomic discomforts related to recurrent work activities outside the cab
	• Work environment hazards
	• Constantly changing temperatures – from warm cabs to out- door environments
	• Frustration on other road users actions
	• Inadequate and/or insufficient tools and devices

3.2 OHS & HFE Challenges in the Nordic Context

After identifying the most common OHS & HFE challenges based on the scientific literature, we categorised the challenges in three groups. The categorisation was formed inductively by identifying possible connections between the challenges. The categories selected were formed by the mode of work (driving or non-driving) and environmental conditions. A simplified categorization is presented in Fig. 1.

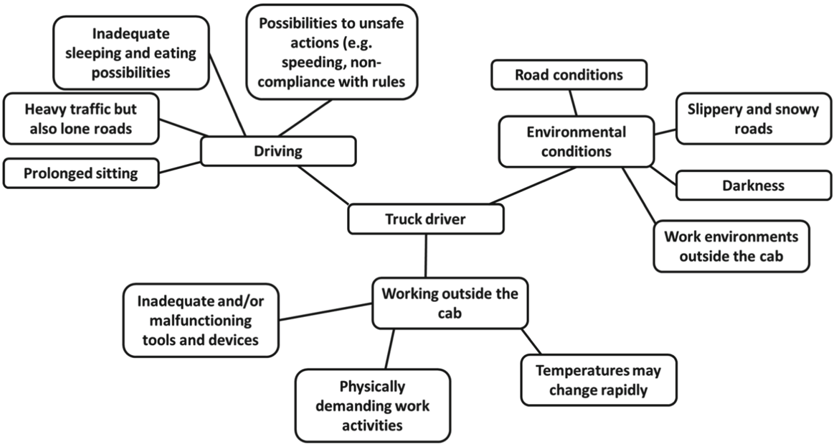


Fig. 1. OHS challenges in three categories.

Environmental conditions cover not only road conditions during driving but also different work environments outside the cab, i.e. the truck body structures and cargo spaces and customers' work premises, courtyards and public places, such as streets.

The Driving category illustrates the duality of the driving work. The drivers may face – even during a single work shift – both heavy traffic and lone roads. Driving requires concentration during the whole work shift. Long distances and lone roads may arouse ethical decision making situations; for instance whether or not violate working hours or speed limitations. Long distances are also associated to prolonged sitting. Prolonged sitting combined with whole-body vibration has been identified causing adverse health effects, such as low-back pain.

In addition to driving, truck drivers perform various work tasks outside the cab. In SH transportations these tasks are most often associated to unloading and loading activities at home terminals, cargo spaces and/or at customers' premises. As an opposite of sitting, working outside the cab includes various physical work activities, such as manual materials handling and movements in and out of the cargo space and cab. These are all associated with accidents and incidents in the literature.

Nordic environments bring some specialities to OHS management. For example, the driver may face icy and snowy grounds; from which parts are sanded whilst some parts are left unsanded. This raises possibilities to slipping and falling. In addition, the devices and tools used to ease and fasten the delivery work, such as roll cages, are usually not designed to be used in snow. Temperature changes affect the driver especially at wintertime. The driver constantly changes from warm cabs to outdoor environments; requiring continuous thermoregulatory adjustments.

3.3 Considerations on Possible OHS&HFE Challenges at IMTs and HCTs

As our scoping review reveals, no research has been made on the OHS & HFE challenges related to IMTs and HCTs. Based on our categorisation above, we emphasise the potential for OHS & HFE research in IMT and HCT contexts by providing nine new research challenges (see Table 3). We see that there is a need to study drivers' actions during driving. Especial interest should be given to cognitive analyses on driving behaviour and on perceived stress. While working outside the cab we highlight the need to provide more in-depth knowledge on possible new risks that the larger and heavier trucks may bring about and whether any new risks could be identified on the collaboration between different stakeholders while the transportation mode is changed. Further, we highlight the Nordic conditions and social sustainability in general as research topics.

Table 3. OHS research challenges at IMTs and HCTs.

	IMTs	HCTs
Driving	• Perceived stress; is the driver aware what s/he is transporting?	• Cognitive and working posture analyses on driving larger trucks
	• Perceived stress on load securing	• Perceived stress; does the driver fear that s/he is causing risk or discomfort to other road users or risk to the freight to be damaged
Working outside the cab	• OSH & HFE risks during the change of the transportation mode	• Do larger truck body structures bring out any new risks for occupational accidents?
Nordic conditions	• How are the Nordic special challenges related to OHS & HFE managed at long distance IMTs?	• How does the HCT driver perceive winter conditions at roads?
		• How do icy and snowy road conditions affect HCT drivers driving behaviour?

While for instance HCTs and IMTs are considered in general as sustainable transportation solutions, we highlight the social sustainability dimension as a topic for further research in the future. The third dimension of sustainability; social sustainability is often disregarded, while environmental and economic dimensions are paid more attention. Questions such as the drivers' perceptions on the reductions for drivers needed and general road safety perceived by other road users may arouse when social sustainability is discussed.

4 Conclusions

Fluent and safe transportations are a basic prerequisite for modern societies. Unsafe trucking can affect at various levels. In addition to road safety in general, unsafe actions while trucking can lead to various kinds of hazards to others as well. The drivers as frontline operators naturally confront different risks for hazards and accidents, but also other people such as by-passers' and other road users' health may be endangered. In addition to human suffering, unsafe actions and accidents give rise to monetary costs reaching from the employer and employee levels all the way to the society level. Thus, a need to improve OHS and HFE at transportations is justified. Truck drivers OHS and HFE challenges were summarized in this scoping review. Especial interest was given to two transportation modes, HCTs and IMTs. Nine new research challenges were identified as future research topics. These research challenges open up a discussion on possible upcoming future OHS and HFE problems. In addition, to reach the discussion outside OHS and HFE attention is paid to the social sustainability dimension.

References

Anderson, N., Smith, C.K., Byrd, J.L.: Work-related injury factors and safety climate perception in truck drivers. Am. J. Ind. Med. **60**(8), 711–723 (2017)

Bergqvist, R., Monios, J.: The last mile, inbound logistics and inter- modal high capacity transport – the case of Jula in Sweden. World Rev. Inter. Transp. Res. **6**(1), 74–92 (2016)

Chandler, M.C., Bunn, T.L., Slavova, S.: Narrative and quantitative analyses of workers' compensation-covered injuries in short-haul vs long-haul trucking. Int. J. Injury Control Saf. Promot. **24**(1), 120–130 (2017)

Dul, J., Bruder, R., Buckle, P., Carayon, P., Falzon, P., Marras, W.S., et al.: A strategy for human factors/ergonomics: developing the discipline and profession. Ergonomics **55**(4), 377–395 (2012)

Grant, M., Booth, A.: A typology of reviews: an analysis of 14 review types and associated methodologies. Health Inf. Libr. J. **26**(2), 91–108 (2009)

Huang, Y.-h., Zohar, D., Robertson, M.M., Garabet, A., Lee, J., Murphy, L.A.: Development and validation of safety climate scales for lone workers using truck drivers as exemplar. Transp. Res. Part F: Traff. Psychol. Behav. **17**, 5–19 (2013)

Murphy, L.A., Huang, Y.-h., Robertson, M.M., Jeffries, S., Dainoff, M.J.: A sociotechnical systems approach to enhance safety climate in the trucking industry: results of an in-depth investigation. Appl. Ergon. **66**, 70–81 (2018)

Reiman, A., Väyrynen, S., Putkonen, A.: Truck drivers' work systems in environments other than the cab—A macro ergonomics development approach. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management - Solutions and Industrial Cases, pp. 97–110. Springer, Cham (2015)

Reiman, A., Forsman, M., Målqvist, I., Parmsund, M., Lindahl Norberg, A.: Risk factors contributing to truck drivers' non-driving occupational accidents. Int. J. Phys. Distrib. Logist. Manag. **48**(2), 183–199 (2018)

Rikhardsson, P.M.: Accounting for the cost of occupational accidents. Corpor. Soc. Responsib. Environ. Manag. **11**(2), 63–70 (2004)

Sandin, J.: Effects of higher capacity vehicles on traffic safety in Sweden. In: HVTT14 Future Pathways Conference, Rotorua, New Zealand, 15–18 November (2016)

Schulte, P., Vainio, H.: Well-being at work – overview and perspective. Scand. J. Work Environ. Health **36**(5), 422–429 (2010)
van der Beek, A.J.: World at work: truck drivers. Occup. Environ. Med. **69**(4), 291–295 (2012)
Ye, Y., Shen, J., Bergqvist, R.: High-capacity transport associated with pre- and post-haulage in intermodal road-rail transport. J. Transp. Technol. **4**(3), 289–301 (2014)
Zwetsloot, G.I.J.M., Aaltonen, M., Wybo, J.-L., Saari, J., Kines, P., De Beeck, R.: The case for research into the zero accident vision. Saf. Sci. **58**(Oct), 41–48 (2013)

Information for Managers and Experts or Communication with All Employees Within Organisations and Networks – Case HSEQ

Seppo Väyrynen[1(✉)] and Heli Kiema-Junes[2]

[1] Faculty of Technology, University of Oulu, Oulu, Finland
seppo.vayrynen@oulu.fi
[2] Faculty of Education, University of Oulu, Oulu, Finland
heli.kiema@oulu.fi

Abstract. The planned interplay of many assets is needed for production. Efficient and effective interplay is enabled by human assets, social and technical skills, communication and leadership within the psychosocial work community, in addition to formal management. The above issues are herein dealt with in the context of an industrial case, a Finnish cluster of big companies employing the Health, Safety, Environment and Quality Assessment Procedure (HSEQ AP, www.HSEQ.fi). The HSEQ AP provides a comprehensive picture of a supplying company's capabilities, consisting of a description of management and systems, including many alphanumeric or categorised performance indicators and other information for managers and experts. One leadership approach emphasises the following drivers of excellence: (1) involving employees in management issues raises their motivation, competence and confidence; (2) this participation provides humans with many useful skills; (3) these skills are then available for employers' needs. Therefore, more human-centred communication and participation are needed in work organisations, related to the HSEQ issues and contexts as well.

1 Introduction

Production depends on the skills and competences of human assets, and an interplay of other tangible and intangible assets. This paper deals with the efficient, effective and responsible contributions of human assets. As far as all assets for production are concerned, the holistic quality of work life (QWL) (cf. Royuela et al. 2008) is increasingly seen as one of the most important factors in maintaining high quality companies and work communities. In the greater business context, entire value networks play a key role in companies' success, as do networks' QWL and productivity management. Excellent and consistent management of the network requires consideration of the entire picture: a company purchases services and executes production alongside many supplier–employer parties operating in a shared workplace. The management of information and communication, the unity of different employee communities and the participation of all contributing organisations and all their people comprise important prerequisites for successful business. Involving all employees in these issues is seen as an important starting point for excellence. This has been clearly

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 439–446, 2020.
https://doi.org/10.1007/978-3-030-48021-9_49

emphasised by the following chain of drivers of excellence: (1) involving all employees in management issues raises their motivation, competence and confidence; (2) this participation provides humans with all necessary social and technical skills; (3) these skills are then available for employers' needs (Wilson and Haines 2000). Total participation of all individuals within work places and processes for desired outcomes requires a great deal of shared information and wise communication, both intra- and inter-organisationally.

The role of human assets largely depends on human communication. It can be seen that information (I) and communication (C) act as essential factors and contributors of a workplace's organisation, environment and outcomes. (C) enables interaction that exists and operates between and within work communities, work places and other components, such as humans, tasks and technology, as for instance the current authors have described (Väyrynen and Kiema-Junes 2018). The following definition of (C) is used in this article: 'a process by which information is exchanged between individuals through a common system of symbols, signs, or behaviour' (Merriam-Webster 2018), and 'the technology of the transmission of information (as by print or telecommunication)' (Merriam-Webster 2018). (I) is typically connected to facts, data, knowledge, findings, intelligence and news (Merriam-Webster 2018). In intra-organisational contexts, and even more so in cross- and inter-organisational contexts, the channels and flow of (I) and (C) are essential for effective businesses and the individuals within them (e.g., Kiema et al. 2014). For instance, the practical importance of Health, Safety and Environment (I) depends significantly on consisting of (C) while reporting and collecting data, storing data, processing information, and distributing information to decision makers inside the organisation (cf. Kjellén 2000). Saari (1984) emphasised, in his wide analysis based on empirical findings, that disturbances in the information processes of a workplace and human communication are an important causal factor behind accidents at work.

Regarding effective (C), Glendon et al. (2006) conclude that the trainee (herein employee) is many times more likely to remember presented material if she/he is looking at pictures or watching a demonstration than if she/he is only hearing words or reading relevant material. Glendon et al. (2006) further advise that learning can be significantly boosted by active participation, simulation and performing.

The above key issues of production are dealt with in this paper in the context of an industrial case. This case considers a Finnish cluster of manufacturing and other big companies or company units encountering multiple issues of management related to the Health, Safety, Environment and Quality Assessment Procedure (HSEQ AP; Väyrynen et al. 2012; www.HSEQ.fi). New approaches that the cluster could attempt are sought. Proposals presented are based on the authors' own studies and other researchers' reports, and they aim to be useable in workplaces other than HSEQ AP's companies, as well.

2 Case: HSEQ AP-Related Organisations

Networking is a typical solution for companies of different sizes to combine and manage their contributions in a competitive way in a contemporary business environment. It is typical for employees from several supplying companies or contractors and independent workers to work simultaneously for the same production, such as in the process industry (customer of suppliers). This situation creates new requirements for managing HSEQ issues and achieving desired results within that framework. These requirements are often regulation-based, but many of them are voluntary, business-driven and promotional. Large-scale process industry companies in Finland have developed alongside Research & Development institutions (Väyrynen et al. 2016) and have created the collaborative cluster to apply HSEQ AP to measure and evaluate approximately 200 supplying companies. The objective of the HSEQ AP is to ensure that outside employees in shared workplaces experience the same conditions as internal employees, as well as have knowledge and skills that are good and consistent enough with HSEQ to operate on the principal companies' (more than 10) premises.

Integrated Management Systems (IMS) (cf. Wilkinson and Dale 2007) are used for a company's internal reasons, but also to assure its customers and partners that production, products and services satisfy holistic quality requirements. Responsible organisations also have to be concerned with QWL, such as the work environment and community issues (HS) and the impact on the environment (E). HSEQ management comprises the planning, organisation, control, monitoring and review of these measures; its features as an IMS are described by Kauppila et al. (2015). Diverse facts, figures, observations, hearing sessions and discussions are needed for an HSEQ assessment: the list of main topics and questions answered by representatives of the assessed company is presented by Koivupalo et al. (2015).

HSEQ assessments need to be effectively supported by an Internet-utilising information and communications technology (ICT) system (Fig. 1). Intranet

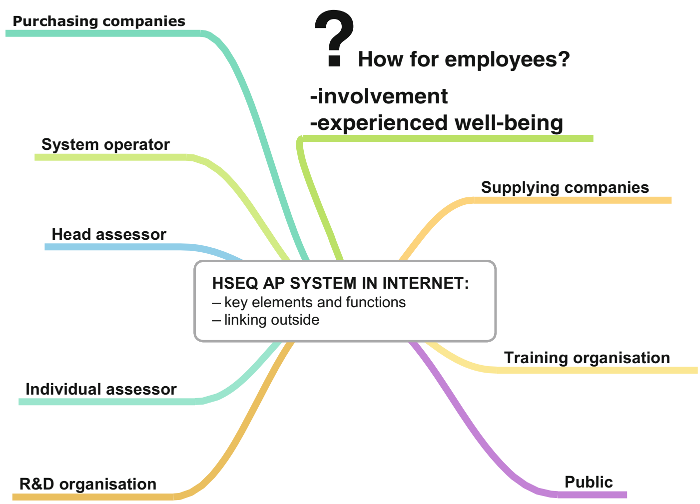


Fig. 1. HSEQ AP is connected to many key stakeholders. The Internet provides the most important channel for sharing information. Communication is carried out or at least supported by the Internet. It should be considered how to include more frequent and deep face-to-face communication about HSEQ AP indicators and issues facing all employees in both the purchasing and supplying companies.

communication about HSEQ is utilised mainly by purchasing companies belonging to the HSEQ AP cluster, the supplying companies, the head assessor and the assessors. The public section of ICT www pages provides information to the above, but is primarily for other representatives of the various stakeholders interested in it (Fig. 1). However, the specific utilisation possibilities of all employees of both the principal purchasing companies and the supplying companies should be considered, as they are key stakeholders, as well. They need to be better involved and considered in enriched and illustrative information regarding HSEQ material collected and in HSEQ AP communication.

3 Psychosocial Factors and Social Skills

Social skills are a key to success, both in specific work organisations and generally in working life. The level of social support can be identified by psychosocial factors in a workplace (Carayon 2009). Social relationships and communication are able to increase engagement and commitment (Schaufeli and Bakker 2004; King 2005). The importance of social skills is emphasised in dialogic leadership, which concentrates on the interactional process of the supervisor and subordinates instead of the individual characteristics of the leader (Hersey and Blanchard, 1979; Isaacs 1999; Yukl 2002). This new approach to management highlights the role of communication and interaction. Isaacs (1999) regards that ideal leadership and management comprises an interactive relationship between the leader and her/his employees. The quality of this relationship has been shown to have a great impact on employee job satisfaction and well-being, reduced staff turnover and innovativeness (e.g. Erdogan and Enders 2007; Loi et al. 2013). The previously mentioned authors posit that the relationship between supervisor and subordinate is an important aspect of employee well-being, as well. We examined supervisor–subordinate pairs in communication skills training for employees and managers, in a recent study in which social and communication skills were shown to increase due to training (Kiema et al. 2014). Communication can support the goals of a business and help to attain these goals (Greenberg and Baron 2003). To encourage fluent communication at the team level, emphasis should be given to cooperative skills, participatory decision-making, wide information and data management and interactive teamwork (Putman 1993).

4 New Ways for Communication of HSEQ AP

A possible leadership and management approach for enhanced communication and co-creation within work communities, like those of HSEQ AP, could include more discussions, meetings, participation, detailed narratives, storytelling and storyboards to find new solutions for improved QWL, including well-being at work and increased productivity (cf. Rajala and Väyrynen, 2013; Kiema et al. 2014).

Glendon et al. (2006) presented many examples of successful safety training, and their key emphases were listed earlier (cf. 1 Introduction). Nowadays, smartphone features enable pictorial and demonstrative abilities that are more effective learning aids than hearing or reading only (Glendon et al. 2006). Tools for technology-mediated communication have been increasingly utilised, including apps for mobile devices for all employees, such as ones available from the App Store or corresponding Android sources. Information and communication by and for all the key stakeholders of the HSEQ AP could comprise the following new ICT-enabled features:

(1) the remote presence of managers, even from the highest level, allowing discussion with work community members even if they are located at different sites (e.g., remote guidance and feedback); (2) the means for empowering employees and improving work conditions, both for productivity and a high QWL). There are many new approaches for operating and communicating better inter-organisationally and supporting all employers' people within the network.

5 Discussion and Conclusions

Glendon et al. (2006) suggest considering the following topics while leading industrial organisations: hierarchy, team structure and performance, degree of centralisation among teams and networks, attitudes and the quality of communication. This final point strongly relates to ICT at the company and personal levels, as well. Specifically, Glendon et al. (2006) encourage us to study top-down communication, bottom-up communication, ease of worker relations and reduced status distinctions to encourage communication and idea sharing. Quantitatively and qualitatively enriched information and frictionless communication about and with the environment and community of every employee are needed to enable a 'better' workplace (cf. Kiema et al. 2014). ICT must be tailored for multi-site jobs, as these are common within HSEQ AP-related companies, for remote work and to link with enterprise resource planning (ERP) systems and social media. New features would make it possible to include the presence of all managers, even from the highest level, possibly remotely and more frequently; remote guidance and feedback related to tasks; and "communication allocation", i.e., whether to discuss with individuals, teams, groups, one company, a group of companies or all employees of the purchasing company and its supplying companies.

We find that (C) at its best promotes 'transpersonal leadership' that is characterised as 'extending or going beyond the personal or individual, beyond the usual limits of ego and personality' (Coyne 2016). Transpersonal leaders are defined as follows (Coyne 2016):

'They operate beyond the ego while continuing personal development and learning. They are radical, ethical, and authentic while emotionally intelligent and caring. They are able to

- embed authentic, ethical and emotionally intelligent behaviours into the DNA of the organisation
- build strong, collaborative relationships, and
- create a performance enhancing culture that is ethical, caring and sustainable'.

A great deal of information is available regarding workplace experts, but the above style of communication is essential for every individual in a workplace. We recommend that HSEQ AP companies, purchasers and suppliers increase HSEQ- related human communication, and devote time and wisdom to these efforts. In the future, (I) related to the strategic and financial effects of work conditions, such as (I) related to accidents, should be provided to and communicated more at work places (cf., Reiman et al. 2018). Humans' social skills and the psychosocial factors of humans and organisations play a central role in the workplace. Open and fluent interactions increase functionality at work and enhance employee engagement and commitment, job satisfaction and well-being as described and reviewed by Kiema et al. (2014) and Väyrynen and Kiema-Junes (2018).

Additionally, Coyne (2016) discusses 'transpersonal leadership' as a mode of a leader 'extending or going beyond the personal or individual, beyond the usual limits of ego and personality'. In an analogical way, we could recommend 'trans-organisational leadership' for company networks. Employees should be enabled to use their best social and technical skills in their own work tasks and close work environment, though in many cases, they face strenuous conditions of changing and mobile tasks.

Other potential approaches suggested in the literature on this theme are as follows: a fair process (Kim and Mauborgne 2003), the so-called 'toolbox meeting' practice within the construction industry (Levitt and Samelson 1993), mobile ICT- like smartphones (Bye 2013; Thomas 2018) and the idea to use ICT as well for more psychosocial purposes, such as praising excellent employees (Horishita et al. 2013).

Generally, we think digitalisation can and should be utilised more. The close onsite interaction of employees with supervisors can be supported by technology, such as high-level visualised updated information, direct video discussion and spoken instruction provided by peers and/or managers and experts on special issues. This could include direct remote demonstrations of what needs to be done, and how. More potential can be found to enhance psychosocially enriched leadership through sophisticated digitalisation. More research is needed in this area. Starting empirical trials might be considered by purchasing and supplying companies within the HSEQ AP cluster. We recommend that the focus of the HSEQ AP's utilisation should be developed to include more communication to and with all about HSEQ issues. So, the future solutions should mean information for managers and experts, and communication with all employees within organisations and networks. In that way a company utilising HSEQ knowhow is forwarding towards excellence, i.e., well-being at work and productivity, corporate social responsibility and sustainability.

References

Bye, R.: Designing mobile user experiences: disruptive innovation in railway asset information. In: Dadashi, N., Scott, A., Wilson, J.R., Mills, A. (eds.) Rail Human Factors: Supporting Reliability, Safety and Cost Reduction, pp. 453–460. CRC, Taylor & Francis Group, London (2013)

Carayon, P.: The balance theory and the work system model… Twenty Years Later. Int. J. Hum.–Comput. Interact. **25**, 313–327 (2009)

Cecich, T., Hembarsky, M.: Relating principles to quality management. In: Christensen, W., Manuele, F. (eds.) Safety Through Design: Best Practices, pp. 67–72. National Safety Council, Itasca (1999)

Coyne, S.: Sustainable leadership; rewire your brain for sustainable success. In: Transpersonal Leadership Series: White Paper Three. Routledge, Taylor & Francis Group (2016). https://www.crcpress.com/rsc/downloads/WP-TL3-2016_Transpersonal_Leadership_WP3_r1.pdf. Accessed 25 Mar 2018

Erdogan, B., Enders, J.: Support from the top: supervisors' perceived organizational support as a moderator of leader-member exchange to satisfaction and performance relationships. J. Appl. Psychol. **92**(2), 321–330 (2007)

Glendon, A.I., Clarke, S., McKenna, E.F.: Human Safety and Risk Management. CRC Press, Boca Raton (2006)

Greenberg, J., Baron, R.A.: Behavior in Organizations, 8th edn. Prentice Hall, New Jersey (2003)

Hersey, P., Blanchard, K.: Situational leadership, perception, and the impact of power. Group Organ. Manag. **4**, 418–428 (1979)

Horishita, T., Yamaura, K., Kanayama, M.: Study of effective praise in train driver's workplace. In: Dadashi, N., Scott, A., Wilson, J.R., Mills, A. (eds.) Rail Human Factors: Supporting Reliability, Safety and Cost Reduction, pp. 614–620. CRC, Taylor & Francis Group, London (2013)

Isaacs, W.: Dialogic leadership. Syst. Think. **10**(1), 1–5 (1999)

Kauppila, O., Härkönen, J., Väyrynen, S.: Integrated HSEQ management systems: developments and trends. Int. J. Qual. Res. **9**(2), 231–242 (2015)

Kiema, H., Mäenpää, M., Leinonen, T., Soini, H.: Peer group counseling as a tool for promoting managers' communication skills in industrial and planning organisations. In: Proceedings of the Fourth International Conference on Advanced Collaborative Networks, Systems and Applications COLLA 2014, Seville, Spain, 22–26 June 2014, pp. 28–33 (2014)

Kim, W.C., Mauborgne, R.: Fair process: managing in the knowledge economy. Harv. Bus. Rev. **81**, 127–136 (2003)

King, W.: Communications and information processing as a critical success factor in the effective knowledge organisation. Int. J. Bus. Inf. Syst. **10**(5), 31–52 (2005)

Kjellén, U.: Prevention of Accidents through Experience Feedback. Taylor & Francis, New York (2000)

Koivupalo, M., Junno, H., Väyrynen, S.: Integrated management within a Finnish industrial network: steel mill case of HSEQ assessment procedure. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management - Solutions and Industrial Cases, Production & Process Engineering, pp. 41–67. Springer, Heidelberg (2015)

Levitt, R., Samelson, N.: Construction Safety Management, 2nd edn. Wiley, New York (1993)

Loi, R., Chan, K.W., Lam, L.W.: Leader-member exchange, organizational identification, and job satisfaction: a social identity perspective. J. Occup. Organ. Psychol. **87**, 42–46 (2013)

Putman, R.: Making Democracy Work. Civic Traditions in Modern Italy. Princeton University Press, Princeton (1993)

Rajala, H.-K., Väyrynen, S.: Participative approach to strategy communication: a case of small- and medium-sized metal enterprises with a review after seven years. Hum. Ergon. Manuf. Serv. Ind. **23**(4), 346–356 (2013). https://doi.org/10.1002/hfm.20322

Reiman, A., Räisänen, T., Väyrynen, S., Autio, T.: Strategic accident reduction in an energy company and its resulting financial benefits. Int. J. Occup. Saf. Ergon. (JOSE) (2018). https://doi.org/10.1080/10803548.2018.1462990

Royuela, V., Lopez-Tamayo, J., Surinach, J.: The institutional vs. the academic definition of the quality of work life. What is the focus of the European Commission? Soc. Indic. Res. **86**, 401–415 (2008). https://doi.org/10.1007/s11205-007-9175-6

Saari, J.: Accidents and disturbances in the flow of information. J. Occup. Accid. **6**, 91–105 (1984)
Schaufeli, W., Bakker, A.: Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study. J. Organ. Behav. **25**, 293–315 (2004)
The Merriam-Webster Dictionary (2018). http://www.merriam-webs-ter.com/dictionary/. Accessed 26 June 2018
Thomas, K.: Wanted: a WhatsApp alternative for clinicians. Br. Med. J. **360**, k622 (2018). https://doi.org/10.1136/bmj.k622
Väyrynen, S., Koivupalo, M., Latva-Ranta, J.: A 15-year development path of actions towards an integrated management system: description, evaluation and safety effects within the process industry network in Finland. Int. J. Strateg. Eng. Manag. **1**(1), 3–32 (2012)
Väyrynen, S., Jounila, H., Latva-Ranta, J., Pikkarainen, S., von Weissenberg, K.: HSEQ assessment procedure for supplying network: a tool for promoting sustainability and safety culture in SMEs, Chapter 5. In: Arezes, P., Rodrigues de Carvalho, P.V. (eds.) Ergonomics and Human Factors in Safety, Management, pp. 83–108. CRC, Taylor & Francis Group, Roca Baton (2016)
Väyrynen, S., Kiema-Junes, H.: Exploring blue- and white-collar employees' well-being at work system: differences in indicators of physical and psychosocial conditions of occupational groups. Int. J. Sociotechnol. Knowl. Dev. **10**(2), 14–34 (2018)
Väyrynen, S., Kisko, K., Filppa, H., Väänänen, M.: Review, framework and likert-scale survey for improving intra-organisational communication of Finnish case companies. In: Proceedings of the Fourth International Conference on Advanced Collaborative Networks, Systems and Applications, COLLA 2014, Seville, Spain, pp. 12–20 (2014)
Wilkinson, G., Dale, B.G.: Integrated management systems. In: Dale, B.G.T., van der Wiele, T., Iwaarden, V.V. (eds.) Managing Quality. 5th edn, pp. 310–350. Wiley-Blackwell, Chichester (2007)
Wilson, J., Haines, H.: Participatory Ergonomics. In: Karwowski, W. (ed.) International Encyclopedia of Ergonomics and Human Factors, vol. 2, pp. 1282–1286. Taylor & Francis, London (2000)
Yukl, G.: Leadership in organisations, 5th edn. Prentice-Hall, Upper Saddle River (2002)

Developing Students' Working Life Knowledge and Skills at All Educational Levels

Maria Lindholm[1(✉)], Seppo Väyrynen[1], Liisa Kiviniemi[2], Eija Rajakangas[2], and Minna-Liisa Myllylä[3]

[1] The University of Oulu, Oulu, Finland
{maria.lindholm, seppo.vayrynen}@oulu.fi
[2] The University of Applied Sciences of Oulu, Oulu, Finland
{liisa.kiviniemi, eija.rajakangas}@oamk.fi
[3] The Oulu Vocational College, Oulu, Finland
minna-liisa.myllyla@osao.fi

Abstract. In addition to strong subject-matter competence in their specific fields after graduating, students also need the understanding and ability to adapt to ever-changing working life, e.g., behaviour as a member of a work community, employer and employee roles, legislation, and good practices. The project *Prepared for working life!* aims to develop students' working life knowledge and skills at all educational levels, and by doing so, also improve the quality of working life in Finland once students begin to work with their knowledge and skills. The project utilises virtual learning tools and has created nationally usable study modules that prepare young adults for their future working life. In Oulu, all three educational levels (university, polytechnic, and vocational college) have collaborated to ensure that the projects succeed regionally. The Oulu project team has also cooperated with the HSEQ Training Park, a new and unique concept, in developing occupational health and safety and well-being at work education.

1 Introduction

In addition to strong subject-matter competence after graduating, students also need the understanding and ability to adapt to changes in their working lives. Working life skills and knowledge about well-being at work will help them to participate in and contribute to their work community and organisation and maintain their own well-being. The European Union, in its report "Mainstreaming occupational safety and health into university education", strongly recommends including and finding the means to teach occupational health and safety issues within universities (EU-OSHA 2010). Reasons for learning, while still in education, the safety, health, and well-being issues related to one's working lifetime can be enumerated as (1) humanitarianism, (2) the law, and (3) cost, like Brauer (1994) did.

The Finnish Institute of Occupational Health conducted a national study related to young people, schools, and the beginning of working life (Nykänen and Klemola 2015). The researchers described the need for occupational health and safety teaching in this way:

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 447–453, 2020.
https://doi.org/10.1007/978-3-030-48021-9_50

> Young workers have a higher risk of occupational accidents compared with older workers. Experiencing occupational accidents or developing work-related disease early in the career may influence young workers' work participation for the rest of their careers. Young people entering working life need to be equipped with effective behavioural strategies and skills that support their occupational safety and well-being.

The University of Oulu conducted a field study in Northern Finland's workplaces on the experienced needs of work sciences education. The study's message was clear: young experts moving into work from university need to be "prepared for working life" (Yrjänheikki and Majava 2015); Table 1 presents the opinions of different key groups in the workplace.

Table 1. Key groups' opinions of teaching work sciences (N = 96) (Yrjänheikki and Majava 2015)

	Management (N = 44)	Occupational safety managers (N = 19)	Occupational safety representatives (N = 33)
Not necessary	0%	0%	0%
Necessary	29.3%	10.0%	16.7%
Very necessary	31.7%	25.0%	43.3%
Extremely necessary	39.0%	65.0%	40.0%
Total	100%	100%	100%

The importance of teaching working life skills has been officially recognised in Finland; the European Social Fund-funded project *Prepared for working life!* (in Finnish, Valmiina työelämään!; VALTE), running from November 2015 to October 2018, aims to develop students' working life knowledge and skills at all educational levels by producing nationally usable study modules to prepare young adults for working life. Educational institutions, led by the University of Turku, have collaborated to ensure that different perspectives are acknowledged. The institutions involved are the Universities of Turku, Oulu, Jyväskylä, Vaasa, and Lapland; the Universities of Applied Sciences of Oulu, Jyväskylä, Vaasa, Saimaa, and Turku; and the Oulu Vocational College.

The VALTE project's goals are to improve training at educational institutions in working life skills and well-being at work, and by doing so, also improve the quality of working life in Finland as the students move in to work. This study aims to explain what the VALTE project has done nationally in Finland, and especially in Oulu (Northern Finland) to respond to the need to develop students' working life knowledge and skills.

2 VALTE Project Nationally

Every educational institution participating in the VALTE project produces one or more study modules. The study modules aim to improve students' knowledge and competence in working life skills to ensure that they know how to act and react at work and when facing working life issues, and are able to work healthily until retirement. The study modules give practical methods to develop, control, and manage occupational well-being at the personal, team, and organisational levels. The themes of the modules include, for example, health and well-being, occupational safety, transitions from education to work, equality and diversity, and management and organisation.

The study modules created during the VALTE project will be accessed through ViLLE, a collaborative education platform developed by a team of researchers at the Department for Future Technologies of the University of Turku. From ViLLE, educational institutes can choose the modules they need, which can be taught face-to-face or with minimal contact between teacher and students as ViLLE allows the modules to utilise virtual learning, which is independent of time and place. The study modules exploit gamification, with exercises and animated games. Most of the materials will also be on the VALTE project's website to allow workplaces to use them, as well as students. The materials on the website include participatory exercises to be used for team and development activities.

The project has organised development days amongst the 11 institutions, held in Jyväskylä, Turku, Tampere, Helsinki, and Oulu, and five public seminars. Two VALTE seminars have been held in Turku, two in Oulu, and one in Jyväskylä, on themes including future working life, working life knowledge and skills, occupational well-being and safety, and education in working life skills. The participants have represented companies and other workplaces, students, teaching staff, developers, authorities, researchers, and research institutions from various parts of society. In addition to organising the seminars, the project's researchers have participated in conferences and other educational events.

The University of Turku has also conducted surveys on working life skills for students, human resources personnel, and graduates. In the last months of the project, the study modules will be finalised and shared with institutions for use outside the project, a closing ceremony held, and the final report written.

3 VALTE Project in Oulu

3.1 University of Oulu

The University of Oulu is an international science university founded in 1958. Its eight faculties are from the fields of biochemistry and molecular medicine, education, humanities, information technology and electrical engineering, medicine, science, technology, and the Oulu Business School. The university has approximately 13,500 students and 2,800 employees (University of Oulu 2018). In the mid-1970s, the university began to widen its work science teaching, and research increased in the late 1980s when a professorship was created in the Faculty of Technology. The Well-being and Productivity research team, from the Industrial Engineering and Management unit in the Faculty of Technology, participates in the VALTE project.

The five-credit study module created by the University of Oulu for VALTE is *Safety of the working environment – the key to well-being and productivity*. Themes include safety management, law and directives, responsibilities in different roles, working conditions and environment, ergonomics, risk assessment, safety measurement, and the benefits of safety. When the study module was piloted, students visited the Health, Safety, Environment, and Quality (HSEQ) Training Park – a safety innovation where visitors can be trained on a practical level to perform safely at work and manage and utilise effective and efficient good practices tailored to individual abilities and limitations.

The HSEQ Training Park is located in Oulu, covers an area of 1.2 ha, and provides more than 20 training points that are full-scale working demonstrations of building and infrastructure construction, manufacturing, energy, rail transport, and industrial services. The training points consist of full-size representations, work-stations (mock-ups), and work-task scenarios with tools and equipment (Reiman *et al.* 2015, 2017). Regarding the HSEQ Training Park, an escape room game is also under development. Problems presented at each training point are related to that point's themes and require numerical and verbal reasoning.

In addition to these contributions, the University of Oulu piloted the study module *Professional interaction and development of the professional identity*, created by the Oulu University of Applied Sciences. The module was tested with students from the university, and feedback gathered from the students and the teacher.

3.2 Oulu University of Applied Sciences

Oulu University of Applied Sciences (OUAS) is one of the biggest universities of applied sciences in Finland. OUAS offers studies in the fields of business, engineering, health and social care, information and communication technology, media and performing arts, and natural resources. OUAS has 9,000 students and over 600 employees and offers bachelor's and master's degree programmes, professional specialisation studies, pedagogic studies at the School of Vocational Teacher Education, and Open University studies, as well as supplementary training.

OUAS has planned and piloted the modules on successful career and professional interaction and identity for the VALTE project. In the study module *For a successful career* (two or three credits), students are taught the demands of working life, learn to design their careers, develop self-knowledge and readiness as a future applicant, identify and market their own know-how, prepare to answer job advertisements, and practice for job interviews. The main approaches of the study module are group work and peer evaluation.

In the study module *Professional interaction and development of the professional identity* (three to five credits), students learn to identify the basic skills of professional interaction, examine their own interactions, become acquainted with their strengths, and strengthen their professional identities. The pedagogic background of the study module is positive learning and pedagogics, and the essential components of the implementation are interaction and reflection. The study module was planned and piloted in cooperation with the University of Oulu.

OUAS is also planning a model to support professionals at the beginning of their careers, in cooperation with managers and young professionals. The model aims to teach skills to manage and promote well-being at work and will be piloted in spring 2018 by managers in health and social care.

3.3 Oulu Vocational College

The Oulu Vocational College (OSAO) is one of the biggest vocational schools in Finland. OSAO is a multidisciplinary vocational college that trains professionals for the labour market and provides up-to-date education in line with the needs of working life. OSAO is one of the leading vocational colleges in Finland, with 14,000 students and 935 staff. Cooperation with the business community is routine and includes on-the-job learning, skills demonstrations, company personnel training, and joint development projects. All OSAO education sectors and units offer services to the public.

For the VALTE project, a group of teachers developed online studies that prepare students for working life and which include knowledge and exercises on employment contracts, working hours, contractual skills, and workplace rules. Project workers also visit companies and workplaces to identify working life skills; this approach is one form of the cooperation between the workplaces and education development. OSAO also collaborates with the University of Oulu and OUAS.

3.4 Collaboration in Oulu

In Oulu, all three educational levels – university, polytechnic, and vocational college – have collaborated in the VALTE project, with monthly meetings of the key project teams, two public seminars for different stakeholders of the working life and students, and collaboration in creating the study modules themselves. The module S*afety of the working environment – the key to well-being and productivity* was tested with a small group of students from the University of Oulu and OUAS. The module *Professional interaction and development of the professional identity* was originally developed at OUAS, and further developed after it was piloted at the University of Oulu.

The Oulu project team also cooperates with the HSEQ Training Park. Students have visited the HSEQ Training Park in multidisciplinary groups consisting of students from all three educational levels. Studying in multidisciplinary groups promotes understanding of others' professional knowledge and skills and strengthens cooperative skills. Most companies employ young adults from all three education levels and the collaboration of all students is therefore a new, powerful tool in teaching working life skills.

The first seminar, *Prepared for working life?* was held at the University of Oulu in November 2016. The 66 participants represented companies and other workplaces, labour market organisations, students, teachers, researchers, developers, and other interested groups, such as authorities and research institutions, from many parts of society. At this seminar, companies and industrial employers' associations noted the strong commitment of the international companies, and practical models and achievements in safety management were emphasised. It was also stressed that the work community skills already learned at educational institutions are of significant importance when seeking employment.

The themes of the afternoon workshops were *The attraction of work*, *Visions of the collaboration of three educational institution levels in working life studies*, *Occupational safety today and tomorrow*, *What working life skills are expected from recent graduates? – Industry and services (public and private)*, and *Professional interaction* *at work in the 2020s – frictionless communication*. The discussion revealed several important viewpoints, such as the importance of students forming an affirmative approach to the workplace to bring out the strengths of their educational institutions and themselves, and having a good attitude, motivation, and the ability to cope with changes in working life. A further recommendation was to link VALTE matters to worldwide sustainable development and corporate responsibility.

The second seminar, *Students to a successful career*, was held on the OUAS and OSAO campus in October 2017. The 80 participants were students, experts in working life, and teachers. The seminar programme consisted of an introduction to the steps to a successful career, panel discussions, and workshops. A panel discussion featured participants from different stages of career and working life, including students, an entrepreneur, a manager and a chief learning officer. Brave choices and actions, networking, and the meaning of attitude were emphasised in the panel discussion.

The themes *The ABCs of working life skills*, *The flow of work*, *Working life skills as a part of studying* and *How to create a workplace of our dreams* were considered in the afternoon workshop. Three important working life skills were raised in *The ABCs of working life skills*: attitude, such as positivity, reliability, honesty, and kindness; cooperative skills, such as communication skills and following common rules; and individual ability to learn and develop. Other viewpoints included the importance of studying working life skills to cope in the workplace, developing versatile social skills, and knowledge about responsibilities.

In the final year of the VALTE project, the Oulu project team aims to pilot and finalise its study modules, strengthen the cooperation between the three educational levels, organise a seminar for teaching staff, and further develop the use of the HSEQ Training Park in studies.

4 Discussion

The VALTE project aims to develop students' working life knowledge and skills at all educational levels and in all fields. This development is important to ensure long careers and a good quality of working life through understanding of the employer-employee roles and rules. With good working life skills, students can become professionals who maintain their well-being, promote the well-being of their organisations, and work safely. They can also provide better leadership for their subordinates and company as well as fellowship with their peers.

The teaching of working life skills should occur over academic years. Students' knowledge should mature, and they need to learn to understand different operational models and recognise the situations in which to apply them. This also helps in engineering asset management, to which safety and managerial skills are related.

Developing education in working life skills is also in the best interests of companies, industries, and other workplaces as good education creates a skilled workforce. Agreement with this point was seen in the seminars organised by the Oulu project team, since there were participants from several industries including technology, construction, trade, municipalities, unions, and white-collar workplaces. Interest in the highlighted themes, such as attitude and cooperation skills, showed that the VALTE project is topical.

The cooperation between 11 educational institutions in the VALTE project ensures that the main points and different perspectives relating to working life skills are acknowledged in the study modules. The results of the VALTE project are not yet measurable; measurement of the effects on students' development during their studies and their actions in working life requires longitudinal monitoring. However, the students who participated in pilot courses have given positive feedback. Participants in the *Safety of the working environment – the key to well-being and productivity* course of the University of Oulu were clearly satisfied and expressed their understanding of the holistic view of occupational safety and health and their knowledge of how to recognise and prevent risks.

The VALTE Oulu project team's activities, with study modules, seminars, and strong cooperation between educational institutions at three levels, are developing students' working life skills both regionally in Oulu and through the use of study modules nationally in Finland. Due to the professorship established at the Faculty of Technology of the University of Oulu, teaching in this field is frequent and expanding. OUAS and OSAO also have long traditions of collaboration with the professional world, and the VALTE project has sharpened working life skills education even further.

References

Brauer, R.L.: Safety and Health for Engineers. Wiley, New York (1994)

EU-OSHA: Mainstreaming occupational safety and health into university education. European Agency for Safety and Health at Work. Publications Office of the European Union, Luxembourg (2010)

Nykänen, M., Klemola, S.: Avaimet työelämään – kehittämishanke, Helsinki (2015)

Reiman, A., Airaksinen, O., Väyrynen, S., Aaltonen, M.: HSEQ training park in Northern Finland - a novel innovation and forum for cooperation in the construction industry. In: Väyrynen, S., Häkkinen, K., Niskanen, T. (eds.) Integrated Occupational Safety and Health Management - Solutions and Industrial Cases, pp. 145–153. Springer, Switzerland (2015)

Reiman, A., Pedersen, L.M., Väyrynen, S., Sormunen, E., Airaksinen, O., Haapasalo, H., Räsänen, T.: Safety training parks – cooperative contribution to safety and health trainings. Int. J. Constr. Educ. Res. **15**(1), 19–41 (2017)

University of Oulu: University of Oulu (2018). http://www.oulu.fi/university/node/34711. Accessed 2 Feb 2018

Yrjänheikki, E., Majava, J.: Työhyvinvoinnin yliopistotasoinen opetus (Well-being at work education at the university). University of Oulu, Oulu (2015). (in Finnish)

Part VI: Human Capital and Organizational Development

Contemporary Issues of Intellectual Capital: Bibliographic Analysis

Olga N. Korableva[1,2(✉)], Nikolay Gorelov[3], and Olga Kalimullina[4]

[1] Institute of Regional Economic Studies of Russian Academy of Science, St. Petersburg State University, S-Petersburg, Russian Federation
on.korableva@gmail.com
[2] ITMO University, Saint Petersburg, Russian Federation
[3] Saint Petersburg State Economic University, Saint Petersburg, Russian Federation
gorelov33@inbox.ru
[4] The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, S-Petersburg, Russian Federation
chemireva@mail.ru

Abstract. Intellectual capital management is a key factor in the transition to the digital economy. The rapid development of the IoT, cloud computing, Big data and other technologies leads to the creation of a complex interdisciplinary environment - industry 4.0. However, at the moment, the role of intellectual capital is not defined in this context. The purpose of this study is to analyze the existing work in the field of intellectual capital and the digital economy, to identify and systematize the main research problems affected by the authors, and to identify the most promising research problems.

1 Introduction

Currently, the main digital economy resources are the intangible assets, consisting of information, knowledge, innovation and creativity. The performance of the expenses on research and development, their results in the form of technologically advanced products and their application in the production of traditional goods determines the productivity of the entire economy. The competitive advantage of an enterprise is no longer connected with the physical capital, such as land, equipment or facilities, but is based on the value created by the intangible assets and knowledge of the enterprise, its intellectual capital. Thus, the intellectual capital is one of the most important concepts of the creation of the competitive advantages in the digital economy. The increasing role of intellectual capital can be proved by the patent statistics. So, in 1995, 1 million patent applications were filed worldwide, and since then their number has increased. In 2011, applications exceeded 2 million. In 2016, more than 3.1 million patent applications were filed worldwide (WIPO 2017). However, the intellectual property rights market is not the only indicator of the development of the innovation and research leadership of the country, but the level of its development reflects the degree of development and transparency of managing the intellectual activity results. The human capital, as an integral part of the intellectual capital, can also serve as an indicator of the

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 457–464, 2020.
https://doi.org/10.1007/978-3-030-48021-9_51

change in the IC. Currently, according to the Global Human Capital Report 2017, only 62% of global human capital have been developed by now.

Undoubtedly, in conditions of dynamic innovative development, the intellectual capital is an integral part of any business, and its importance for many companies – market leaders exceeds the market capitalization (Korableva et al. 2017). Management of such an economy requires the development of new strategic approaches and practice-oriented techniques, both at the macro and micro levels (Korableva and Kalimullina 2016; Kurbanova et al. 2012). In the framework of this field of research, the objective of great importance is the formation of new conceptual foundations and methodology for management of the intellectual capital, first of all for the companies focused on innovations. The relevance of the research on the problem of evaluation and management of the intellectual capital is beyond doubt. However, what areas of research in the field of IC have been presented at the moment? What issues have not been adequately addressed or covered at all? To answer these questions, a systematic bibliographic analysis of the existing works in the field of intellectual capital in the context of the digital economy was carried out in order to determine the main themes and areas of research in this field, as well as their subsequent systematization and identification of promising areas of research not yet covered by modern scientific works.

2 Methodology

The study consists of two stages: 1. Literature search aimed at identification of the initial set of questions; 2. Structured classification of research questions and problems, as well as the identification of the fields for further research. At the stage of literature search, the analytical method is used, as well as the semantic method of text analysis, which make it possible to determine the semantic core of the article being analyzed. To carry out a structured classification of the research questions, a whole set of methods of economic research is used to provide a systematic approach to the study of the problem, as well as the metaplan approach. The methodological basis of the research includes the methodological principles, the theoretical provisions and conclusions contained in the fundamental scientific works on the problems of research of the intellectual capital.

2.1 Literature Search

In the course of the study, a structured literature review is conducted. The expanded version of the Scopus database, the world's largest universal abstract database, is used for this purpose. The search for such word combinations as "intellectual capital", "intellectual capital and digital economy", "IC and digital economy" is performed in the headings of articles. The search is limited to the documents published during the last 5 full years, and in 2018 (as of the date of search) (2013–2018). The sources and a basic set of keywords are expected to provide a good sample of literature. The completeness of the results can not be guaranteed at this stage, taking into account only the headings and a limited set of keywords. However, this is considered sufficient for this stage of the study. Also, the study of the documents found as a result of cross-referencing and provided by the Scopus database is performed. In the course of the study, a headline is studied first and a decision is made on the relevance of the article, then, if the article is the subject to the header check, the keywords and the abstract are studied. If the article passes the second check, the full text of the article is loaded into the semantic text analysis program released by Advego (https://advego.com), which allows to identify the semantic core of the text and, if the analysis of the semantic core confirms the compliance with the main keywords, the complete article is studied, otherwise the article is eliminated.

2.2 Classification

The resulting set of articles is carefully analyzed with the extraction of potential research problems. Each potential problem is singled out and marked. Since the problems are singled out from the individual articles, the resulting set could contain the coincidence of topics formulated at different levels of abstraction and, as a result, be too large and difficult to analyze. To avoid such a situation, a structured classification of the research problems is carried out. For the classification, the Metaplan approach, the sorting method based on the adjustment of the group discussion is chosen (Howard 1994).

3 Results

3.1 Literature Search

The search for the articles by the keywords resulted in the identification of 466 works. Of these, 13 documents were found by the keywords “intellectual capital AND digital economy”. 31 documents were found by the keywords “IC and digital economy”. More than 6000 thousand documents (2013–2017) were found by the keywords “intellectual capital”, so the search for these keywords was limited to 2017–2018 years. Thus, 422 documents were found by “intellectual capital”. After further study of the headings of the articles and their abstracts, 58 articles were selected for the semantic core analysis. The example of the semantic core analysis of the article, written by Asiaei and Jusoh (2017) is presented in Table 1.

Table 1. The example of the semantic core analysis (the article, written by Asiaei and Jusoh 2017)

Phrase/word	Amount	Frequency, %
Performance	174	1.72
Measurement	126	1.24
Intellectual capital	119	1.17/2.35
Organization	82	0.81
Diversity	77	0.76
Organizational performance	57	0.56/1.12

For the semantic core analysis, the text of the article was used without abstract, keywords and a list of references. The data on words with a frequency of mention greater than 0.5% is presented in Table 1. As a result of studying the semantic core of the text, 31 articles were selected for the full text study. The reading of these articles resulted in the final selection of 11 documents. The search for cross-references resulted in 8 possible useful articles, of which 4 were useful. The cross-search of these articles did not lead to further conclusions. The further analysis of the 15 articles obtained (Agostini and Nosella 2017; Asiaei and Jusoh 2017; Demuner Flores et al. 2017; Doucek et al. 2017; Gavrilova et al. 2017; Indiran et al. 2017; Kotarba 2017; Labra and Sánchez 2017; Ramírez et al. 2017; Secundo et al. 2017; Călin et al. 2017; Martins and Lopes 2015; Erickson and Rothberg 2015; Besley and Peters 2013; Widiatmoko and Indarti 2017) showed that six of them brightly reflect the problems of research. Of these, a total of 14 potential problems were identified. The remaining nine articles did not have such direct evidence, but provided the relevant materials, of which another 3 potential problems were identified. This allowed the authors to compile a list of 17 potential research problems.

Thus, in the sample set, several groups of research problems were identified, as well as several potential areas of research. The first group of potential problems is related to the methodology of research and evaluation of the intellectual capital. The totality of these problems is the least studied in comparison with others. The intellectual capital is the most important factor determining the economic growth (Korableva and Litun 2014). However, there are very few studies on the assessment of the intellectual capital at the country level. Therefore, a special interest is the study (Labra and Sánchez 2017) on the definition, selection and classification of models for the assessment and management of intangible assets at the country level. The study (Doucek et al. 2017) of new trends in the transformation of the world economy on the way from the economy of goods to the knowledge economy also is of great interest.

The second group of problems is related to the generation and formation of human capital (the basis for the formation of the IC), this includes the problems of transferring the knowledge from universities to external stakeholders, the access to information and the distribution of intellectual capital in the digital economy, as well as the assessment of the intellectual capital of the educational institution (university). This group of problems is widely reflected in modern studies (Di Berardino and Corsi 2018; Secundo et al. 2017). The assessment of the intellectual capital directly at a university also seems very relevant, since the knowledge is the main product, as well as the contribution of these institutions. The intellectual capital and the university performance in developing countries are being discussed by the scientists from different countries (Cricelli et al. 2018; Ramírez et al. 2017). Also within the framework of this group of problems, the problem of loss of the human capital in the system of progress evaluation (Martin-Sardesai and Guthrie 2018).

The third group of problems is related to the management, analysis and control of the IC and also includes a wide range of diverse studies (Agostini and Nosella 2017; Wahyuningtyas et al. 2018). In modern literature, there is a gap associated with the measurement indicators of each component of the intellectual capital (Indiran et al. 2017; Demuner Flores et al. 2017). The impact of the knowledge-based assets on the measurable indicators has been extensively studied in the literature on intellectual capital, but little is known about the role of the organizational control system in the facilitation of the intellectual capital management as the most strategic asset for the enterprises. The study (Asiaei and Jusoh 2017) considers the role played by the performance measurement system (in terms of the diversity of dimensions) in the relationship between the intellectual capital and the organizational indicators. The next group of questions is related to the impact of the corporate governance and company characteristics (profitability, leverage, company size, company age) when disclosing the intellectual capital (Widiatmoko and Indarti 2017; Kotarba 2017; Wudhikarn 2017).

3.2 Classification

Almost the same problems of research can be considered by various authors, so they were grouped, which resulted in the identification of 17 unique potential research questions that were conditionally divided into 3 groups: “Methodology for research and evaluation of the IC”, “Education and human capital”, “Management, analysis, control of the IC. “ The resulting structured list is presented in Table 2.

4 Discussion

Despite the wide coverage in the scientific literature of the problem of management of the intellectual capital, its conceptual framework has not been sufficiently developed and there is, in fact, no efficient methodology for assessment of the intellectual capital and making the managerial decisions. Also, the relevant research topics in this context are the insufficient understanding of the process of generation of the IC, which often hinders the enforcement of property rights and the generation of the intellectual capital of the company, as well as the variety of methods for measuring the intellectual capital of the company, each of which has its advantages and disadvantages. The creation of a unified methodology for the formation and evaluation of the intellectual capital is of practical interest on the part of modern companies. Therefore, one of the topical fields of research is the development of a methodology, based on an integrated approach to the analysis of the information field, associated with the intellectual capital and its transformation into a knowledge model. In general, the development of a new approach and methodology will, on the one hand, theoretically generalize the emerging problems of the capitalization of intellectual activity and formulate them in scientific terms, and on the other hand, it will largely contribute to improving the labor productivity and business competitiveness.

Table 2. Groups of actual research questions

Education and human capital
The problem of access to the information and distribution of the intellectual capital in the digital economy
The problem of assessment of the intellectual capital of an educational institution (university)
The problem of knowledge transfer from the universities to external stakeholders and society as a whole
Identification of the key non-material elements and indicators that should be included in the model of intellectual capital of the universities
Investigation of the impact of human, organizational and relational capital in the digital economy
Management, analysis, control of the IC
Investigation of the relationship between the level of intellectual capital and profitability of the companies
The problem of the influence of the corporate governance and the characteristics of the company when disclosing the intellectual capital
Determination of the relationship between the intellectual capital, strategy, competitive intelligence
Evaluation of the intellectual capital of the company
Evaluation of the performance of the intellectual capital of small and medium-sized enterprises (SMEs)
Compliance control and corporate culture as an effective management of the intellectual capital
Methodology for research and evaluation of the IC
Methodology for assessment of the intellectual capital at the micro level
Methodology for assessment of the intellectual capital at the country level
The role of various components of the intellectual capital
Comparison of the three aspects of the intellectual capital
Methodology for comprehensive analysis of the intellectual capital by industry
New trends in the transformation of the world economy on the way from the economy of goods to the knowledge economy

5 Conclusion

The proprietary knowledge of the company, as well as the application method, are currently the decisive factor for achieving the optimal potential for the development of the digital economy and creation of a stable social environment. At first glance, the digital economy and the intellectual capital are two related concepts, but what studies describe and research this relationship? The literature analysis revealed 17 main problems of the study of the intellectual capital, as well as the fact of absence of the unified methodology for the formation and assessment of the intellectual capital in the digital economy. Moreover, the problems of improvement of the role played by the intellectual capital in the digital economy in modern literature have not yet been considered. Speaking about the fields of further research in this area, it should be noted that at the moment there are many theories of the IC formation, various scientists distinguish different components of the IC, differently describe the relationship between them. Also, there are a lot of IC evaluation methodologies, which should be analyzed and systematized, which will serve as the basis for the formation of a unified methodology for the assessment of the IC in the digital economy. Moreover, the study of the structural components of the intellectual capital will contribute to the further development of the stakeholder theory and the search for the new ways of the interaction of all the stakeholders in the formation of a knowledge-based economy.

Acknowledgment. The research is supported by the RFBR grant 16-29-12965/17.

References

Agostini, L., Nosella, A.: Enhancing radical innovation performance through intellectual capital components. J. Intellect. Cap. **18**(4), 789–806 (2017)

Asiaei, K., Jusoh, R.: Using a robust performance measurement system to illuminate intellectual capital. Int. J. Account. Inf. Syst. **26**, 1–19 (2017)

Besley, T.A.C., Peters, M.A.: Re-imagining the Creative University for the 21st Century, pp. 1–188 (2013)

Călin, A.M., Ciobanu, G., Iovițu, M.: The access of the population to information and communications technology - an opportunity for development of the digital economy. Qual. - Access Success. **18**, 107–113 (2017)

Chen, Y., Cook, W.D., Li, N., Zhu, J.: Additive efficiency decomposition in two-stage DEA. Eur. J. Oper. Res. **196**(3), 1170–1176 (2009)

Cricelli, L., Greco, M., Grimaldi, M., Llanes Dueñas, L.P.: Intellectual capital and university performance in emerging countries: evidence from Colombian public universities. J. Intellect. Cap. **19**(1), 71–95 (2018)

Demuner Flores, M.R., Saavedra García, M.L., Camarena Adame, M.E.: Measurement of intellectual capital in the banking sector: application of skandia and vaic models. Innovar **27** (66), 75–89 (2017)

Doucek, P., Fischer, J., Novotný, O.: Digital economy. In: IDIMT 2017: Digitalization in Management, Society and Economy - 25th Interdisciplinary Information Management Talks, pp. 33–40 (2017)

Di Berardino, D., Corsi, C.: A quality evaluation approach to disclosing third mission activities and intellectual capital in Italian universities. J. Intellect. Cap. **19**(1), 178–201 (2018)

Erickson, G.S., Rothberg, H.N.: A longitudinal look at strategy, intellectual capital and profit pools. J. Intell. Stud. Bus. **5**(2), 5–13 (2015)

Gavrilova, T., Alsufyev, A., Kokoulina, L.: Knowledge management trends in the digital economy age. Commun. Comput. Inf. Sci. **745**, 464–473 (2017)

Han, D., Han, I.: Prioritization and selection of intellectual capital measurement indicators using analytic hierarchy process for the mobile telecommunications industry. Expert Syst. Appl. **26**, 519–527 (2004)

Howard, M.S.: Quality of Group Decision Support Systems: a comparison between GDSS and traditional group approaches for decision tasks. Ph.D. thesis Eindhoven University of Technology (1994)

Indiran, L., Khalifah, Z., Ismail, K.: A model for intellectual capital of business incubators. Adv. Sci. Lett. **23**(9), 8450–8457 (2017)

Korableva, O., Kalimullina, O.: Strategic approach to the optimization of organization based on the BSC SWOT matrix. In: Proceedings of the International Conference on Knowledge Engineering and Applications, ICKEA 2016, Singapore, 28–30 September 2016, pp. 212–215 (2016)

Korableva, O.N., Razumova, I.A., Kalimullina, O.V.: Research of innovation cycles and the peculiarities associated with the innovations life cycle stages. In: Proceedings of 29th IBIMA Conference, Vienna, Austria, 3–4 May 2017, pp. 1853–1862 (2017)

Korableva, O., Litun, V.: The potential of transitive economies' growth based on innovative strategy. WSEAS Trans. Bus. Econ. **11**(68), 725–736 (2014)

Kotarba, M.: Measuring digitalization-key metrics. Found. Manag. **9**(1), 123–138 (2017)

Kurbanova, E., Korableva, O., Kalimullina, O.: Enhancing the effectiveness of asset management through development of license management system on the basis of SCCM 2012 program by microsoft company. Paper Presented at the ICEIS 2018 - Proceedings of the 20th International Conference on Enterprise Information Systems, vol. 2, pp. 171–178 (2018)

Labra, R., Sánchez, M.P.: Intellectual capital of nations: a comparative analysis of assessment models. Knowl. Manag. Res. Pract. **15**(2), 169–183 (2017)

Martins, M.M., Lopes, I.T.: Intellectual capital and profitability: a firm value approach in the European companies. In: Proceedings of the European Conference on Knowledge Management, ECKM, pp. 496–503 (2015)

Martin-Sardesai, A., Guthrie, J.: Human capital loss in an academic performance measurement system. J. Intellect. Cap. **19**(1), 53–70 (2018)

Ramírez, Y., Manzaneque, M., Priego, A.M.: Formulating and elaborating a model for the measurement of intellectual capital in Spanish public universities. Int. Rev. Adm. Sci. **83**(1), 149–176 (2017)

Secundo, G., Elena Perezs, S., Martinaitis, Ž., Leitner, K.H.: An intellectual capital framework to measure universities' third mission activities. Technol. Forecast. Soc. Change **123**, 229–239 (2017)

The Global Human Capital Report (2017). https://www.weforum.org/reports/the-global-human-capital-report-2017

Widiatmoko, J., Indarti, M.G.K.: The influence of corporate governance and company characteristics on intellectual capital disclosures. Adv. Sci. Lett. **23**(8), 7059–7061 (2017)

Wahyuningtyas, R., Astuti, Y., Anggadwita, G.: Identification of intellectual capital (IC) within micro-, small-and medium-sized enterprises (MSMEs): a case study of Cibuntu Tofu Industrial Center in Bandung, Indonesia. Int. J. Learn. Intellect. Cap. **15**(1), 51–64 (2018)

WIPO: World Intellectual Property Indicators 2017. World Intellectual Property Organization, Geneva (2017)

Wudhikarn, R.: Determining key performance indicators of intellectual capital in logistics business using Delphi method. In: 2nd Joint International Conference on Digital Arts, Media and Technology 2017: Digital Economy for Sustainable Growth, ICDAMT 2017, pp. 164–169 (2017). 7904955

Maintenance Education in Engineering Programs on Bachelor and Master Level: Evidence from Finland and Sweden

Mirka Kans[1(✉)] and Lasse Metso[2]

[1] Department of Mechanical Engineering, Linnaeus University, Växjö, Sweden
mirka.kans@lnu.se

[2] Industrial Engineering and Management, Lappeenranta University of Technology, Lappeenranta, Finland
lasse.metso@lut.fi

Abstract. This paper discusses the need for maintenance related training in higher education and investigates the maintenance related education offered by engineering programs in Finland and Sweden. Main study objects are Finnish and Swedish Mechanical and Industrial engineering programs on both bachelor and master level. The study covers, for the selected programs, full programs in maintenance, single courses and parts of courses in which maintenance plays a role. In Finland there are in total 42 universities and applied science schools offering 115 programs within Mechanical or Industrial engineering. Of those, 17 programs contain some sort of maintenance related training. The corresponding figures for Sweden are 23 universities and applied science schools offering 87 programs within Mechanical or Industrial engineering, and 10 of these programs contains maintenance related education. For reviewing the educational contents, data was collected from course syllabuses; for each course the content and expected learning outcomes were analysed and categorised. The maintenance related education in the studied programs is in general low; less than 15% offer maintenance courses. The content in the maintenance related courses differs greatly: concept of maintenance, information systems in maintenance, reliability, life-cycle management, condition monitoring and management of maintenance are covered. For increasing the maintenance topics in higher education, the development of appropriate study material and joint online courses are suggested.

1 Introduction

Engineers of today require a holistic understanding of products and processes. Consequently, engineering education should provide the possibility to acquire these kind of competences. The worldwide initiative CDIO (Conceive-Design-Implement-Operate) was developed in collaboration with academia, industry and students for supporting the realisation of a modern engineering education, Crawley et al. (2008). According to CDIO, the engineering students have to understand both engineering fundamentals and the full product life cycle, from the conceiving of an idea to the operations of the finished product. A mechanical engineering student for instance needs to understand

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 465–474, 2020.
https://doi.org/10.1007/978-3-030-48021-9_52

not only how to design and construct products, but also how the products are operated and maintained. Similarly, an industrial engineering student needs understanding of how to optimise the whole business, from production and logistics to maintenance and asset management. This paper will focus on education regarding the operations and maintenance phase.

The topic of maintenance related education in engineering programs and maintenance training is not well researched: only four publications were found during a structured review. In Heilman and Heilman (2011) findings from an interview survey of maintenance training in the Finnish forest industry stress the importance of competence development for current as well as new employees. However, this article does not address the higher education and how this should be designed for preparing prospective new employees. Lai (2010) surveyed building-related education in Hong Kong, and no program was found that focused on operations and maintenance of buildings. Lai reports on strong interest in such education amongst the respondents. Knezevic (1997) describes a postgraduate education program within maintenance, and the master thesis by Nerland (2010) describes a survey conducted on maintenance vocational training in Norwegian industry in collaboration with the Norwegian Maintenance Organization. The author concludes that the skills level was satisfying but with improvement possibilities, and that maintenance education and training should be developed further on high school, higher education as well as on post-training level.

In this paper, two main research questions will be addressed: What should the engineering students know about maintenance (knowledge and skills), and What kind of maintenance related training (what and how) do the mechanical and industrial engineering students get during their education? For answering question one a literature addressing engineering education as well as the standard EN15628:2014 Maintenance, Qualification of maintenance personnel and the European Federation of National Maintenance Societies (EFNMS) instructions and certificates of requirements for maintenance professionals will be utilized. For addressing question two, the Finnish and Swedish education systems are presented, and thereafter studies regarding engineering education that comprises maintenance training conducted in the previously mentioned countries are accounted for.

2 Maintenance Knowledge and Skills

This section describes the skills and knowledge requirements for the higher education level as described in literature and standards. EFNMS has developed curricula for maintenance training for the maintenance technician as well as for the maintenance manager. According to EFNMS (1998), a maintenance manager should be skilled in following topics: Management and, Reliability performance of production plants, Maintenance information systems, and Maintenance methods and techniques. A curriculum is proposed based on the EFNMS requirements in Nerland (2010) for maintenance training on high school, higher education and post-training level. The headings suggested for the higher education level are: Maintenance objectives and strategies, Maintenance terminology, Addressing failures, Operational safety and reliability, Condition based maintenance, Health, safety and environment, Modern maintenance concepts, Indicators, Maintenance logistics, Maintenance software, Economy, organization, Teamwork and communication, RCM, and Proactive and predictive maintenance. The postgraduate education program within maintenance denoted the MIRCE academy described in Knezevic (1997) consisted of following modules: Maintenance in life cycle engineering, Mathematics of maintenance, Methods and tools of maintenance analysis, Reliability and safety, Maintainability engineering, Supportability engineering, Economics of maintenance, Maintenance integration, Design for maintenance, and Maintenance management. The standard EN 15628 (2014) defines seven competence areas of a maintenance manager: 1) to define and develop maintenance policies according to company strategies; 2) to define processes and tools to support maintenance tasks; 3) to define, manage and develop the organizational model of maintenance; 4) to ensure the levels of availability, reliability, maintainability, supportability, safety and quality required for the entire useful life of assets; 5) to ensure appropriate management and continuous improvement of maintenance; 6) to ensure and control the compliance with maintenance and company budget, the respect of the planned maintenance tasks and the proper condition of assets; 7) to define strategies, policies and criteria for performance management of contractors and for the definition of maintenance materials requirements. Table 1 summarises the knowledge and skills of a maintenance manager according to the literature and standards.

Table 1. Summary of maintenance knowledge and skills of a maintenance manager

Topics	EFNMS	Nerland	Knezevic	EN15628
Terms and standards		x		
Management and organisation	x	x		x
Maintenance systems and planning	x	x	x	x
Maintenance engineering	x	x	x	
Reliability engineering	x	x	x	
Heath, safety and environment		x		x
Maintenance economics	x	x	x	x

3 The Swedish and Finnish Education System

Both the Swedish and Finnish higher education are structured according to the Bologna Process. The first cycle covers years 1–3, the second cycle years 4–5, and the third cycle years 6–9 (PhD level). The workload needed for one full year of studies is correspondent to 60 credits. In Sweden, all degrees are issued in accordance with the The Higher Education Ordinance, Swedish Council for Higher Education (2013). The general qualification for the first cycle is the Degree of Bachelor, which fulfils the general admission requirements to second-cycle studies, i.e. qualification within the first cycle of at least 180 credits. The second cycle general qualifications are the Degree of Master. Both a one-year master (60 credits) and a two-year master degree (120 credits) exists. A Degree of Doctor is normally achieved after four years full-time studies in Sweden. In addition to the general qualifications, a number of professional qualifications are awarded, amongst them two within engineering. The Degree of Bachelor in Engineering is issued after three years of studies. A five year qualification, Degree of Master in Engineering, is also issued. The latter is a popular engineering programme in Sweden, and the admission requirements are higher than for the general qualifications (more advanced maths and science is required). 49 higher education institutions exist in Sweden. Of these, 16 are full universities and 33 university colleges. Main part of the institutions are public, and the education is publicly funded, with few exceptions. In 2018, 22 of these institutions offer 140 Bachelor of engineering programs, and 15 institutions offer 145 Master of engineering programs, Swedish Council for Higher Education (2018). Of the 140 Bachelor of engineering programs, 37 are within the subject area Mechanical engineering, and 13 within Industrial engineering. 18 and 19 are the corresponding figures for the Masters of engineering programs. In the Sect. 3.1 the maintenance related education in these 87 programs is described.

The Finnish higher education system is governed and financed by the Ministry of Education and Culture (Ministry of Education and Culture). The degrees offered at universities of applied sciences are Bachelor's degrees (210–270 credits, 3,5–4,5 years) and Master's degrees (60–90 credits, 2 years). In universities, the degrees are Bachelor's degrees (180 credits, 3 years), Master's degrees (120 credits, 2 years), and third-cycle postgraduate degrees. The calculated timeframe to complete a doctorate is four years of full-time studies. The Finnish higher education system is composed of 40 universities and universities of applied sciences, most of which operate under the Ministry of Education and Culture. In total, the administrative branch of the Ministry of Education and Culture manages 14 universities and 23 universities of applied sciences. The three exceptions are Åland University of Applied Sciences, the Police University College and the National Defence University, which operate under the Ministry of the Interior. In 2018, 30 of these institutions offer 228 Bachelor of engineering programs, and 19 institutions offer 122 Master of engineering programs. Of the bachelor of engineering programs 49 are in Mechanical engineering, and 21 in Industrial engineering. Of the master of engineering programs 20 are within the Mechanical engineering and 25 in Industrial engineering. In the Sect. 3.2 the maintenance related education in these 115 programs is described (Ministry of Education and Culture, Opintopolku).

3.1 Maintenance Education in Mechanical Engineering and Industrial Engineering Programs in Sweden

For collecting information regarding Swedish conditions, three types of sources were used: 1) **The web portal www.antagning.se provided by Swedish Council for Higher Education.** Education programs and courses open for admission are announced here. The search engine was used for finding available engineering programs with start date 2018. This resulted in a first list of totally 110 programs, including duplicates. 2) **The official University web pages.** The main investigation was made on the university web pages. First, a broad search using keywords maintenance and reliability was conducted. All hits indicating programs and courses were investigated further, and the program and course syllabuses were downloaded or copied from the web page. Thereafter, the mechanical and industrial engineering programs were investigated in detail using the headings, subpages and search engines provided for that purpose. 3) **Phone calls to program responsibles.** Programs that have undergone recent changes or where it is mentioned in the syllabus that maintenance related skills are acquired but where no maintenance related course was to be found were investigated further by inquiring the program responsible through phone.

Maintenance related education is found in ten of the 87 programs. The scope of maintenance related education differs greatly, from being a topic in a course to whole programs focusing on maintenance management and engineering. Even if few Mechanical engineering programs on bachelor level contain maintenance related education, two programs exist which are dedicated to the topic; the Bachelor Programme in Maintenance Engineering at Luleå University of Technology and the Degree Programme in Industrial Technology and Production Maintenance at Royal Institute of Technology. These programs constitutes of courses in the maintenance area of 74 credits and 57 credits respectively. The program in Luleå has been offered for over a decade and covers various aspects of maintenance and its management, but also reliability engineering and life cycle costing. The program offered by Royal Institute of Technology starts in 2018 and has a sustainability focus, and covers concepts of maintenance, reliability and dependability, and in addition courses covering the digital and automated factory. In addition to these programs, the Manufacturing engineering program at University of Skövde includes a course in maintenance and operation reliability. This course covers the concepts of maintenance and reliability, and especially the economic aspects and improvement work. In the five-year master variants, two programs offered at Mälardalen University include maintenance related courses. One covers maintenance and dependability while the other course is regarding Total quality Management, Risk management and Lean production. In total five Industrial engineering programs include maintenance related courses, whereas three are in bachelor programs and two in master programs. Reliability engineering is offered at Skövde University addressing reliability analysis and design, but also software quality. Reliability and maintenance are in focus in the course offered for Industrial engineering students at the bachelor program at Linnaeus University. In the economy and production technology program at Chalmers a course covering quality and reliability control is given, covering Total Quality Management, Total productive Maintenance, and Reliability Centered Maintenance. In the five year master program at Umeå University a course in reliability is offered, covering reliability theory and Markov processes. Luleå University of Technology offers a master program in Industrial and Management Engineering, and the civil engineering profile includes a course in operation and maintenance, and their applications for infrastructure. Table 2 summarizes the Swedish findings. Except for the Mechanical and Industrial engineering programs, maintenance courses are found in programs focusing on *Automation engineering*, *Energy technology and environment*, *Civil engineering with railway specialisation*, and *Production development and management*.

Table 2. Maintenance related education in Sweden, Mechanical/industrial engineering

	Program and institution	Maintenance related education (credits)
Mechanical engineering, bachelor level	Manufacturing Engineer, University of Skövde	6
	Bachelor Programme in Maintenance Engineering, Luleå University of Technology	75
	Degree Programme in Industrial Technology and Production Maintenance, Royal Institute of Technology	57
Mechanical engineering, master level	Master Program in Engineering – Dependable Systems, Mälardalen University	About 5 of 10*
	Master Program in Production and product design, Mälardalen University	7.5
Industrial engineering, bachelor level	Study Programme in Industrial Engineering and Management, University of Gävle	7.5
	Industrial Engineering and Management, Linnaeus University	7.5
	Economy and production technology, Chalmers	About 4 of 7.5*
Industrial engineering, master level	Master of Science in Industrial Engineering and Management, Umeå University	7.5
	Master Programme in Industrial and Management Engineering, profile civil engineering, Luleå University of Technology	7.5

*The number of credits is an assessment based on the course syllabus

3.2 Maintenance Education in Mechanical Engineering and Industrial Engineering Programs in Finland

For collecting information about Finnish conditions, following types of sources were used: 1) **The web portal** www.minedu.fi **provided by Finnish Ministry of Education and Culture.** The portal offer general information about Finnish education system. 2) **The web portal** https://opintopolku.fi **provided by Finnish National Agency for Education and Finnish Ministry of Education and Culture.** Education programs and courses open for admission are announced here. 3) **The official University web pages.** The main investigation was made on the university web pages. The mechanical and industrial engineering programs were investigated in detail using the headings, subpages and search engines provided for that purpose. On other programs the search was done using keywords maintenance and reliability.

Maintenance related courses are found in 17 of the mechanical engineering and industrial engineering programs offered by both universities of applied science and universities. The scope of maintenance is a part of course to a couple of courses in the same institution. Häme University of Applied Sciences has a maintenance module of 15 credits. The module is a general introduction to the core concepts of maintenance. Themes are reliability-focused maintenance, production information systems and maintenance budgets. In Kajaani University of Applied Sciences the Bachelor Programme in Production Technology offers 29 credits of maintenance courses. The content of courses are introduction to maintenance, project work, maintenance in mining, vibration and diagnostics. The University of Applied Science of Jyväskylä had in 2015 a "Service Lifecycle Management" program in master level containing in total 60 credits, but the program is no longer given. That program included 28 credits maintenance related courses. In other universities of applied science maintenance education consists of a course or a couple of courses mainly focused to bachelor level. The themes are basics of maintenance, industrial maintenance, maintenance and safety, reliability engineering and life-cycle management, preventive maintenance, CMMS, technical diagnostics, and vibration mechanics. The University of Oulu offers five maintenance related courses in different areas of industries on both master and bachelor level, 5 credit/course. The content of the courses is introduction to maintenance, maintenance of machines, building and infrastructure maintenance. Tampere university of Technology has four maintenance related courses on both bachelor and master level about condition monitoring and diagnostics. Lappeenranta University of Technology has two courses of maintenance in master level about maintenance management and costs. Other universities have only a single course or part of courses in maintenance. However, the *forestry*, *paper technology, process technology, marine technology, mining, energy engineering, technology, communication and transport* and *building technologies* programs have some maintenance courses in their own area of technology. Table 3 summarizes the Finnish findings in Mechanical and industrial engineering programs.

Table 3. Maintenance related education in Finland, Mechanical/industrial engineering

	Program and institution	Maintenance related education (credits)
Mechanical engineering, bachelor level	Bachelor Programme in Mechanical Engineering, Tampere University of Technology	6
	Bachelor Programme in Mechanical Engineering, Karelia University of Applied Sciences	3
	Bachelor Programme in Production Technology, Kajaani University of Applied Sciences	29
	Bachelor Programme in Mechanical Engineering, University of Oulu	10
	Bachelor Programme in Mechanical Engineering, Metropolia University of Applied Sciences	5

(*continued*)

Table 3. (*continued*)

	Program and institution	Maintenance related education (credits)
Mechanical engineering, master level	Master Programme in Mechanical Engineering, University of Oulu	15
	Master of Science in Mechanical Engineering, Lappeenranta University of Technology	4
	Master Programme in Mechanical Engineering, Tampere University of Technology	10
Industrial engineering, bachelor level	Bachelor Programme in Automation Engineering & Industrial Management, Häme University of Applied Sciences	15
Industrial engineering, master level	Master of Science in Industrial Engineering and Management, Lappeenranta University of Technology	6

4 Discussion and Conclusions

For the maintenance related education, engineering students need holistic understanding as well as subject specific knowledge. Table 1 is a comprehensive summary of relevant topics, covering maintenance terminology, system, engineering and management (subject specific themes) as well as maintenance with respect to reliability, safety, health and economy (holistic understanding). In general, maintenance has a small role in mechanical engineering and industrial engineering education in the studied countries. Most programs do not contain maintenance at all. In Sweden Luleå University of Technology and the Royal Institute of Technology and in Finland Häme University of Applied Sciences offer a larger combination of maintenance education while other institutions offer a single or a couple of courses of maintenance within the programs. The content of courses covers the basics of maintenance to advanced topics of maintenance. Maintenance skills and knowledge presented in Table 1 can be covered in the maintenance-focused programs listed above, but in the single courses, maintenance skills and knowledge are only partly achieved.

The main differences between Finland and Sweden in maintenance education are the length of programs and the number of universities giving maintenance courses. In Finland, maintenance related courses are given in more universities than in Sweden. However, maintenance programs in Sweden have much wider scope and more credits than in Finland, and more courses for Industrial engineers. In Finland, a few universities have more than two courses topics about maintenance. For example, Häme University of Applied Sciences offer maintenance module of 15 credits and university of Oulu total 25 credits about maintenance topics but not a program of maintenance. Two basic courses is given in bachelor level and three courses in master level. This paper focus only mechanical engineering and industrial engineering programs, which is a delimitation. The review shows that maintenance courses are given in other programs as well – usually a single basic course and maybe some courses in their own technology e.g. about maintenance of ships, railway infrastructure or electrical systems.

The results are to be seen as a snapshot in time. While talking with colleagues at other universities, it is evident that changes are made in the programs depending on the resources and competencies available. If a person leaves the university, the ability to provide sufficient education within maintenance might be lost, resulting in changes in the curriculum. The opposite could also be seen; the Royal Institute of Technology in Sweden is for instance starting up a new program within the area. It was acknowledged in Nerland (2010) that more support is needed for maintenance related education, for instance in the form of syllabuses, teaching material and textbooks. The standard EN 15628 provides a good basis for the syllabus for maintenance related education, but from an asset management perspective, a similar syllabus covering a wider area than maintenance management and engineering might be a better alternative. As for textbooks, it is important that such are available in the native language, and on the correct level of education. Moreover, as teachers with the competence and ability to train maintenance related topics is missing in many universities, textbooks supporting the general engineering teacher in the introduction of maintenance would be needed. With sufficient and appropriate material, the probability that maintenance and reliability is included as a topic in a related course, such as one in quality management or sustainability, would increase. Another option is to provide online courses that could be incorporated into the education programs without the necessity to acquire the competence in maintenance management and engineering at the individual universities. Here, the possibilities that lies in MOOCS, Massive Open Online Courses, is promising. The MOOC could be directed either to the engineering student, or to the engineering teacher. Collaboration in developing such would be not only of national, but of international interest.

References

Crawley, E.F., Brodeur, D.R., Soderholm, D.H.: The education of future aeronautical engineers: conceiving, designing, implementing and operating. JSET **17**(2), 138–151 (2008)

European Federation of National Maintenance Societies: The Requirements and Rules to Achieve an EFNMS Certificate as a European Expert in Maintenance Management (1998). http://www.efnms.eu/committees/european-certification-committee/

Finnish Standards Association: SFS-EN 15628 Maintenance. Qualification of maintenance personnel (2014)

Heilmann, P., Heilmann, J.: Competence management in maintenance: case – Finnish forest company. Manag. Res. Rev. **35**(1), 4–13 (2011)

Knezevic, J.: Industry driven postgraduate maintenance education: MIRCE approach. J. Qual. Maintain. Eng. **3**(4), 302–308 (1997)

Lai, J.H.K.: Building operation and maintenance: education needs in Hong Kong. Facilities **28** (9/10), 475–493 (2010)

Ministry of Education and Culture: Higher Education and Science (2018). http://minedu.fi/en/higher-education-and-research. Accessed 2 Mar 2018

Nerland, A.S.: Competence within Maintenance. [Master thesis]. Norges Teknisk-Naturvitenskapelige Universitet, Fakultet for Ingeniørvi- tenskap Og Teknologi, Institutt for Produksjons- Og Kvalitetsteknikk (2010)

Opintopolku (2018). https://opintopolku.fi/wp/fi/. Accessed 2 Mar 2018

Swedish Council for Higher Education: The Swedish Higher Education System (2013). https://www.uhr.se/globalas-sets/_uhr.se/bedomning/diploma-supplement/the_swed_high_system.pdf

Swedish Council for Higher Education: Sök utbildning på alla Sveriges universitet och högskolor - Antagning.se (2018). https://www.antagning.se/se/start. Accessed 20 Feb 2018

Managing Competence in Naval Asset Management: Professionalising Defence's Cadre of Asset Managers for Ships and Submarines

Tobias Lemerande(✉)

BMT, Hackney, SA, Australia
tlemerande@bmtdt.com.au

Abstract. Defence's sustainment policy for naval vessels incorporates asset management principles and methods and recognises the need for experts who specialise in asset management as essential to delivering and maintaining maritime capability. For warships and submarines, asset managers constantly weigh technical challenges and costs of maintenance and modernisation against operational and functional benefits to meet the Operating and Support Intent, life cycle management objectives and Seaworthiness requirements. A naval asset management skills model has been developed that focusses on technical compliance, operational capability, and business-related goals intrinsic to Fleet Life Cycle Management. Naval enterprise sustainment efforts are a combined endeavour between Navy, Australian Public Service and defence industry contractors but has no identified or documented minimum asset management capabilities for life cycle management of either individual vessels or entire classes. While Defence acknowledges the need for asset management within the naval enterprise, it has not established criteria for codifying or formally recognising required competencies directly related to life cycle management of naval vessels. National recognition and professional certification formally credits the learning, development and experiential requirements each warship and submarine life cycle management professional should have as a means to become officially sanctioned asset management specialists within Defence's maritime sustainment community. Moreover, the nature of naval vessel asset management necessitates formal recognition within Capability and Acquisition Sustainment Group's Sustainment Management Career Pathway as well as accreditation through the Australian Quality Framework. The US Navy's Port Engineer Program provides a solid example on which Australia's naval enterprise can base its own scheme. This paper outlines a method to recognise naval asset management specialists by codifying experience, education, and training requirements for inclusion in Defence's Sustainment Management Professionalisation and Certification Framework and provides recommendations for establishing a valid career path for Warship and Submarine Life Cycle Managers.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 475–482, 2020.
https://doi.org/10.1007/978-3-030-48021-9_53

1 Introduction

After nearly a decade of research investigating ship maintenance practices and continual improvement activities, Defence's Capability and Acquisition Sustainment Group (CASG) formally identified asset management as a core competency for sustainment of Defence materiel. CASG has mandated that sustainment efforts must align to internationally recognised asset management methods, principles and practices contained in the ISO 5500X series (CASG 2017). Asset management specialists are key personnel who should be accountable for delivering capability across a warship's life cycle (CoA 2016). Australia's national naval enterprise needs competent professionals to specialise in life cycle management of naval ships. The Asset Life Cycle Manager (ALCM) concept contends that competent maritime professionals should be assigned as dedicated individuals responsible for managing a specific Royal Australian Navy (RAN) ship (i.e. surface ship or submarine) on a long-term basis; ALCMs are key elements in the Fleet Life Cycle Management (FLCM) concept (Lemerande 2018b). The ALCM competency model, shown in Fig. 1, incorporates best practices from asset management academia and practitioners, commercial ship management and existing programs used by the United States Navy (USN) and United States Coast Guard (USCG) (Lemerande 2018a).

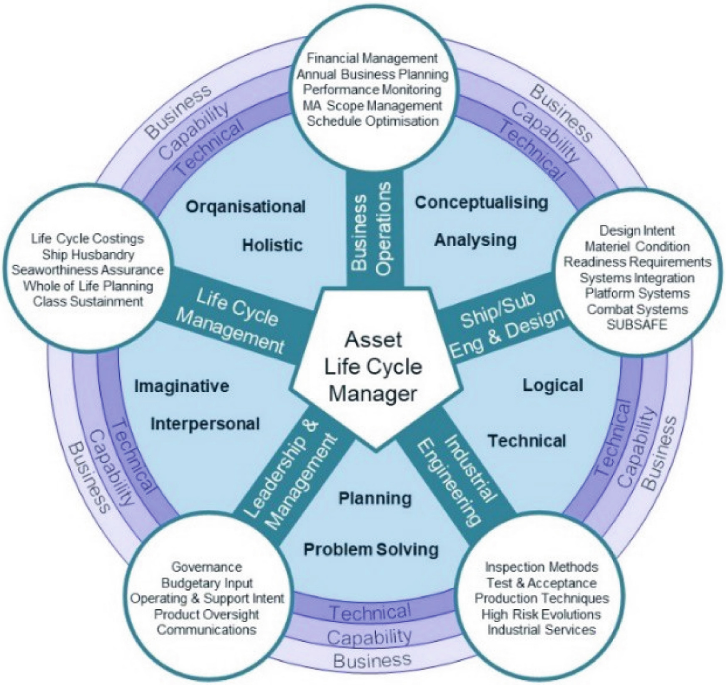


Fig. 1. Competency model for naval enterprise asset life cycle managers

CASG's Sustainment Management Professional Competency Framework (SMPCF) provides an appropriate method for articulating the career path, competencies and qualifications for asset management generalists within Defence (DMO 2014) but it fails to provide the necessary constituent components and factors ALCMs require for life cycle management of a modern naval ship. As a high-level document, the SMPCF provides guidance and requires nondescript sustainment managers to undertake generic competency-based asset management training that can be applied to facilities and infrastructure, land vehicles, utilities, aircraft or any other item categorised as a Defence asset. Warships and submarines represent arguably the most complex and expensive assets within the Australian Defence Force (ADF). Each ship is an intricate and complex system of systems designed to operate in harsh and wide-ranging maritime environments. Each RAN ship is expected to deliver full capability across a 30–40 year service life. Asset management training, qualifications and certifications listed in the SMPCF is woefully inadequate for ALCMs charged with life cycle management of a warship or submarine. A more comprehensive program that addresses competency requirements is needed within CASG's framework. Professionalising ALCMs using elements identified in the competency model could easily be added to the SMPCF. This will formally identify experience, expertise, training, qualifications and certifications needed to document competence for such an important position.

Establishing an ALCM program based on the USN Port Engineer Program (PEP) but tailored to meet naval enterprise needs, would establish an irrefutable world-class qualification that will absolutely deliver greater value to Australia's national naval enterprise. The USN formalised its PEP to establish a rigorous program that is both objective and demanding. It is a proven and effective model that can be tailored by Defence and implemented in Australia to meet CASG and RAN asset management needs. This paper proposes the professionalisation of ALCMs using the competency model as the foundation and the USN's highly successful PEP as a proven model that can be used as the cornerstone for a program in Australia. A brief discussion covering the competency model is provided to explain ALCMs' knowledge domains, areas of expertise and cognitive abilities. Elements covering education, experience, credentialing and qualifications that contribute to the professionalisation of ALCMs is then discussed for consideration. The paper concludes with recommendations as to how Defence can incorporate professionalising the ALCM cadre into the SMPCF for asset management specialists.

2 Competencies

Competencies are a "collection of qualities, abilities, skills and other capacities of the staff needed for successful performance" (Skorkova 2016) in an associated role. Competency models are developed from the identified competencies. An asset manager must have competence in engineering and technical specialties associated with the assets as well as being a competent business manager that can provide clear direction through financial awareness and adept communication skills. Persons responsible for life cycle management require knowledge across technical, business and asset management related aspects that can meet the enterprise's asset-related needs (Hastings

2015). Competence related requirements in ISO 55001 ensure that individual asset management positions are filled by people who have the knowledge, skills and behaviours to perform their duties and can actively demonstrate those attributes to fulfil requirements needed in specific asset management roles (AMC 2014). Strategic asset managers apply engineering, strategic management and financial skills to develop, plan and maintain assets in an economically responsible and sustainable manner over the assets' life cycles. The ALCM concept aligns closely to stewardship theory for asset management in that it includes specific aspects of asset governance and oversight. As front-line technical life cycle management experts, ALCMs could easily fill the strategic asset manager role for a naval ship. These persons are expected to be highly trained and skilled and exhibit a wide-range of competencies across various specialty areas.

Competence is critical to achieving a state of professionalism which is characterised as specialised expertise and an ideological approach to the work that is "based on clear principles, including a commitment to the interests of the client" (Mills 2014) with individuals being paid for full-time work in that specialty. Competence management requires the appropriate combination of thinking and practical skills accompanied by knowledge and comprehension of asset-related activities. Competence should be treated firstly as an official authority (by position) and secondly as a personal authority that an individual maintains at a particular point in time. The competence may be 'professional' meaning it is related to a specific profession (i.e. marine engineer or naval architect) or 'managerial' in that it relates to managerial processes and relations. Professions have been characterised by the following qualities: specialist knowledge; required credentials for inclusion; regulated activities; and a common set of values binding the group together (Susskind and Susskind 2015). A program that professionalises asset management specialists has been cited as a critical success factor to asset management implementation in Defence (DSTG 2015).

3 Professionalisation

Professionals are often described as persons with degrees of specialisation in discrete fields that relate to the development of new knowledge or skills (Lilleker and Negrine 2002). Professionalisation refers to a social process an occupational group takes regarding one or more elements of an ideal type of profession (Vollmer and Mills 1965). Professionalisation has five constituent components: the work is performed full-time and the job is considered permanent; an organisation encourages established training through a scheme, regimen or formal schooling; a professional association forms around the occupation; the job title is protected by law; and there is a formal and established code of ethics (Wilensky 1964). Through an appropriate mix of education, experience and recognised qualifications, ALCMs can attain a recognised pedigree through a formal certification process that acknowledges competence and professionalises the ALCM role in CASG's SMPCF.

3.1 Education

ALCMs should have nationally recognised engineering degrees. Education and academic accomplishment hold a prevailing power in professionalisation in technical disciplines. University level education provides a level of assurance of minimum engineering competency, both theoretical and practical. An engineering undergraduate degree provides the basis for not only sound engineering practice but also meets requirements necessary for recognition by professional engineering bodies and associations like Engineers Australia (EA) or the Institute of Marine Engineering, Science and Technology (IMarEST). In many countries, post graduate training is a very efficient way of improving the qualifications of maritime professionals because it helps introduce them to the latest advances and developments in the maritime sector. Formal education should also include targeted short courses that specialise in specific business, technical or maritime capability aspects that support the ALCM job description. Several Australian universities offer advanced courses in sustainment and engineering asset management, which would also support attaining or maintaining competency. Industrial management training following undergraduate coursework should not be discounted; it can be a viable option for providing valuable training for maritime specialists outside of the university setting. Non-resident business-related short courses are readily available from Australia Institute of Business (AIB) or Australia Institute of Project Management (AIPM). Basic courses in asset management and ISO 55001 are available from registered training organisations. (These competency-based training courses are currently listed within the SMPCF and provide a basic approach to asset management but are wholly inadequate for ALCMs.) Commercial training providers like Life Cycle Institute partner with universities to offer a complete range of courses covering different aspects of life cycle management. Courses directly related to submarine engineering and design, warship capability management and FLCM are also available from commercial vendors within Australia. Special courses developed specifically for ALCMs can provide focussed attention to better meet naval enterprise needs.

3.2 Experience

Experience should be measured qualitatively and quantitatively. Qualitative experience can be measured by years of working in related fields and other types of jobs performed while serving in specific maritime asset management roles. Involvement in Australian naval operations as a member of a ship's company can provide valuable experience for an ALCM. Seagoing experience in allied navies provides an appropriate level of naval operations knowledge with ties close enough to the RAN that would be sufficient. Past work in seagoing maritime operations, either commercial shipping or deep-water drilling would also provide a substantial level of applicable maritime experience directly relatable to the ALCM position. Less applicable, but still noteworthy, would be life cycle engineering and management experience in commercial port/marine operations or land-based asset intensive fleet industrial operations in rail & transport, infrastructure, utilities or commercial aircraft sectors. Past experience serving as either a Chief Engineer, Port Engineer or Ship Manager in the commercial shipping industry would be invaluable and heavily desired in ALCMs. Qualitatively, this would provide a certain level of assurance that an ALCM has the relevant experience and expertise to competently perform the assigned duties. Quantitatively, experience can be measured by the number of years an individual has served in previous roles or by the number of maintenance periods that have been managed and the total maintenance and operations budgets for those specific vessels as well as the size and complexity of vessels managed. Previous work with commercial contract management could provide the appropriate business-related experience ALCMs need.

3.3 Third Party Recognition

Third party credentials provide supporting objective evidence of competence and compliance against established, recognised and accepted standards. Achieving and maintaining recognition from established authorities outside of Defence lends greater credence to claims of education, experience and expertise in areas specifically identified as critical to ALCMs' roles and responsibilities. Modern credentialing contributes to legitimatising professionalised occupational groups based on the possession and application of unique technical qualifications, knowledge and skills that are crucial and indispensable to highly skilled professions. Formal credentials, such as licences, can be used as a method to certify mastery in a specific body of knowledge. Credentials that address business, technical and maritime/naval operations are needed to support the tremendous tasks that have been outlined in the ALCM competency model. Such credentials can be attained through various Australian organisations.

ALCMs are expected to be highly experienced and thus should have an AMSA issued Engineer Class 1 licence to validate a person's competence in structural, mechanical and electrical shipboard systems. A qualification in marine surveying signifies a person's competence in understanding and applying established standards to determine a vessels' compliance. Each ALCM should be a Chartered Professional Engineer (CPEng) with EA. Formal credentialing from AIPM signifies experience, knowledge and competence in high risk, high profile and challenging projects. Formal certification in asset management from Australia's Asset Management Council is recognised nationally and internationally and attests to an individual's knowledge and skills and demonstrates a person's ability to improve asset performance through application of asset management methods, practices and principles. Certificates of competency in either Marine Engineering or Weapons Electrical Engineering attained through active service in the RAN would be highly desirable for an ALCM. Formal recognition in engineering, business, project management and specific maritime industry activities will be necessary for the ALCM cadre to be established as a recognised authority in life cycle management of naval ships.

3.4 Formal Qualification and Recognition by Defence

The USN PEP has four levels of certification. Each stage of certification requires candidates to satisfactorily complete a demanding program that includes formal course work, self-study, practical demonstration of activities and successfully passing standardised written examinations (USN 2014). A similarly tiered qualification program could easily be developed for ALCMs. ISO 55001 requires an organisation to ensure asset management professionals "are competent on the basis of appropriate education, training, or experience" (SA 2014). Defence policy requires formal asset management certification for all sustainment management roles. Certification should be based on an assessment against known competencies in sustainment and from AQF accredited Asset Management competencies (DMO 2014). The ALCM qualification must go beyond this and should be developed to cover all competency model elements while also providing rigour to withstand scrutiny by naval enterprise stakeholders. Establishing an ALCM Qualification Program inside Defence will provide an objective qualification that is specific to asset management for naval vessels and will also provide a career path for the cadre of ALCMs.

4 Conclusion

ALCMs require education and experience, credentials recognised by professional entities and organisations (outside of Defence) to provide an objective assessment of knowledge, skills and abilities. Managing ALCM competencies through a formal program provides assurance that assigned ALCMs not only understand their role but that they have the fundamental elements required to perform lifecycle management duties and responsibilities for the assigned ship. They should also have formal qualifications specific to the ALCM role inside of Defence. ALCMs can be professionalised through establishing a formal program that encompasses education, experience, third-party credentialing and formal qualifications recognised and acknowledged by Defence as required for the ALCM role. Using the USN PEP as a model, Australia's naval enterprise can easily develop and implement its own professionalisation regime to produce elite professional life cycle managers for some of the most complex naval assets in the world.

References

Asset Management Council: Companion Guide to ISO 55001 Edition 1.1, Victoria (2014)

Capability and Acquisition Sustainment Group: Functional Policy (PM) 001: Sustainment Management in Capability and Acquisition Sustainment Group Vers 2.0. Commonwealth of Australia, Canberra, 11 August 2017 (2017)

Commonwealth of Australia: Interim Capability Life Cycle Manual. Department of Defence, Canberra (2016)

Defence Materiel Organisation, DMO Sustainment Management Professionalisation and Certification Framework. Canberra, 2 September 2014

Defence Science and Technology Group: Effective implementation of asset management in Defence, DSTG Minute signed by A. Zelinsky, CDS/OUT/2015/87 of 17 September 2015 (2015)

Hastings, N.A.J.: Physical Asset Management: With an Introduction to ISO 55000. Springer International Publishing, Switzerland (2015)

Lemerande, T.J.: Developing asset life cycle managers: essential skills for managing complex naval assets in the 21st century. In: E-Proceedings of AMPEAK 2018, Hobart, Tasmania, 16–18 April 2018 (2018a)

Lemerande, T.J.: Leading the naval asset management effort with a framework for fleet life cycle management. In: E-Proceedings of AMPEAK 2018, Hobart, Tasmania, 16–18 April 2018 (2018b)

Lilleker, D.G., Negrine, R.: Professionalization: Of What? Since When? By Whom? Harvard Int. J. Press/Polit. **7**(4), 98–103 (2002)

Mills, S.: Professionalisation: Of What, Since When, and By Whom. Working Paper No. 29, University of Sydney (2014)

Mills, D.E., Brown, K., Waterhouse, J.: Asset management stewardship: the effectiveness of public-private mix governance structures. In: 2008 First International Conference on Infrastructure Systems and Services: Building Networks for a Brighter Future (INFRA), pp. 1–5. IEEE, November 2008

Skorková, Z.: Competency models in public sector. Proc.-Soc. Behav. Sci. **230**, 226–234 (2016)

Standards Australia: Asset Management – Management Systems – Requirements, AS ISO 55001: 2014. Council of Standards Australia, Sydney (2014)

Susskind, R., Susskind, D.: The Future of the Professions: How Technology Will Transform the Work of Human Experts. OUP, Oxford (2015)

United States Navy, Port Engineer Training Manual: A User's Guide. USN, United States (2014)

Vollmer, H.M., Mills, D.L.: Some comments on "The professionalization of everyone?". Am. J. Sociol. **70**(4), 480–481 (1965)

Wilensky, H.L.: The professionalization of everyone? Am. J. Sociol. **70**(2), 137–158 (1964)

Master Student Teaching and Training in an International Context

Jan Frick(✉)

University of Stavanger, Stavanger, Norway
jan.frick@uis.no

Abstract. Erasmus plus Strategic partnership are a program for collaboration between European Universities and companies. The projects Colibri 2014–2017 (Collaboration and Innovation for Better, Personalized and IT-Supported Teaching) and EPIC 2017–2020 (Improving Employability through Internationalization and Collaboration) both have goals of testing and improving teaching methods. The Colibri project goal was to test methods in delivering a common industrial relevant course for Master students from both a business and an engineering background. The test course was "Future Internet Opportunities" that merged internet technologies with internet usage and entrepreneurship. The industrial cases came from a wide range of industries and organizations. Given the Colibri results, the project EPIC started in September 2017. The purpose is to extend the experiences in Problem Based Learning Projects from an international course into Master (and Bachelor) Thesis work.

1 Introduction

For more than a decade now, there has been a strong increase in the proportions of students in European Union studying in a foreign country. At the same time, we see a strong increase in the number of students completing their tertiary education. Thus, student groups attending higher education are becoming more diverse due to growing student mobility and due to the growing proportion of the population entering higher education. The growing diversity combined with a growing proportion of the population completing tertiary education leads to greater differences in the students' academic background, learning styles and cultural values within academic institutions across Europe. This diversity challenge the effectiveness of our pedagogical approaches, and suggest that we should adapt courses to the needs and abilities of each student. This includes a need to improve the industrial relevance of the student work since the number of students with any previous industrial experience is decreasing. However, a recent study shows that student diversity in Europe may be a pretext for *boosting* academic performances, not only a challenge to overcome: In a survey of 670 international business students and 288 domestic business students, it was found that students' efforts to integrate academically[1] correlated positively with academic

[1] The following efforts to integrate academically correlated with academic performances; Adjusting to the options at the university, emotional adjustments, and a sense of attachment to the university.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 483–490, 2020.
https://doi.org/10.1007/978-3-030-48021-9_54

performances. The scores of foreign and domestic students did not differ significantly (Rienties et al. 2012).

We also see that recent trends indicate that an increase in student diversity does not seem to affect collaboration through the Internet negatively. The frequency of internet use among young people in the European Union is high, and less varied, compared to the population in general. In 2015 the frequency of daily internet user varied between 37% and 92%, with an average of 67%, among the general population in the 28 EU member states. The proportion among the young population (16–29 years) varies between 69% and 98%, with an average of 89% in the 28 member states (Eurostat 2015).

2 Colibri and Blended Learning

Into this context we applied blended learning in the course entitled "Future Internet Opportunities", which was developed and run as a deliverable in the Erasmus plus project "Colibri"[2]. The overall idea of the Colibri project is to implement new and innovative teaching methods, and to establish a Living Lab[3] of students from different universities, topics and countries/cultures for a systematic testing and evaluation of these methods. The course targets a relatively small group of master students from universities[4] in seven countries. There are at least four students per university, and some of the students also have an origin from other parts of the world. The students have either an engineering background with different specialization or a business strategy background. The varied backgrounds of the students, and geographical spread of their home universities, provided us with valuable experiences regarding impact of blended learning techniques. Our findings are based on written and oral student evaluations, grades, and internal evaluation by teachers and other Colibri-participants.

The "blended learning" described cover learning sessions, interactions between students, and interactions between students and instructors. Much of the literature recognizes that blended learning is not just about finding the right mix of technologies, or increasing access to learning, it is about rethinking and redesigning the teaching and learning relationship. Blended learning offers possibilities to create transformative environments that effectively facilitate critical and creative thinking skills.

Blended learning requires a coordinated effort from both the teaching faculty and the leading administrators of the academic institution in policies, planning, scheduling, and supportive functions for students and teachers. In the Colibri case, we take most of these institutional requirements for granted as the course is set up by a multinational

[2] Colibri was funded by Erasmus+ in the EU. The acronym stands for "Collaboration and Innovation for Better, Personalized and IT-Supported Teaching".

[3] The concept "living-lab" originates from the MIT Media-Lab applying a user-centered research methodology for observing, prototyping, validating, and refining solutions in a real life context.

[4] Colibri partner universities were located in Aalborg, Stavanger, Istanbul, Barcelona, Bydgoszcz, Riga, and Hamburg.

and specially dedicated teaching staff, supported by staffs of the participating universities. The goal of Colibri is to support the development, transfer, and implementation of innovative practices within higher education, and to develop creative and innovative teaching methods (Lopez et al. 2015). These teaching methods include online modules that utilize a variation of text, videos, quizzes, peer learning, etc., and Problem Based Learning in international groups on problems provided from real companies.

2.1 Course Content

"Future Internet Opportunities" was divided into 4 distinct parts (see flowchart Fig. 1). Each of the ten modules had an introductory level that all student had to follow. These modules covered topics from Internet technologies like hardware and software for different types of communication, Internet usage and application, Security, Internet Architecture, and Entrepreneurship. The introductory level was followed by a basic level where each student choose five modules, and then an advanced level where students selected two modules. In a five-day midterm workshop the students presented one of their advanced modules. In the midterm week students also trained in presentation techniques, and company cases were handed out. The cases address real problems provided by companies that in 2015, 2016 and 2017 were located in Athens, Barcelona, and Berlin. The students then solved these cases in groups consisting of mixed nationalities and educational backgrounds. During this part of the course they were at home and had to communicate via Internet. The last part of the course was a 5-day workshop where the students are challenged by teachers on how to interpret and present their assignment. At the final day of the workshop each group presented their solution/paper and defended it before the other students and the collegium of teachers. All students had English as a second language and there was a mix of engineering and business students, but the majority had an engineering background.

An overlap between the work on the online modules and the projects, was productive to:

- explore different learning tools and learning strategies,
- question any assumptions made during the course,
- utilize the module content in problem based learning cases,
- learn how to learn individually, and collectively, face-to-face and mediated through electronic networks.

This covers what Chris Argyris and Donald A. Schön refer to as single-loop learning, double-loop learning, and deutero-learning. ((Argyris and Schön 1978); (Easterby-Smith et al. 2004)). The virtual start-up meeting and the transnational project groups were important arenas for socializing as well. This was possible because each of the virtual meetings were preceded by face-to-face meetings at host universities before the start-up meeting, and in five-day workshops before the network-based work in the transnational project groups.

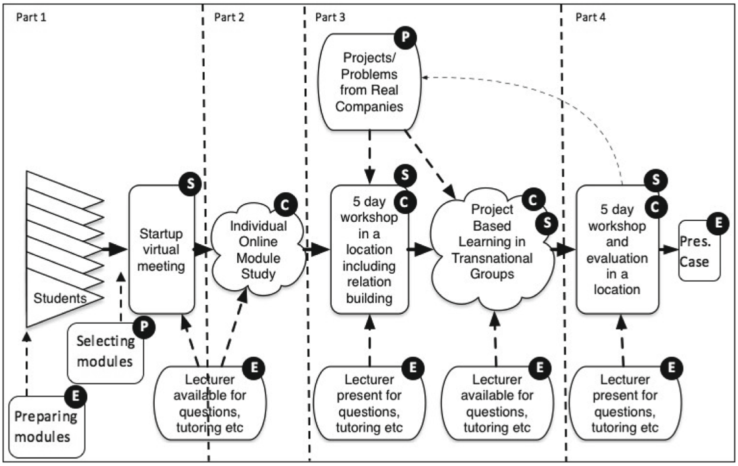


Fig. 1. Flowchart of the Future Internet Opportunities Grad Course. The capital letters refer to the section on "Elements of blended learning in the course" below. "P" = "personalize", "S" = "socialize", "E" = "explore new pedagogical approaches", and "C" = "collaborate".

2.2 Colibri Experiences

All students provided feedback after finalizing the course. The general feedback was quite good as they liked to meet other students from other universities and to have variation in how modules were taught. The main issues emphasized from this individual feedback were the following:

1. the usage of the module content in Problem Based Learning provided a different understanding than just remembering the module content that learning
2. working together with students from different European and non-European nations gave them a better foundation for international employment and higher probability for employment in general.

Experiences for our evaluation are gathered by observations, written feedback, quality committee meetings (with one student from each country), and feedback in common evaluation meetings with all students present. We have identified some critical strengths in this project:

– The course literature covered not only technology or business, but aimed to cover the range from basic technology via its utilization to the business case. This got a unanimous approval from the students as their home institutions tends to keep these issues separate in different courses.
– The course accepted students with both a technical and a business background, and they select modules depending on their background. This mix of profiles was regarded an advantage of the students when they worked on the case from real companies in the last part of the course.

– With an exception of Aalborg University, the students are not familiar with complex cases from real companies in problem based learning. But to investigate these in a multinational group with a mix of modules was received positively by the students.

The positive feedback from students seemed to be related to the high degree of diversity in this course: Students commended the diverse background of those participating in the course, the cross-disciplinary curriculum, and the realistic case assignments covering real industrial problems in different countries. These characteristics are well adapted to meet the more diverse and mobile student population in the EU.

3 Issues in Higher Education and Team Learning

We see that the experiences in this course relate to well-known mechanisms in the literature on higher education and team learning:

Communication and trust in global virtual teams: Students seemed to improve collaboration due to the personalization and socialization activities. This is in line with Jarvenpaa and Leidner who emphasize the need for social communications that complements, rather than substitutes, task communications in virtual teams (Jarvenpaa and Leidner 1998). The variation in background might have limited the learning experiences but the coming together with people they have only seen on net meetings seems to encourage learning.

Realistic approach to entrepreneurial education: Students seemed to get a better understanding on relation between technology, its use, and business value creation. This was much due to collaboration on cases for real companies but it also included a wider scope due to the multinational background. An example of that is that the companies that provided the problem reported benefits from the presentation made by the student group. Entrepreneur processes must be allowed to be creative and devote time to open-ended problems and question conventional wisdom. Kirby points out that one of the crucial elements in education programs for entrepreneurs is to involve student teams in problem-solving based on real-world situations (Kirby 2004).

Effective virtual learning environment: Students depended on this to become prepared for the Midway seminar, and to collaborate between the seminars. The students knew that their work in the advanced modules was to be presented in the midway seminar, and later that they in the Problem Based Learning had to collaborate to produce a result at the end of the course, so they needed to use the virtual environment in a productive way. This dependency on team collaboration is recognized as something that stimulates the use of virtual meetings. It has been found that virtual teams that are high on trust, and that are faced with task uncertainty early in the team's life, were the best problem solvers (Jarvenpaa and Leidner 1998). This suggests that dependency on team collaboration is useful in virtual learning environments.

Project based group learning: Most students are not used to this pedagogic approach, but finds it an improvement as it provides a process oriented learning as a contrast to an exam oriented approach. Project or problem based learning are more oriented on a process approach than actual reproducing knowledge at an exam (Kirby 2004). We provided the students with real world issues that they had to sort out, analyze, and come up with a solution or opinion about. It was complex cases with no premade solution available and the students knew that their work would be presented to the companies.

Individual learning contribute to team learning: The project based learning in groups have a better learning process as students have to explain their views for the other students in the groups to achieve a common result. When we observed the learning process in the course we saw that the students started with a personal learning as they are used to. But during the personalization and socialization process this changed. And the case solved in the last part of the course had much more elements from team learning as they needed to understand and then explain to the other group members from other countries and backgrounds. The link between individual learning and team learning, by which individuals influence shared mental models, is described in the seminal article on how individual learning affects team learning (Kim 1998).

The overall impression from Colibri also included that blending student background in terms of engineering with business students, and students from different countries improved the problem-solving in Problem Based Learning Projects. And also blending the learning methods into this context gave very high scores in student feedback in terms of employability and international relevance.

4 The EPIC Continuation

Given the Colibri results, the project EPIC (Improving Employability through Internationalization and Collaboration) started in September 2017. The purpose is to extend the experiences in Problem Based Learning Projects from an international course into Master thesis work. We also accept Bachelor thesis and similar project work. In EPIC we have 8 universities from Aalborg, Riga, Bydgoszcz, Enschede, Barcelona, Kayseri, Hamburg, and Stavanger, and companies from Barcelona and Berlin. Students are set up in transnational teams to solve problems stated from real companies. We have 2 companies as project partners, but the problems comes from several other companies also.

Compared with Colibri we reduce the emphasis on modules for learning as we assume that students have relevant background or can add on a need to know basis. We keep emphasis as in Colibri on updating knowledge on group work, problem based learning projects, and presentation skills. And we increase the tutoring side of the work which include a main supervisor from the students home university, in some cases an additional supervisor from the university where the student have a student to collaborate with, and a more active contact person in the companies compared with Colibri.

There are some challenges in the EPIC setup that we partly expected. As all students have to deliver their work according to the rules of their home university, we get student teams with high variation in requirements. As example, we have one bachelor student that have to deliver in June working together with a Master student that have to deliver in October. A consequence is that only the first part of the master student work is a collaboration. There are also different expectations about being located at the company or not, travels, and defense of thesis.

5 Summary

The Colibri project got a letter from the EU commission in June 2018:

"Your project **2014-1-DK01-KA203-000764- "Collaboration and Innovation for Better, Personalized and IT-Supported Teaching"** has been selected as a "success story" by a panel of experts from the Directorate-General for Education, Youth, Sport and Culture of the European Commission. "Success stories" are finalized projects that have distinguished themselves by their impact, contribution to policy-making, innovative results and/or creative approach and can be a source of inspiration for others. The selection of your project as a success story was made on the basis of rigorous criteria regarding its quality, relevance and results.… As a consequence of this selection, visibility and acknowledgement will be given to your project, for instance on our websites, social media, and when preparing documentation for conferences or other events with high-ranking attendance.«

The project EPIC ended its first cohort summer 2018. Some of the results are excellent and some has issues. The positive side includes better integration with industry and increased effort by the students which we think relates to Master thesis being more important for both students and companies than a course. What we need to look into includes communication and socialization across the cohort which relates to having only one physical workshop (the students does not expect to meet all the others again,) and the nature of a Master thesis which is regarded much more as an individual effort by many students than a course. There is also an issue on how to combine business students and engineering students when they report to and get evaluation by their different home universities even if they work on the "same" project with the same company.

References

Argyris, C., Schön, D.A.: Organizational Learning: A Theory of Action Perspective. Addison-Wesley Publishing Company, Cambridge (1978)

Easterby-Smith, M., Antonacopoulou, E., Simm, D., Lyles, M.: Constructing contributions to organizational learning. Argyris and the next generation. Manag. Learn. **35**, 371–380 (2004)

Eurostat: Eurostat. Table Internet activities – individuals (2015). http://ec.eu-ropa.eu/eurostat/data/database?node_code=isoc_ci_ac_i.015. Accessed 4 May 2016

Jarvenpaa, S.L., Leidner, D.E.: Communication and trust in global virtual teams. J. Comput.-Mediat. Commun. **3**, 1–36 (1998)

Kim, D.H.: The link between individual and organizational learning. The Strategic Management of Intellectual Capital, pp. 41–62 (1998)

Kirby, D.A.: Entrepreneurship education: can business schools meet the challenge? Educ.+ Train. **46**, 510–519 (2004)

Lopez, J.M., Frick, J., Kirikova, M., Solé-Pareta, J., Pedersen, J.M., Tran, N.: The Colibri Project: overcoming diversity in blended e-learning activity preparation. In: Proceedings of the 43rd SEFI Annual Conference 2014: Diversity in Engineering Education: An Opportunity to Face the New Trends of Engineering. SEFI: European Association for Engineering Education (2015)

Rienties, B., Beausaert, S., Grohnert, T., Niemantsverdriet, S., Kommers, P.: Understanding academic performance of international students: the role of ethnicity, academic and social integration. High. Educ. **63**, 685–700 (2012)

Part VII: Asset integrity, Maintenance, and Logistics

Development of Modern Maintenance Management Strategy for Complex Manufacturing Assets

Abdu Shaalan(✉), David Baglee, Michael Knowles, and Salla Marttonen-Arola

University of Sunderland, Sunderland, UK
bg60so@research.sunderland.ac.uk,
{david.baglee,michael.knowles,
salla.marttonen-arola}@sunderland.ac.uk

Abstract. Modern manufacturing organizations are designing, building and operating large, complex and often ‘one of a kind’ assets, which incorporate many different electrical, electronic, hydraulic and mechanical systems and components. Due to this complexity and the need to react quickly to changes in production, there is a need for more advanced strategies to ensure effective and efficient high maintenance and high availability. Modern maintenance strategies including Total Productive Maintenance and Reliability Centred Maintenance (RCM) have proven successful. With the increase in complexity and a large number of interconnected components, however, the use of one single strategy may not provide the necessary detailed system to support a maintenance task selection. The paper will propose that for complex assets with a large number of systems, a new framework will be proposed utilising RCM, and include additional supporting systems, including Value Stream Mapping and Cost Benefit Analysis.

1 Introduction

Manufacturing equipment has become more complex to remain competitive, and to produce higher quality products quicker and cheaper. The increment in the manufacturing process complexity is shown in the high number of processes the product has to go through to reach high quality and reliability. This will result in an enormous amount of machines engaged together in the production line to achieve one complete successful production cycle. The optimal maintenance of modern, complex assets with various systems and interdependencies requires aligning the maintenance recommendations, provided by the asset or component manufacturers, with the historical maintenance data of the system. However, in practice many manufacturing organizations must deviate from the operating conditions and loads recommended by their component manufacturers when operating complex ‘one-of-the-kind’ assets. This leads to unexpected reliabilities and breakdown times. On the other hand, the availability of historical maintenance data often causes issues. When available, these data tend to provide understanding about the asset health, age and reliability over the years and the type of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 493–499, 2020.
https://doi.org/10.1007/978-3-030-48021-9_55

maintenance applied in the assets. It can be argued that these complex, unique assets require tailored maintenance approaches, combining features from various existing strategies to optimally integrate the management of asset reliability and financial performance. In this paper the features of the existing maintenance strategies, especially RCM and TPM, are combined with the principles of Value Stream Mapping (VSM) and Cost Benefit Analysis (CBA) to introduce a framework for a more comprehensive maintenance strategy for complex assets. The research has been conducted in collaboration with an industrial manufacturing company, which operates and maintains a large, unique and complex asset.

2 Literature Review

Maintenance strategies developed since 1900 can be categorised into three different generations (see Table 1).

Table 1. Maintenance evolution.

Period	Equipment characteristics	Maintenance philosophy
1st Generation (1900–1939)	- Simple and easy to repair	Corrective Maintenance
2nd Generation (1945–1979)	- More complex - High maintenance cost in relation to operation cost	- Preventive Maintenance - Total Productive Maintenance (TPM) - Total employees participation - Team based approach to maintenance
3rd Generation	- High complexity - Expensive downtime - Inventory management and Just-In-Time usage - High quality products required	- Condition Based Maintenance (CBM) - Reliability Centred Maintenance (RCM) - Computer aided maintenance management - High focus on equipment availability and reliability

Corrective Based Maintenance (CBM) is the base strategy for maintenance and the first been used in the manufacturing sector. CM tends to restore the equipment into operation condition without any consideration to analysing the failures, which took place. ***Preventive Maintenance (PM)***, on the other hand, is a maintenance scheduling strategy that tends to apply periodic maintenance tasks based on fixed periods on assets to lower the equipment's breakdown rate. ***Total Productive Maintenance (TPM)*** is considered an effective maintenance philosophy, as it tends to improve machine reliability and employee morale (Rahman 2015). The main goal of TPM is to eliminate the barrier between operators and maintenance staff by engaging machine operators in performing simple maintenance tasks and ensuring the engagement of all various levels of management (Ng et al. 2011). TPM tends to perform its function by using nine main pillars:

1. General housekeeping, or 5S
2. Autonomous maintenance,
3. Focused maintenance,
4. Planned maintenance,
5. Quality management,
6. Education and training,
7. Safety, health and environment,
8. Office TPM, and
9. Development management (Rahman 2015).

Each of these pillars tends to engage with the system in a specific way, but in return, they all tend to affect each other in increasing the whole performance. According to Baglee et al. (2016), TPM increases equipment's OEE by eliminating waste, increasing availability and enhancing the productivity in general by increasing employees' discipline and the commitment of different management levels towards improving the workplace. Chan et al. (2005) presented the impact of TPM implementation at an electronics manufacturing company with 83% improvement in the equipment's productivity. Bon and Lim (2015) applied TPM in the automotive sector which resulted in decreased downtime and increased machine availability.

Carretero et al. (2003) presented ***Reliability Centred Maintenance (RCM)*** as a method used to identify the required PM plan for a complex system. In 1992, the first RCM standard (GJB1378) was created and started to be used in practice (Cheng et al. 2008). RCM is commonly used in sectors such as nuclear power, aviation, rail, metallurgical industries, shipping and water distribution (Selvik and Aven 2011). With the continuous development of the strategy, different new updated models for RCM have been developed, including RCM 2 (Moubray 1991), Stream lined RCM (Bookless 1999). In addition, Vishnu and Regikumar (2016) presented RCM as optimum mixture of preventive, reactive, proactive, condition based and time-based maintenance, that tends optimize the optimum PM plan for a complex system with best cost efficiency (Fig. 1).

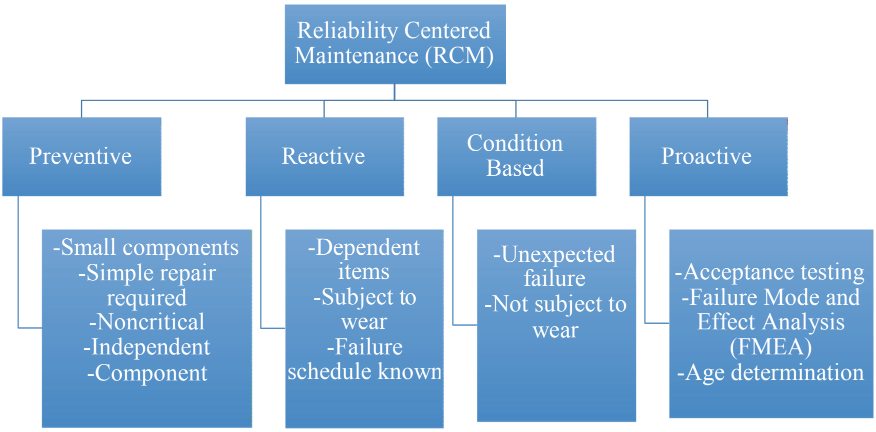


Fig. 1. RCM as a combination of maintenance strategies (Sainz and Sebastian 2013).

One of the main things to be considered when implementing RCM is that the design phase of the inspection manual and procedures should be executed by a technical team that has both field and design experience. This will ensure that the technicians' tasks are aligned with the design guidelines. Another key aspect for sufficient maintenance implementation is the proper assessment of the system design for maintainability (Igba et al. 2013). Key factors in RCM implementation include proper training for the RCM team (Cheng et al. 2010).

According to Moubray (1991), RCM analysis works mainly on finding the sufficient answers for the following seven questions:

1. What are the functions and associated performance standards of the equipment in its present operating context?
2. In what ways can it fail to fulfil its functions?
3. What is the cause of each functional failure?
4. What happens when each failure occurs?
5. In what way does each failure matter?
6. What actions to prevent each failure?
7. What actions if a suitable preventive task could not be found?

As a Japanese philosophy, ***Value Stream Mapping (VSM)*** is defined as a tool that helps the manufacturer to visualize and have more understanding of the flow of materials and data in a production line. It also contributes in visualizing time and materials waste in an entire manufacturing system. VSM is involved in all of the process steps, both value added, and non-value-added activities are analysed. The goal of VSM is reorienting the production practices in the work place to align with lean thinking and establish future improvement plans to eliminate waste in the manufacturing system (Black and Phillips 2010). Toivonen and Siitonen (2016) stated that VSM tends to improve the value stream by focusing on the whole system instead of individual parts. Patel et al. (2015) presented multiple case studies on the implementation of VSM in different sectors including railway services and manufacturing assembly line and more that had high impact on the reduction of work in progress inventory, processing time, lead-time, and delivery time for the whole system.

Cost Benefit Analysis (CBA) is a method used by decision makers to evaluate the efficiency of a proposed or currently running policy from an economical perspective. CBA is dealt with as one of the conventional methodologies in economic analysis for evaluating project alternatives and proposed changes. CBA is known as a systematic and rational decision support tool, which makes it one of the most accepted tools in economic analysis. Furthermore, CBA enables comparing and capturing costs and benefits of any investment the project might have (Djukic et al. 2016).

3 Proposed Framework

The combination of the mentioned concepts along with considering the condition and the specification of the case study asset will help in the development of a hybrid framework that will help to overcome many difficulties that currently face the advanced maintenance practices. These difficulties will include, increasing the maintenance plan efficiency through the perfect maintenance plan optimization, increasing the equipment availability and better cost optimization (Fig. 2).

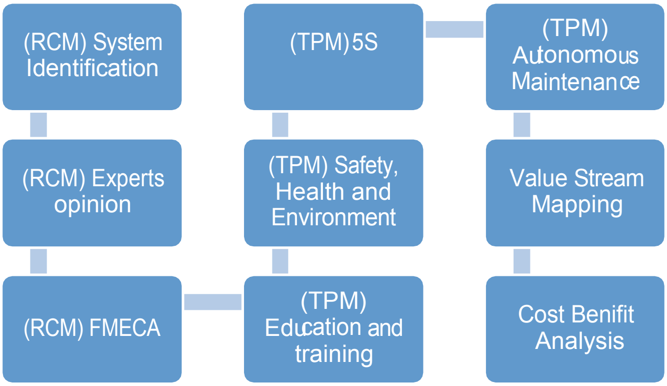


Fig. 2. Proposed framework

The proposed framework tends to mix certain pillars from different maintenance strategies in a specific order that is believed to have optimum efficiency implication. The framework starts with identifying the system under study and understanding its mode of operation as well as all the associated aspects related to production. Expert opinions are crucial, as they highlight different aspects and issues that are hard to identify through analysis alone.

By the stage of applying Failure Mode Effect and Criticality Analysis (FMECA), in depth understanding of all system's equipment and related machines will be understood and critically assessed. Specific Education and Training requirements must be identified and addressed Safety, Health and Environment impact training provided to increase awareness of all different levels of employees of all the associated implications related to it for best environment preserving and safety practices. First interaction with the work place is believed to be through the application of 5S practices, that tends to target the workplace Order, Organization, Tidiness and maintain high standards for the whole work place at all times. The followed application is Autonomous Maintenance (AM), which takes place to engage operators in simple, low-criticality machines with simple maintenance tasks, adding lubricant for example in order to ease the load of the maintenance experts and allow them to engage with more complex tasks. VSM will engage to have better visualisation of the system and identify wastes across the system related to inventory for instance and help optimizing the system for waste reduction process. CBA is crucial pillar of the framework as it helps in calibrating and monitoring the system financially to suitable cost analysis without effecting the equipment and machines efficiency. The company understudy has one of a kind equipment that is considered one of the largest manufacturing assets in the modern times. The significance of the mentioned strategies was concluded from the analysis of the extensive Interviews with different employees' levels and the recorded failures and maintenance history. Data analysis highlighted big gap in knowledge regarding the process and the essential maintenance tasks to overtake on time as well as, the shortage in the available staff. Work place environment and standards were considered significant issues that exists across the site.

With the rapid development in the manufacturing sector and the continuous introduction of advanced technology to the sector, it became harder to rely only on one maintenance strategy to cover all related aspects to each organisation. Due to that, a framework combines different maintenance strategies introduced to cover wider prospective, increase quality and efficiency from machines prospective, human factor prospective, and cost prospective, in alignment with waste reduction and work place efficiency improvement using specific pillars of TPM, RCM, VSM and CBA. Future work on the proposed framework will introduce new technology to its different pillars for more efficient results.

References

Bon, A., Lim, M.: Total productive maintenance in automotive industry: issues and effectiveness. In: Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management, Dubai, United Arab Emirates (UAE) (2015)

Baglee, D., Jantunen, E., Sharma, P.: Identifying organisational requirements for the implementation of an advanced maintenance strategy in small to medium enterprises (SME). In: Proceedings of First International Conference on Maintenance Engineering, Manchester, UK (2016)

Black, J., Philips, D.: The lean to green evolution. Ind. Eng. (IE) **42**(6), 46–51. 6p. 2 (2010)

Bookless, C., Sharkey, M.: RCM-streamlined RCM in the nuclear industry-the maintenance optimisation tool (MOT) is the streamlined version of RCM that has been developed by British Energy and used in their power stations to. Maint. Asset Manag. **14**(1), 27–30 (1999)

Chan, F.T.S., Lau, H.C.W., Ip, R.W.L., Chan, H.K., Kong, S.: Implementation of total productive maintenance: a case study. Int. J. Prod. Econ. **95**(1), 71–94 (2005)

Cheng, Z., Jia, X., Gao, P., Wu, S., Wang, J.: A framework for intelligent reliability centered maintenance analysis. Reliab. Eng. Syst. Saf. **93**(6), 806–814 (2008)

Carretero, J., Pérez, J.M., García-Carballeira, F., Calderón, A., Fernández, J., García, J.D., Lozano, A., Cardona, L., Cotaina, N., Prete, P.: Applying RCM in large scale systems: a case study with railway networks. Reliab. Eng. Syst. Saf. **82**(3), 257–273 (2003)

Djukic, M., Jovanoski, I., Ivanovic, O., Lazic, M., Bodroza, D.: Cost-benefit analysis of an infrastructure project and a cost-reflective tariff: a case study for investment in wastewater treatment plant in Serbia. Renew. Sustain. Energy Rev. **59**, 1419–1425 (2016)

Igba, J., Alemzadeh, K., Anyanwu-Ebo, I., Gibbons, P., Friis, J.: A systems approach towards reliability-centred maintenance (RCM) of wind turbines. Procedia Comput. Sci. **16**, 814–823 (2013)

Moubray, J.: Reliability-Centered Maintenance. Industrial Press Inc., New York (1991)

Ng, K.C., Goh, G.G.G., Eze, U.C.: Critical success factors of total productive maintenance implementation: a review. In: 2011 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 269–273. IEEE, December 2011

Patel, N., Chauhan, N., Trivedi, P.: Benefits of value stream mapping as a lean tool implementation manufacturing industries: a review. Int. J. Innov. Res. Sci. Technol. **1**(8), 53–57 (2015). ISSN 2349-6010

Rahman, C.M.: Assessment of total productive maintenance implementation in a semiautomated manufacturing company through downtime and mean downtime analysis. In: 2015 International Conference on Industrial Engineering and Operations Management (IEOM), pp. 1–9. IEEE, March 2015

Sainz, J.A., Sebastián, M.A.: Methodology for the maintenance centered on the reliability on facilities of low accessibility. Procedia Eng. **63**, 852–860 (2013)

Selvik, J., Aven, T.: A framework for reliability and risk centered maintenance. Reliab. Eng. Syst. Saf. **96**(2), 324–331 (2011)

Toivonen, T., Siitonen, J.: Value stream analysis for complex processes and systems. Procedia CIRP **39**, 9–15 (2016)

Vishnu, C.R., Regikumar, V.: Reliability based maintenance strategy selection in process plants: a case study. Procedia Technol. **25**, 1080–1087 (2016)

Design for Intelligent Maintenance: A Potential Reference Standard Complies with Industry 4.0 Requirements

Idriss El-Thalji(✉), Santiago Echeverri Duque, and Helge Nordal

Department of Mechanical and Structural Engineering and Materials Science, University of Stavanger, Stavanger, Norway
idriss.el-thalji@uis.no

Abstract. The cost and profitability related to operation and maintenance phase is significant for several industrial applications. Thus, the first proactive maintenance management action is to design/out the maintenance work by designing out the critical failure modes and causes. The dependability standards (IEC 60300) provides the framework and methodology to design for maintainability at project phase. However, the risk that physical asset will fail is always there as there are changes in the operating and loading conditions, which might initiate new failure modes. The philosophy of industry 4.0 is to develop smart asset to enable a real-time monitoring of the dynamic asset behaviour. In this context, the dependability standards (IEC 60300) need to be updated to consider the technical requirements that support the intelligent maintenance process. Therefore, the purpose of this paper is to present a potential reference standard related to "design for intelligent maintenance" that comply with industry 4.0 requirements. This work illustrates the progress toward a unified standard body for dependability in industry 4.0, which might lead to significant changes in the current state of the art in designing industrial assets. For example, among the 80,000 sensors that are attached to modern oil and gas platforms, a few ones are generating data for health monitoring and maintenance purposes, the majority applied for detecting operational anomalies and control.

Keywords: Industry 4.0 · Intelligent maintenance · Maintenance design · Systems dependability · Maintenance standards

1 Introduction

The cost and profitability related to operation and maintenance phase is significant for several industrial applications (Alsyouf 2007). Thus, the first proactive maintenance management action is to design-out the maintenance work by designing out the critical failure modes and causes. The dependability standards (IEC 60300) provides the framework and methodology to design for maintainability at project phase. However, the risk that physical asset will fail is always existing as there are changes in the operating and loading conditions which might initiate new failure modes. According to a study executed by ARC Advisory Group (Rio 2015) only 18% of asset failures are associated with age, hence, the remaining 82% of failure occurrence is not avoided

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 500–510, 2020.
https://doi.org/10.1007/978-3-030-48021-9_56

through the application of preventative maintenance. In order to avoid such large percent of failures, the failure modes and causes should be sufficiently understood. Based on the 34 projection scenarios of maintenance in digitised manufacturing (Bokrantz et al. 2017) that the philosophy of maintenance in industry 4.0 is to develop smart asset to enable a real-time monitoring of the dynamic asset behaviour (process, health, product and context data). Such maintenance scenario can be considered as smart learning process where the real/time data will help us to learn how to maintain our assets. However, such purposes requires a smart design. For example, despite the fact that an offshore oil rig include around 30,000 sensors monitoring operational parameters (Manyika et al. 2015), none of the big data generated is integrated in an intelligent maintenance system, but only applied for detecting operational anomalies and control (Manyika et al. 2015). Therefore, the standards related to maintenance engineering shall be updated to support such industrial revolution.

The dependability standards (IEC60706-2 2006) Maintainability of equipment-Part 2-Maintainability requirements and studies during the design and development phase describes maintainability and maintenance support requirements and allocation process. This standard has high potential to consider the technical requirements of maintenance in industry 4.0 context. The basic technical requirement is mainly related to develop a cyber-physical asset, where the physical behaviour (process, health, product, and context) will be perceived, transmitted into the cyber space and processed and configured to meet the user expectation (i.e. effective decision making). Answering the questions of what are the requirements of maintenance in industry 4.0 context and do we have a unified standard guiding toward how such requirements are allocated in the design process at the project phase, are main targets of this paper. Therefore, the purpose of this paper is to explore a potential reference standard related to "design for intelligent maintenance" that comply with industry 4.0 requirements. The paper starts with a brief presentation of what are required to develop an intelligent maintenance system, followed by a critical exploration and review of the related existing standards.

2 Design for Maintenance: Any Update?

To answer the question, "does intelligent maintenance has similar architectures, layers and standards as industry/manufacturing 4.0?" let us start with industry 4.0 architecture. The standards, patents (Trappey et al. 2016) (Trappey et al. 2017), technologies and applications (Lu 2017) related to industry 4.0 are reviewed. It is obvious that most of the literatures describe industry 4.0 as an architecture of several layers. The most well-known architecture for industry 4.0 is RAMI 4.0 (VDI/VDE 2015), (Zezulka et al. 2016) which consists of six layers. For manufacturing/production 4.0, the purpose of developing smart asset is to get data related to process and analyse it smartly to support the decision making process. Maintenance has slightly different purpose where the data related to health, process, product and context are smartly collected and analysed to support the maintenance decision making process. Thus, the cyber/physical relationship is similar. However, the type of perception instruments i.e. sensors (numerical measurements) and controller, descriptive text e.g. reports and analytics type are different. The 5C model (connection, conversion, cyber, cognition, configuration) was proposed (Lee et al. 2015) to illustrate the development of intelligent maintenance architecture and layers. For example, watchdog agent platform has been further developed to comply with industry 4.0 technologies (Bagheri et al. 2015).

It is easy to observe the inconsistency among the definition of those layers e.g. conversion layer for some models include the functions of pre/processing the data (big data analytics) and for other models it includes high-level computation e.g. health assessment. Therefore, to avoid such confusing definition, we present the intelligent maintenance as shown in Table 1, based on the process flow and related functionalities. In summary, in the physical space, the asset layer (e.g. physical machines and human workers) will generate data related to maintenance (health and performance measurements, descriptive notifications and reports) which will be acquired (perception layer e.g. sensors, controller) and transmitted (connection layer) into the cyber space.

Table 1. Cyber-physical maintenance

Space	Layers	Main functions
Physical space	Business	Production, operations, maintenance, supply chair, marketing
	Asset and maintenance operations/executions	Platform, Pump, operators, maintenance staff
	info/data perception	Data acquisition e.g. Vibration sensor, reports, notifications
Transmission	Data between physical/cyber spaces	Communications, networks
Cyber space with interface to physical space	Cyber space	Cloud solution
	Conversion (data)	Data manipulation
	Computation (information)	State detection, descriptive analytics
	Cognition (knowledge)	Diagnostic, predictive and prescriptive analytics, Health and prognostic assessment
	Support decision making (optimised solutions)	Maintenance optimisation
	Maintenance management	Maintenance programme planning, capacity planning, spare part planning, scheduling
	Configuration	User interfaces, automatic actions

In the cyber space, several analytical functionalities will be performed: conversion (data per-processing, big data analytics), computations (state detection analytics, performance analytics), cognition (health diagnosis and prognosis assessment analytics, maintenance optimisation and decision support analytics, maintenance management analytics), and configuration (e.g. user interfaces, automatic actions, actuators). The configuration layer aims to transmit (via the transmission layer again) the required actions from cyberspace to physical space. The whole process looks like a closed loop from physical to cyber to physical space again.

3 Design for Intelligent Maintenance Complies with Industry 4.0 Requirements: The Update

Moving toward new standards of design for intelligent maintenance should help us to answer the following questions: How healthy asset I have designed e.g. architecture, machines, interfaces? How smart the existing design is in managing its health over the lifetime? How smart it should be? How smart the maintenance support is and how smart it should be? The new/updated standards shall guide the maintenance process to utilise the potential advantages of having the cyber/physical assets and internet of things (IoT). Cyber physical asset means that you have smart system which can provide data about its behaviour (performance, health, and event), where IoT enables us to acquire/transit/access such data and act remotely. Moreover, it enables the real-time learning process and past-time learning process (historical data) to support the new maintenance perception to be predictive and proactive.

Do we have standards that help us to answer those questions? The exploration of the existing standards related maintenance in industry 4.0 context is summarized in Table 2. Table 2 illustrates several issues related to design for maintainability of smart assets, design for smart maintenance support, which both will be discussed respectively.

Table 2 Existing standards for potential intelligent maintenance standard

Layer title and number	- Existing standards
1. Design for intelligent maintenance during project phase	- Maintainability design (IEC60706-2, 2006) - Criticality analysis: NORSOK STANDARD Z-008 (2001) - Maintenance support: IEC 60300-3-14:2004 - Engineering of system dependability: IEC 60300-3-15:2009 - Maintenance support services: IEC 60300-3-16:2008 - The technology roadmap for industry data dictionary structure, utilization and implementation (IEC TR 61908)
2. Allocation of intelligent maintenance during execution project phase	- IEC 60706-5:2007
3. Maintenance strategic planning and programmes development	- Dependability management IEC 60300-1 (2014), IEC 60300-3-4:2007, Lifecycle costing: IEC 60300-3-3:2017 - Integrated logistic support (ILS) management system IEC 60300-3-12:2011, IEC 60300-3-14:2004

(*continued*)

Table 2 (*continued*)

Layer title and number	- Existing standards
3.1 Reliability centred maintenance programme	- RCM: IEC 60300-3-11:2009, IEC 60300-3-10 (2001-01)
3.2 Condition/based maintenance programme	- (IEC60706-2, 2006), ISO 17359:2018 - MIMOSA/OSA-CBM (MIMOSA, 1998–2018)
3.3 Other maintenance pro-grammes e.g. inspection	- RBI: IEC/ISO 31010:2009, DNV-RP-G101
4. Asset and perception	- There are several standards related to several applications describing the physical system components e.g. NORSOK STANDARD P-100
	- ISO 14224 (2006) Petroleum and natural gas industries - Collection and exchange of reliability and maintenance data for equipment - IEC 60300-3-2 (2004) Dependability management - Part 3: Application guide - Section 2: Collection of dependability data from the field - IEC 60319 (1999) Presentation and specification of reliability data for electronic components - Industrial-process measurement and control - Data structures and elements in process equipment catalogues (NEK EN 61987) - Programmable controllers (NEK IEC 61131) - Automatic identification and data capture techniques (ISO 15459) - Smart transducer interface for sensors and actuators (ISO/IEC/IEEE 21450 and ISO/IEC/IEEE 21451) - Sensor networks (ISO/IEC 20005, ISO/IEC 29182, ISO/IEC 30101, ISO/IEC 30128)
5. Connection (physical to cyber)	- Industrial communication networks (NEK IEC 61918, IEC 61784) - Wireless communication networks (NEK EN 62657, IEC 62591, IEC 62601) - Engineering data exchange format for use in industrial automation systems engineering (IEC 62714) - Security techniques (ISO/IEC 27000) - Network security (ISO/IEC 27033) - Guidelines for the management of IT Security (ISO/IEC 13335)
6. Cyber space and analytics	- Telecommunications and information exchange between systems (ISO/IEC 13157) - Data structure (ISO/IEC 29161) - Standard data element types with associated classification scheme (IEC 61360)

(*continued*)

Table 2 (*continued*)

Layer title and number	- Existing standards
6.1 Conversion	- Integration of industrial data for exchange, access and sharing (ISO/TS 18876) - Exchange of characteristic data (ISO/TS 29002) - Enhanced communications transport protocol (ISO/IEC 14476) - Cloud Data Management Interface (ISO/IEC 17826) - Function blocks (FB) for process control and electronic device (NEK EN IEC 61804)
6.2 computation	- Cloud computing (ISO/IEC 19944)
6.3 cognition	- Statistical methods in process management (ISO 22514) - Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) ISO/IEC AWI 23053
6.4 Decision support	- Industrial automation systems and integration ISO/TS 10303-1486:2011
6.5 Maintenance management	- Asset management: ISO 55001:2014, ISO 55002:2014 - Digital factory framework - IEC TR 62832
6.6 Configuration	- Enterprise-control system integration (NEK EN 62264) - Enterprise integration - Framework for enterprise modelling, ISO 19439:2006 - Batch control (IEC 61512)

It is clear that the most relevant existing standard related to “design for maintenance” is the Maintainability requirements and studies during the design and development phase (IEC60706-2 2006). This standard handles two issues: the design for maintainability (i.e. physical asset e.g. platform) and design for maintenance support (i.e. maintenance system e.g. onshore, offshore support services). The first issue was handled in the standard to guide the developers to answer the following question: How healthy is the asset that I have designed e.g. architecture, machines, and interfaces? Whereas, the guiding question for the second issue is: how smart is the existing design in managing its health over the lifetime? The standard guides the answering of those questions via four main development phases: maintenance concept planning, preliminary design, detailed design and critical design review.

The maintenance concept planning is the first phase where maintenance is seen as a supportive function for the production process. Thus, the maintenance design shall first meet the operational requirements. Therefore, the traditional standard IEC 60706-2 (2006-03-20) starts with determination of operational requirements and environmental conditions. For example, a desired operational scenario for unmanned platform is a cycle of three months of uptime operations and two weeks staffed support.

Second phase aims to estimate, using the reliability and maintainability analyses, how much maintenance is needed for specific physical system e.g. platform, to achieve the desired operational scenario. Then, the maintenance support analysis and allocation aims to define which maintenance system support is required for such designed system to achieve the desired operational scenario. Knowing the desired operating scenarios is the first step, however, the expected operational scenario should be explored to estimate the operational use cases (operating and loading conditions of specific site) beside analysing the historical failure data (if it is available) to estimate 'time-to-failure' of each critical component. Knowing the "time to failure rate" for each critical component might be utilised using maintenance modelling and simulation tools to estimate the overall maintenance event timeline for the whole system e.g. platform. This provides a good indication of how many maintenance events is expected over the lifetime and do such events disturb the desired operational scenario. If the expected maintenance scenario is not acceptable, a second iteration of development shall be taken. In this iteration, the failure modes and causes that generates those critical maintenance events shall be investigated to determine whether they can be designed-out at this stage or detected and monitored during the operational phase. Otherwise, the inspection programme for the undetected failure modes shall be developed. Those traditional three ways to handle the undesired maintenance events will influence the expected overall maintenance event timeline (which was not accepted by designer) and simulation tool can help to estimate those enhancements and see how far/near they are from the desired operational scenario. The existing standard needs to include the maintenance modelling and simulation to guide the developers to estimate the potential maintenance needs for specific physical design under different scenarios (Virtanen 1998), (Hagmark and Virtanen 2009), (Virtanen et al. 2014).

Third phase of the standardised analysis is to define the maintenance support systems e.g. RCM, CBM. The implementation of maintenance programmes is supported by several standards e.g. IEC 60300-3-11:2009, IEC 60300-3-10 (2001-01) for RCM. The projections of maintenance in industry 4.0 indicate that several types of data (process, health, product, and context) will be collected and transmitted into a cyber space where several analytics are applied to extract knowledge help us to take optimal decisions. Thus, all maintenance programmes is expected to get data and analyse them automatically in smart manner that is traceable with the required decisions. However, the volume and sensitivity of process and health measurements are the most significant which make CBM programme in the hotspot. In fact, the cognition layer which involves several different standards related to several monitoring techniques and analytical procedures, has been summarized in Table 3. We tried to explore, in more details, the existing standards for the potential analytics that might be used to detect the present state of the asset behaviour, diagnose and describe the abnormality type, location and severity, and predict the future state.

In the end, the standard IEC60706-2 2006 guides us to generate two documents after performed the standardised procedures. First one is the maintainability definition report where the maintenance needs and physical allocations are determined.

Table 3. Prognostic and health assessment related standards for intelligent maintenance

	General/vibration	Vibration	Lubricant analysis	Acoustic emission	Ultrasound	Thermography	Performance
General guidelines	ISO 13372:2012 ISO 17359:2018 ISO2041:2009						
Data processing	ISO 13374-1:2003 ISO 13374-2:2007	ISO 13373-1:2002 ISO 13373-2:2016 ISO 18431-1:2009 ISO 18431-2:2004 ISO 18431-3:2014 ISO 18431-4:2007	ISO/CD 14830-1 [Under development]	ISO 22096:2007	ISO 29821:2018	ISO 18434-1:2008 ISO/DIS 18434-2 [Under development	ISO 18129:2015
Data communication	ISO 13374-3:2012						
Data presentation	ISO 13374-4:2015						
Data interpretation and diagnostic techniques	ISO 13379-1:2012 ISO 13379-2:2015 Part 2: Data-driven applications	ISO 13373-3:2015 ISO 13373-5:2015 [Under development					
Data interpretation and prognostic techniques	ISO 13381-1:2015						
Qualification and assessment of personnel	ISO 18436-1:2012 ISO 18436-2:2014		ISO 18436-4:2014 ISO 18436-5:2012	ISO 18436-6:2014	ISO 18436-8:2013	ISO 18436-7:2014	
Mechanical systems	ISO 13373-7:2017	ISO 13373-7:2017					
Electrical systems	ISO 20958:2013	ISO 13373-9:2017					
Structural systems							ISO 16587:2004

Second document is the maintenance support definition where all maintenance programmes that will support the required maintenance needs are determined.

4 Discussions and Conclusion

Therefore, it can be concluded that maintenance 4.0 exist in theory and practice. However, it has no unified standard for neither its definition nor procedures. Regards the standardisation, the history and present show us that the German industry is leading the growth toward industry 4.0 vision. For example, the German government created the roadmap for implementing the strategic initiative industry 4.0 (Kagermann et al. 2013) and roadmap for standardisation of industry 4.0 (DIN/DKE 2016).

The recommendations of the fedical goverment encouraged the business associations BITKOM, VDMA and ZVEI to establish the Plattform Industrie 4.0 in 2013. The Plattform Industrie 4.0, today, has a total of over 300 active players from 159 organisations (European_Commission 2017). These associations are working hard with the International Standardisation Organization to create a standards for industry 4.0 (mostly will be based on RAMI 4.0). Regarding the standardisation of intelligent maintenance, the issue is further challenging.

The updated standards for "design for intelligent maintenance" compare to the tradition standards "design for maintenance" shall consider three evolving issues. First, it should adopt the new cyber/physical process of sending and processing data and performing actions. Even thought, the idea of industry 4.0 is to automate and digitalise the working processes, the manual operations shall be potentially replaced by the utilisation of smart sensors and actuators. Thus, maintenance will have more physical inputs (assets) to be allocated into the design architecture as more perception systems (sensors, controllers), connection systems (intra and inter communication) and configuration systems (actuators, maintenance robots e.g. drones) are installed. Thus, maintenance asset architecture becomes one important layer of the total design architecture that will be constructed and installed in the project execution phase. The 'maintenance asset architecture' should be included in the maintenance definition document. This issue might highlight the retrofit solutions as less cost effective solution compare to the early allocation at project development phase.

Second, there are advance analytics (e.g. maintainability modelling and simulation) to perform the traditional maintainability and maintenance support techniques. The maintenance modelling and simulation shall enable the developers to estimate the potential maintenance needs for specific physical design under different scenarios. It is usually challenging for maintenance manager to convince the design team to invest in developing a maintenance programme and to procure the specified items (usually cost reduction actions at procurement process are targeting the supportive asset e.g. maintenance asset). Thus, there is a need to develop simulation models that help the maintenance managers to estimate the potential benefits (e.g. reduction in cost, enhancements in revenue) of such investment.

Third, the standardised analytics are lacking, particularly, in the maintenance cognition layer where the knowledge is extracted from the processed information and utilised in decision support process. Therefore, the standards ISO 13373 (Vibration condition monitoring), ISO 13374 (data processing, communication and presentation), ISO 18431 (signal processing) shall be updated to include advance analytics. It is worthy to highlight that technical requirements are not only the sensors and data transmission instruments. In order to get the benefit of such smart data, the technical requirements related to smart use (analytics and configuration) is highly required. Therefore, the updated standard of design for intelligent maintenance shall consider the technical requirements to perceive smart data and facilitate the smart utilisation e.g. analytics of such data.

The maintenance design process will be more challenging, as we are targeting more smart systems. Thus, we need to be smart enough in designing what shall work in a smart manner. For example, scheduling a maintenance event for a critical component requires an effective prognosis algorithm to predict the future fault severity, which requires effective diagnosis algorithm to determine the current state which by itself relying on how effective fault detection technique is used. The effectiveness of fault detection process is matter of effective acquisition and connection systems. Selecting all those algorithms, techniques and systems requires good understanding of how the fault will be physically initiated and propagated and which symptoms we might be detected over the time. In this way, the maintenance system can be smart as it collects what it needs. Some reports highlighted that in some cases the smart systems are collecting huge amount of data where just 1% is taken into the decision support process (Manyika et al. 2015). This is very acceptable for exploration purposes and their related businesses, but not for monitoring purposes. Designing such smart system means also knowing the specifications of the physical and digital assets, which you are going to procure, as early as possible. However, this is not always the case in the reality.

Even though, the development in intelligent manufacturing is much rapid compare to intelligent maintenance within the context of industry 40, but industry 4.0 is in general lacking a unified standard. Therefore, this work is timely to illustrate the progress toward a unified standard body for dependability in industry 4.0, which might lead to significant changes in the current state of the art in designing industrial assets and support the whole development of industry 4.0.

References

Alsyouf, I.: The role of maintenance in improving companies' productivity and profitability. Int. J. Prod. Econ. **105**, 70–78 (2007)

Bagheri, B., Yang, S., Kao, H.-A., Lee, J.: Cyber-physical systems architecture for self-aware machines in industry 4.0 environment. IFAC-PapersOnLine **48**, 1622–1627 (2015)

Bokrantz, J., Skoogh, A., Berlin, C., Stahre, J.: Maintenance in digitalised manufacturing: delphi-based scenarios for 2030. Int. J. Prod. Econ. **191**, 154–169 (2017)

DIN/DKE: German standardization roadmap industry 4.0. (2016)

European_Commission: Implementation of an industry 4.0 strategy - the German plattform industrie 4.0. (2017). https://ec.europa.eu/digital-single-market/en/blog/implementation-industry-40-strategy-german-plattform-industrie-40

Hagmark, P.-E., Virtanen, S.: Simulation of reliability, availability and maintenance costs. In: Recent Advances in Stochastic Operations Research II (2009)

IEC60706-2: Maintainability of equipment-part 2-maintainability requirements and studies during the design and development phase (2006)

Kagermann, H., Wahlster, W., Helbig, J.: Recommendations for implementing the strategic initiative industrie 4.0: final report of the industrie 4.0 working group. Acatech – National Academy of Science and Engineering (2013)

Lee, J., Bagheri, B., Kao, H.-A.: A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf. Lett. **3**, 18–23 (2015)

Lu, Y.: Industry 4.0: a survey on technologies, applications and open research issues. J. Ind. Inf. Integr. **6**, 1–10 (2017)

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., Aharon, D.: The Internet of Things: Mapping The Value Beyond The Hype Executive Summary. McKinsey & Company (2015)

MIMOSA: Open system architecture for condition-based maintenance (1998–2018). http://www.mimosa.org/mimosa-osa-cbm

Rio, R.: Proactive asset management with IIoT and analytics (2015). https://www.arcweb.com/blog/proactive-asset-management-iiot-analytics

Trappey, A.J.C., Trappey, C.V., Govindarajan, U.H., Sun, J.J., Chuang, A.C.: A review of technology standards and patent portfolios for enabling cyber-physical systems in advanced manufacturing. IEEE Access **4**, 7356–7382 (2016)

Trappey, A.J.C., Trappey, C.V., Govindarajan, U.H., Chuang, A.C., Sun, J.J.: A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0. Adv. Eng. Inform. **33**, 208–229 (2017)

VDI/VDE: Reference architecture model industrie 4.0 (RAMI4.0): Status report. VDI/VDE Society Measurement and Automatic Control (GMA) (2015)

Virtanen, S.: Reliability in product design Specification of dependability requirements. In: IEEE Proceedings Annual Reliability and Maintainability Symposium (1998)

Virtanen, S., Penttinen, J.-P., Kiiski, M., Jokinen, J.: Application of design review to probabilistic risk assessment in a large investment project. In: Probabilistic Safety Assessment and Management, PSAM, Honolulu, Hawaii, 12 June 2014 (2014)

Zezulka, F., Marcon, P., Vesely, I., Sajdl, O.: Industry 4.0 – an introduction in the phenomenon. IFAC (Int. Fed. Autom. Control) **49**(25), 8–12 (2016)

From Prevent to "Predict & Prevent (PnP)": Optimizing Oil and Gas Asset Integrity Decisions

Jawad Raza(✉)

Apply AS, Stavanger, Norway
Jawad.raza@apply.no

Abstract. Asset integrity decisions for (production) critical equipment are mostly qualitative and experience-driven. Preventive maintenance strategies are dominating in oil and gas asset management for many years. Record low oil prices compelled the industry to undergo major organizational and technical transformations. Companies need to find better and effective ways of improving preventive maintenance strategies. One of the main challenges include performing minimum maintenance without compromising safety, availability and reliability requirements. Oil and gas industry is keen in finding innovative solutions to optimize maintenance strategies. As a result, organizations are adapting to intelligent life cycle analytical methods for running asset in optimal and smarter manner. Apply Sørco's Predict and Prevent (PnP) methodology uses equipment (As-is) condition, combined with maintenance history data analytics to precisely predict upcoming maintenance requirements. PnP approach in this paper refer to Apply Sørco's life cycle predictive analytics methodology to retain integrity, reliability and availability of asset. The results of PnP analysis provide decision basis for in-time asset decisions for repairs, inspections, spares, overhauling and equipment modifications. This methodology combines engineering expertise with data analytics. Results from business cases provide useful insights for taking safe, cost-effective and smarter decisions.

Keywords: Life cycle predictive analysis · Historical data · Equipment health assessment · Reliability and availability forecasting

1 Introduction

High critical, high consequent equipment on any offshore platform is usually equipped with sophisticated monitoring and diagnostics (hardware and software) capabilities. High investments are required for establishing such physical and digital infrastructure, making such solutions infeasible for all "High/Medium critical" equipment groups. Offshore oil and gas operations are remote, complex and costly therefore efficient maintenance management is a continuous challenge. Many companies invest in sophisticated technologies to retain highest levels of integrity, in terms of safety, availability and productivity. One of the main challenges is to allocate maintenance budget to the each system/equipment group. The main focus of operation and integrity management team is to minimize the total expenditure and to maximise availability of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 511–520, 2020.
https://doi.org/10.1007/978-3-030-48021-9_57

production resources (Riane, Roux et al. 2009). In addition, equipment age, operation, maintenance and environment have direct or indirect on efficiency of the equipment.

Predict and Prevent (PnP) methodology in this paper refers to Apply Sørco's life cycle predictive analytics approach to retain integrity, reliability and availability of asset. PnP is suitable for high/medium critical equipment with no or less monitoring, or, where monitoring is not feasible option due to high investment cost. The approach aims to extract vital information from historical data combined with status/health of the equipment (or group of equipment). It uses statistical principles to interpret data and forecast upcoming failures and maintenance requirements. Integrity management team can take effective maintenance-related decisions based on such information. This require engineering skills with experience combined with data sciences.

Predictive analytics in a well-known field of mathematics and statistics. It relies on factual quantitative data generated by machines in operation. The data when interpreted into useful information, can lead to smarter and proactive asset decisions. According to (Mobley 2002), predictive maintenance is a philosophy or attitude that uses the actual operating condition of plant equipment and systems to optimize total plant operation. It provides sufficient warning of an impending failure allowing equipment to be maintained when there is objective evidence of impeding failure (Liyanage, Lee et al. 2009). Preventive maintenance strategies are time/calendar-based whereas predictive are condition-based (Scheffer and Girdhar 2004). Increasing awareness on knowledge management for improved performance, with help from latest information & communication technology (ICT), preventive maintenance is being replaced with predictive maintenance (Parida and Kumar 2006).

Life cycle analysis is a methodology used to understand historical failure and repair data, how to obtain such information, and how to turn historical data into probability density function (PDF) and reliability function (Calixto 2016). Several analysis techniques, such as RCM (Reliability Centred Maintenance), FMEA (Failure Modes and Effects Analysis), RBI (Risk-Based Inspections), Failure Tree Analysis (FTA), Safety Integrity Level (Riane, Roux et al.) etc., are widely in use covering both project and operational phase of any asset.

Three basic groups of diagnostics can be *Model-based, Data-driven* and *Hybrid* (combination of model-based and data-driven approaches) (Liyanage, Lee et al. 2009). Predictive analytics can bring large value potentials contributing to continuous improvement in any organization. Offshore oil and gas industry need to rely more on predictive technologies to identify smarter ways for improving maintenance performance. Predict and Prevent (PnP) empowers life cycle predictive analytics of maintenance data, combined with equipment "as-is" condition.

1.1 Cost of Data

Lifeline of all predictive analytics is data. Maintenance management loop introduced by the Norwegian Petroleum Directorate (Oljedirektoratet 1998), emphasizes on the use of data. This data comes from reliability databases, generic libraries, experience, operations and maintenance. It needs to be registered, stored and maintained in large (online/offline) databases. This require human experts, resources (i.e. software, digital infrastructure etc.) and investments. The cost associated with management and usage of such data can be divided into direct or indirect costs. Direct cost include investments required for data storage, hardware integration, data maintenance, hosting etc. Whereas indirect cost include activities related with utilization, interpretation and analysis. These costs may vary from project to operational phase.

As per today, no published literature is available that provide basis for estimated figures for Norwegian oil and gas industry. Based on experiences from new build offshore assets from offshore oil and gas industry, conceptual view of investments cost and value potential is presented below (Fig. 1).

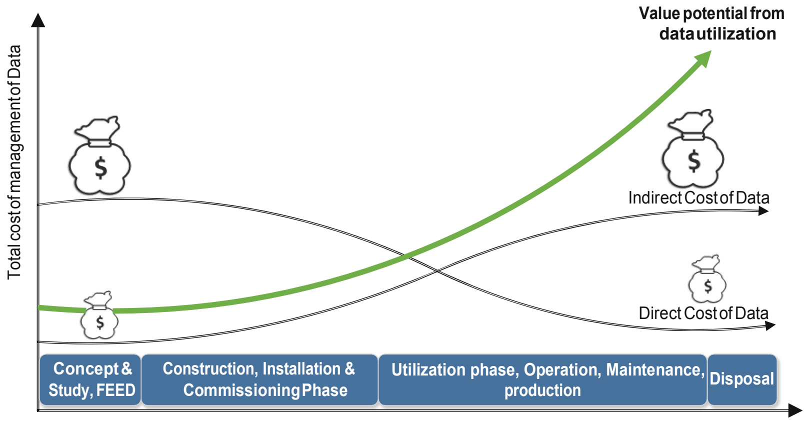


Fig. 1. Cost vs. value in asset life cycle phases

Figure above presents that the investments for infrastructure (direct cost) are high in the project start-up, as this is more or less one-time investment. These costs tend to remain stable (low) when asset is set into operation. Investment at this stage include hardware, software and data acquisition, hosting facilities and digital infrastructure etc. On the other hand, in the operation phase, indirect cost are higher. This include costs associated with usage, interpretation and analyse of collected data.

Simultaneously the potential for extracting valuable information from asset-generated data shows and exponential growth. There is no data available to confirm this trend, however based on experience from offshore assets in operation; this trend is not very far from the reality. Recent focus on digitalization in the oil industry, high digital ambitions, Internet of Things (IoT) and of big-data analytics are the catalyst for this growth.

In oil and gas organizations, reliability engineers are responsible for quality, verification and ownership of the data. Once strategic maintenance plans are developed, data produced by the asset is to be utilized for continuous improvement and life cycle management. It is important that the reporting structure is prepared by the reliability engineer using international reliability standard requirements, refereeing to relevant ISO and NORSOK standards (ISO 2016). For oil and gas assets, criteria and premises for reporting should be defined early in the engineering phase. Failure to do so may kickback in form of re-structuring of the maintenance management system on later stages. This may incur unnecessary cost and challenges for maintenance and reliability department. Data quality, as seen in the industry lacks structure and refinement. This data plays vital role in safety, asset performance and control of the asset.

PnP utilized predictive analytics for smarter maintenance. The level of how organizations benefit from asset data varies from one organization to other. Most organizations have realized the value potential and are using intelligent ways of maintaining their assets. The potential of predictive analytics in not yet fully realized in the Norwegian oil and gas industry.

1.2 Power of Predictive Analytics

As a general rule of thumb, the nature of data determines which modelling method to deploy. A combination of statistical and mathematical techniques are used for PnP lifecycle predictive analytics. In order to fully understand a system's performance data is combined with predictive technologies are analysed to create additional value (Mobley 2002 #1).

Among a long range of possibilities, Weibull method is used in PnP for predictive analysis of maintenance history. It performs well in situations where data is lacking or the quality is questionable (Fig. 2).

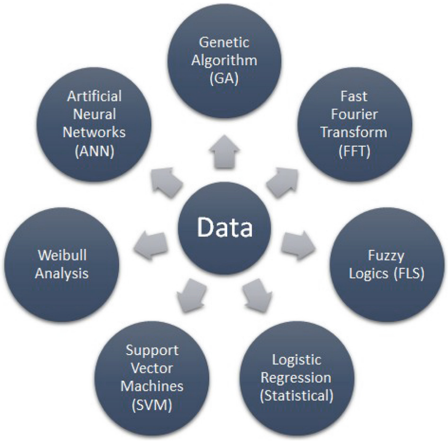


Fig. 2. Data and predictive analytics methodologies choices

Weibull is preferred due to its flexibility and capability to describe many physical modes. It is easy to carry out this analysis since time to failure and preventive replacement details for the failure modes are nearly all that we need (Narayan 2004). This analysis method has many advantages that include graphical solution, suitable for inadequate data, flexibility of working with small samples and accuracy of the results (Abernethy, Breneman et al. 1983). For more complex problems, data-driven techniques such as Artificial Neural Networks (ANN), Fuzzy logics (FLS) and Genetic Algorithm can be more useful.

2 PnP Methodology and Toolbox

Methodology of predict and prevent (PnP) is simple in nature. Basic idea is similar to diagnostics in medical where two main sources of information are:

- Patient's current condition by assessing vital signs (blood pressure, temperature, heartbeat etc.), and
- Patient's medical history

PnP aims to "diagnose" most efficient maintenance strategy based on equipment's current condition and its maintenance history. Figure 3 briefly presents the main processes in PnP methodology that include collecting data, analysing, validating and implementation of results.

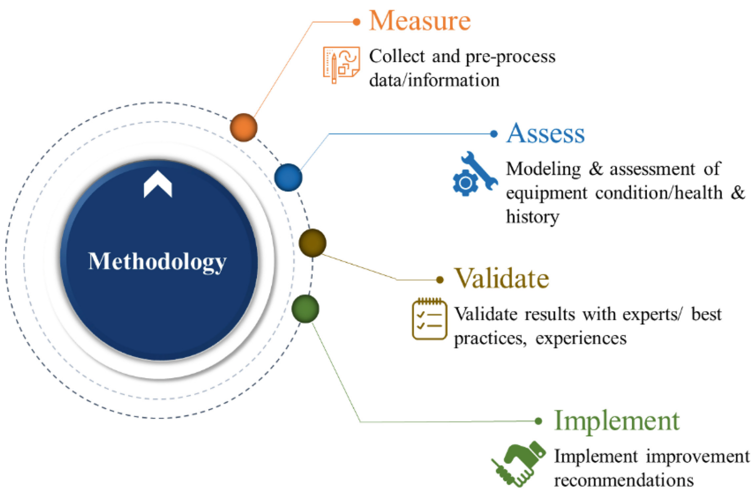


Fig. 3. PnP methodology

In cases where condition data is not available, or monitoring is not feasible, a plug-in Health Assessment & Reliability Toolkit (HART) is used. ADAM (Software) application performs availability, reliability and cost simulations. Output of the inference engine suggest best maintenance strategy. A combination of HART and ADAM is capable of carrying out full PnP analytics.

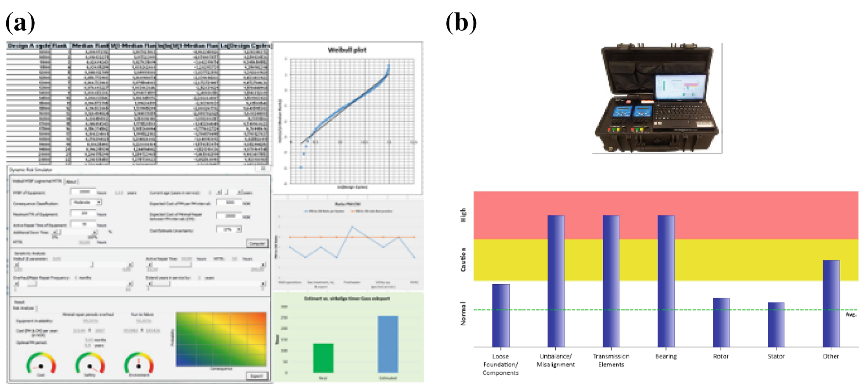


Fig. 4. a. ADAM - the simulation software module. b. HART - as-is health assessment hardware tool

HART interprets voltage & current patterns of the motor using model-based fault detection. Based on the patterns it generates condition assessment report, with warning levels, highlighting major electrical and mechanical, process & energy failures (with accuracy > 90%) (Duyar, Albas et al. 2000). The results provide useful input in energy saving, reduced OPEX, increased productivity and improved process safety (Fig. 4).

ADAM is software application to identify failure distributions, simulate reliability, availability and predict upcoming failures. Cost simulations are performed to identify optimum preventive maintenance intervals.

Results are validated by from domain experts, experiences and reliability engineers. Outcome of the PnP analytics is implemented into maintenance management systems. This provides useful information for future planning of preventive maintenance, corrective maintenance, overhauling, major repairs, modifications, selecting suitable operational strategy and spare parts etc.

Results from PnP lifecycle predictive analytics provide decision support for engineers to optimize their asset's operation and maintenance. It also highlights weaknesses in current maintenance strategies and suggest safer and cost effective solutions.

PnP helps to:

- Identify optimal (safe and cost effective) maintenance strategies
- Predict upcoming failures and repair needs
- Provide input for more informed decisions regarding repair vs. replace
- Predict need for future spares
- Assess equipment start-mid and late-life/End-life assessments with respect to maintenance requirements

3 Description of the Model

Asset data is collected from equipment, systems, and processes. Maintenance data is extracted from maintenance management system whereas condition data is collected from existing condition monitoring systems or via HART. This data is modelled and

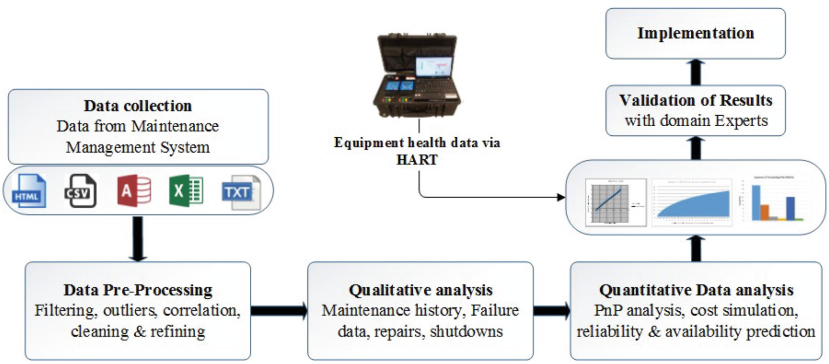


Fig. 5. Generic model PnP analytics

results are used to make asset decision for reliability assessment and logistic support. PnP model is shown in Fig. 5.

Data is collected from maintenance management system in different format and types. This data is pre-processed processed for classification, extracting associations, relationships and trending. Weibull plots from maintenance records are generated to identify probabilities. The two-parameter Weibull equation is simple and is suitable for many applications. Weibull analysis is used for cost simulation, reliability and availability estimations.

4 PnP Life Cycle Analysis Results

PnP analytics helps to understand failures, how these are distributed and their occurrence. Failure modes, causes, down time and repair time are extracted from recorded maintenance history from maintenance management system. Quality of data vary from case to case. In most cases, it is a mix of qualitative and abstract information. Statistical techniques are used to pre-process and clean data. PnP analytics is performed for failures that are observed in the lifetime of the equipment. Other useful information from modifications, previous upgrades, repairs etc. provide additional information used to refine and engineering results.

Below is an example of extracting useful trends from data from an offshore platform. The collected data includes 5 years of operation and maintenance history with 15 failures observed failures.

Starting from qualitative analysis, the failure histogram highlights that most failures occur between 16000–24000 run-hours. Data shows high variance with 95% confidence level that failure occur in range of 1,5–2,1 years. Mean failure time is statistically calculated to be 15477 h. Another important performance indicator is preventive vs. corrective work. In this case, the ratio of preventive to corrective work is 10:1 (Fig. 6).

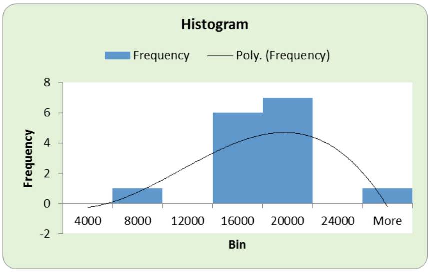


Fig. 6. Histogram of failures from maintenance data

The Norwegian industry practice ratio of preventive to corrective work is 3:1 whereas according to world class standards, this ratio should be 6:1 (Imam, Raza et al. 2013). The ratio simply is an indication of balance between preventive and corrective work. Weibull Probability Density Function (PDF) of the historical data represents proportion of cumulative failures. Weibull PDF and Predicted reliability at time (t) is shown below.

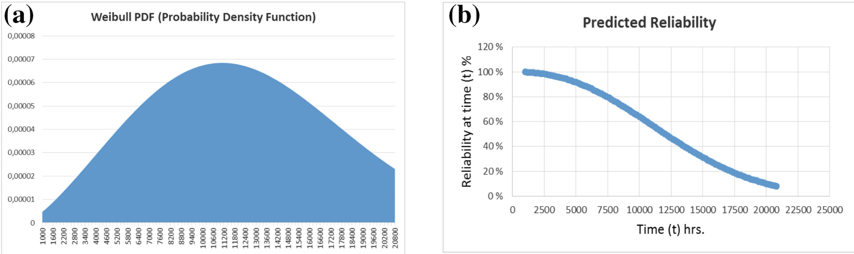


Fig. 7. a Weibull PDF. b. Predicted reliability

PDF (Fig. 7a) of failure data shows peak at 5000 h, cumulative failure probability is 8%. This correspond to chance of only one failure in 5000 h. As a result, predicted reliability of the system @5000 h is 92%. The reliability gradually reduces with time at 10000 h, the reliability it reduced to about 60%. Decrease in reliability means that probability of failure is increasing; it does not mean that equipment will fail at 10000 h. This requires identification of suitable preventive tasks and intervals to keep reliability to higher levels. Weibull plot Shape parameter, Beta (β), determines which member of family of Weibull failure distributions best fit or describes the data. Whereas Weibull characteristic life or scale parameter (α) is percentile of the failure, also denoted as MTTF. For the case β is calculated to be 2,35 whereas α 14058 h. Beta (β) of 1 is regarded as useful life with constant failure rate. In the case under observation, high value of Beta shows that the equipment has passed its useful life period.

These two parameters are used to estimate reliability and availability of the equipment. In order to do so, preventive and corrective maintenance costs, β, lifecycle information, MTTF and MTTR are input for ADAM. The real cost figures were not available for the case; therefore estimated costs are used based on experience.

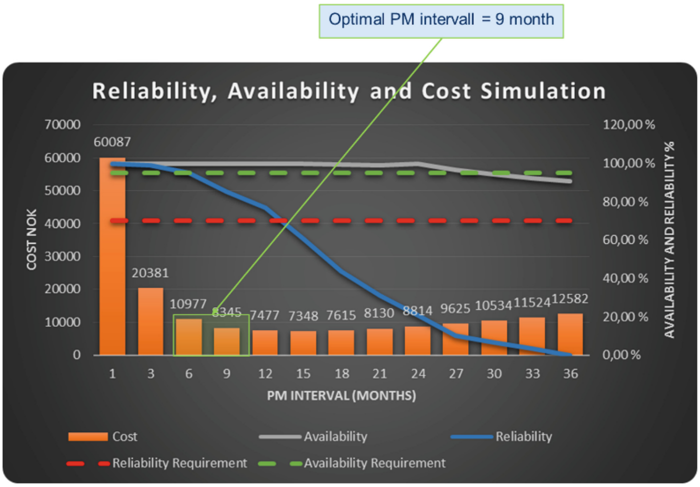


Fig. 8. Graphical representation of PnP results from a case

Corrective cost are assumed to be twice as preventive cost in this case. In other cases, corrective cost can be 5 or even 10 times higher than the preventive cost.

Results from PnP analytics in ADAM suggest the best cost-effective preventive maintenance strategy. Such graphical presentation is easy to explain and shows when the best time for maintenance intervention is (Fig. 8).

In this scenario, for simulation purpose, the upper reliability limit is set to be 95% whereas lowest at 70%. The graph above shows availability of the equipment is not an issue (close to 100%) with very low variation over time. Whereas reliability trend beyond 12 months drops drastically. The best maintenance intervention interval in this case is suggest to be between 6–9 months. These results are validated with team of domain experts. Further is to review the current maintenance strategies and identifying how to optimize the maintenance intervals in a manner without sacrificing safety and risk.

In this case, data from HART was not available. The data from HART provide information about upcoming failures, which are used to suggest the upcoming preventive maintenance tasks.

5 Conclusions

Optimizing maintenance strategies is a living process. To target the real issues and for in-time decisions, sophisticated tools and methodologies are required. Prevent and Predict (PnP) lifecycle Predictive methodology combines equipment health and maintenance history to suggest future preventive and corrective maintenance decisions. In many cases from offshore oil and gas industry, asset maintenance history data is not fully utilized. Outcome of Weibull analysis are reliability, availability and cost predictions. Work on PnP analytics is ongoing and it's full implementation is yet to be explored. The results from the cases so far highlight greater value potential that lies in asset data. When asset data is combined with equipment health (as-is), the upcoming failures can be predicted with high accuracy.

The case presented in this paper, captured a potential of more than 40% savings per annum in preventive maintenance by optimizing existing maintenance strategies. PnP analytics helps in predictive benchmarks that forms basis for smarter maintenance decisions. It also fits well with online remote predictive analytics using data from cloud, which is one of the emerging challenges in the industry.

References

Abernethy, R.B., Breneman, J., Medlin, C., Reinman, G.L.: Weibull analysis handbook. Pratt and Whitney West Palm beach fl Government Products DIV (1983)

Calixto, E.: Gas and Oil Reliability Engineering: Modeling and Analysis. Gulf Professional Publishing (2016)

Duyar, A., Albas, E., Durakbasa, O.T., Serafettinoglu, A.H.: Model-based fault detection system for electric motors. Google Patents (2000)

Imam, S., Raza, J., Ratnayake, R.C.: World class maintenance (WCM): measurable indicators creating opportunities for the Norwegian Oil and Gas industry. In: 2013 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE (2013)

ISO: NS-EN ISO 14224:2016 petroleum, petrochemical and natural gas industries - collection and exchange of reliability and maintenance data for equipment ISO (2016)

Kapur, K.C., Pecht, M.: Reliability Engineering. Wiley, Hoboken (2014)

Liyanage, J.P., Lee, J., Emmanouilidis, C., Ni, J.: Integrated e-maintenance and intelligent maintenance systems. In: Handbook of Maintenance Management and Engineering, pp. 499–544. Springer (2009)

Mobley, R.K.: An Introduction to Predictive Maintenance. Butterworth-Heinemann, Oxford (2002)

Narayan, V.: Effective Maintenance Management: Risk and Reliability Strategies for Optimizing Performance. Industrial Press Inc., New York (2004)

Oljedirektoratet: Basisstudie vedlikeholdsstyring - Metode for egenvurdering av vedlikeholdsstyring, p. 82 (1998)

Parida, A., Kumar, U.: Maintenance performance measurement (MPM): issues and challenges. J. Qual. Maint. Eng. **12**(3), 239–251 (2006)

Riane, F., Roux, O., Basile, O., Dehombreux, P.: Simulation based approaches for maintenance strategies optimization. In: Handbook of Maintenance Management and Engineering, pp. 133–153. Springer (2009)

Scheffer, C., Girdhar, P.: Practical Machinery Vibration Analysis and Predictive Maintenance. Elsevier, Amsterdam (2004)

Life End Stock Need Estimation for Repairable Spare Components of Obsoleting Fleet by Simulation

Tauno Toikka(✉), Jouko Laitinen, and Kari T. Koskinen

Tampere University of Technology, Korkeakoulunkatu 10,
33720 Tampere, Finland
{tauno.toikka,jouko.laitinen,kari.t.koskinen}@tut.fi

Abstract. Fleets of systems are often using common spare components. Occasionally when the manufacturing of some component will end a sufficient size of stock of components is needed in advance to cover the remaining life of the fleet. In this study the computer simulation model was built for making predictions about the fleet availability and spare part stock development concerning the whole life span of the fleet. The computer algorithm included algorithmically the corporation fleet running, repair, transport and storage rules. All event durations have been assumed to be stochastic that is random but evolving from the specific distribution. Failure probability distribution functions for the simulation were generated from the historical data of the components of the systems of the fleet. As a contribution, this study introduces a method of estimating failure probability density function for each failure count individually. This approach makes possible to capture the effect of actual repair process to the probability of the next component failure in simulation.

1 Introduction

Manufacturing of any component will eventually end. Still in some cases it is desired to continue the use of the system, or even the fleet of systems, that uses the component as a spare component. If the post-production is excluded from the options, then one plausible solution is to have a sufficient stock of the specific components as a spare parts before the end of manufacturing. This approach arises a challenge with selecting the stock size since too small stock size will terminate the fleet availability too early and on the other hand too large stock is a wasting of capital.

One efficient way of estimating the stock need in the future is by constructing a simulation of the component usage in future where all events are stochastics. This approach arises a challenge when the spare components are repairable. Current methods of the field do not establish well how to estimate failure probability of the repairable component and further on how to implement the feature of reparability in simulation.

In this study the failure time distributions of the spare components are estimated not only as a function of runtime, that is the general approach, but also as a function of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 521–529, 2020.
https://doi.org/10.1007/978-3-030-48021-9_58

failure count. Since the failure times are clearly the most significant time events in the component life, we are convinced that this approach will improve the simulation accuracy.

The remainder of this paper is organized as follows. Section 2 introduces the framework of parametric failure count and non-parametric failure count and describes the procedure of deriving failure count related failure probability by using failure count and the concept of mean time to next failure. Section 3 introduces how these methods are applied to the data of fleet of aircrafts and the summarizes the results. Section 4 concludes this paper.

2 Framework of Deriving Failure Count Related Failure Probability

2.1 Parametric Failure Count

The failure count is same as the repair count since by default the assumption is that the component will be repaired after the failure until the discard. The estimation of failure count is based on the Nelson-Aalen plot, presented in the Fig. 1 (left). In the figure vertical axis has N(t) that is the failure count as a function of time. Horizontal axis is a component life time. T_1 is the running time of the component recorded before the new component did fail first time, T_2 is the running time of the component starting from the reimplementation of the component after the first repair, till the next failure. The pattern does repeats itself with $T_3, T_4, \ldots T_n$. In practise the Nelson-Aalen plot can be constructed from the historical failure-repair-failure data of the component.

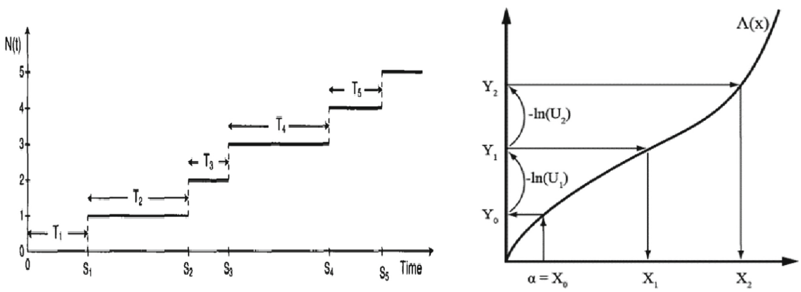


Fig. 1. The principle of Nelson-Aalen plot (left) (Rausand and Hoyland 2004) and continuous model for failure count process (right) (Hagmark and Laitinen 2012)

In this study we denote the analytical form of Nelson-Aalen as $\Lambda(t)$ and use equation for it, suggested by (Rausand and Hoyland 2004):

$$\Lambda(t) = \left(\frac{t}{\alpha}\right)^{\beta} \tag{1}$$

The Eq. 1 is called as a Power Law model.

In order to derive failure count from incomplete real life failure data the non-parametric failure count model is needed. The model is introduced next.

2.2 Non-parametric Failure Count

Non-parametric estimate for Nelson-Aalen can be derived from

$$\Lambda_v = \sum_{r=1}^{v} \frac{S_r}{H_r} \Delta t \tag{2}$$

where

$$S_v = \sum_j \left(\frac{Lap\left(a_j, b_j, t_{v-1}, t_v\right)}{b_j - a_j} k_j + \left(t_{v-1} < b_j < t_v\right) c \right) \tag{3}$$

and

$$H_v = \sum_j Lap(a_j, b_j, t_{v-1}, t_v) \tag{4}$$

where $\mathrm{Lap}(a, b, t, s)$ is the length of the joint interval (a, b) ∩ (t, s) (Hagmark and Laitinen 2012).

In Eq. 3 and Eq. 4 the parameters a, b, k and c are start of the recorded data period, end of the recorded data period, number of events in the period, and event type, respectively. The event type can be either failed (c = 1) or running (c = 0). The interval $]t_{v-1}, t_v]$ is an arbitrary but constant length time interval, that is preferred to be small for high accuracy.

2.3 Mean Time to Next Failure

In order to derive failure count related failure probability we can use the concept of mean time to the next failure (MTNF), that is described next.

Let's consider the continuous failure count function $\Lambda(\mathrm{x})$ so that $\Lambda(0) = 0$, $\lim_{\mathrm{x}\to\infty} \Lambda(x) = \infty$ and $\Lambda(x)$ is strictly increasing. Now let's assume that the single failure count is exponentially distributed with mean 1. Based on the assumption the count step will be $-\ln U_i$, where $U \in [0, 1]$ is the random seed and i is the count. The situation is illustrated in Fig. 1 (right).

In Fig. 1 (right) the counting process $Y_0, Y_1, Y_2\ldots$ (failure counts) is a Homogeneous Poisson Process (HPP) since the counting steps are exponentially distributed. On the other hand counting process $X_0, X_1, X_2\ldots$ (failure times) is a non-homogeneous Poisson process when the $\Lambda(x)$ is non-linear.

From the Fig. 1 it can be seen that the failure time of the *n:th* failure is

$$\begin{aligned} X_n =& \Lambda^{-1}(Y_n) \\ =& \Lambda^{-1}(\Lambda(X_n)) \\ =& \Lambda^{-1}(\Lambda(X_{n-1}) - ln(u)) \end{aligned} \tag{5}$$

and the time between the two subsequent failures is

$$X_n - X_{n-1} = \Lambda^{-1}(\Lambda(X_{n-1}) - ln(u)) - X_{n-1} \tag{6}$$

Thus in more generally the duration to the *n:th* failure time starting from arbitrary time moment $t > 0$ is

$$X_n - t = \Lambda^{-1}(\Lambda(X_{n-1}) - ln(u)) - t \tag{7}$$

from which it can be seen that the mean time to next failure (MTNF) at time moment t is

$$MTNF(t) = \int_0^1 \Lambda^{-1}(\Lambda(t) - ln(u))du - t \tag{8}$$

By applying the parametric Power-Law Eq. 1 to the $MTNF(t)$ of the Eq. 8 we will get

$$MTNF(t) = \int_0^1 \alpha((\frac{t}{\alpha})^\beta - ln(u))^{\frac{1}{\beta}} du - t \tag{9}$$

In order to apply the $MTNF(t)$ in practical simulation we will use discretized version of the function above, that is

$$MTNF(t) = \sum_{u=0..\Delta u}^{1} \alpha((\frac{t}{\alpha})^\beta - ln(u))^{\frac{1}{\beta}} \Delta u - t \tag{10}$$

2.4 Failure Count Related Failure Probability

In this study the cumulative failure probability for the component is estimated by two parametric Weibull function

$$F(t) = 1 - e^{-(\lambda t)^\gamma} \tag{11}$$

where γ is called scale parameter and δ is called shape parameter.

MTNF for Weibull is

$$MTTF = \frac{1}{\lambda}\Gamma\left(\frac{1}{\gamma}+1\right) \tag{12}$$

where Γ is the Gamma function (Rausand and Hoyland 2004).

Practical field data describing failure-repair history can be implemented with the Eq. 2, further on from which the parametric $\Lambda(t)$ of the Eq. 1 can be determined by curve fitting. Additionally Mean Time to Next Failure $MTNF(t)$ can be determined by the Eq. 9 and the Eq. 10. Since now we have runtime dependent measures $\Lambda(t)$ and $MTNF(t)$ we can trivially combine them to $MTNF(\Lambda)$. Further on from the Eq. 12, Eq. 11 we can determine $\lambda(\Lambda)$ and $\gamma(\Lambda)$ for the Eq. 11 since $MTNF = \int_0^\infty F(t)dt$ must be always satisfied and thus having $F(\Lambda, t)$. The function $F(\Lambda, t)$ is called here as failure count related cumulative failure probability.

Now the failure count related cumulative failure probability $F(\Lambda, t)$ can be used to simulate failure times of the components in fleet. From these times, and a set of additional delay times simulated by conventional manners, we can determine the fleet availability and further on the stock need. This is done next.

3 Availability and Stock Need Simulation Using Aircraft Fleet Data

The method of the failure count related cumulative failure probability have been established as a part of larger spare component stock need simulation. In the simulation the aim was to determine the availability of the fleet and the stock size development for further analysis of the domain experts.

In the simulation all the events were stochastics, that is random but from specific distribution. The simulated time events were: component failure recognition, component removal, component transport from operation field to repair shop stock, component transport to repair shop stock to repair, component maintenance, component repair, component transfer from repair to repair shop stock, component transfer from repair shop stock to operation field, component installation to system, and component runtime before failure. Beside these, the simulation also included algorithmically the rules concerning the management of the components: when discarded, number of simultaneously maintainable components, number of simultaneously transportable components, annual fleet operation time requirement, number of components in system, number of systems in fleet, maintenance interval, maintenance margin, fleet size and so on. Out of the all of the time events, the runtimes before the failure were far the most significant, and thus having the special attention here.

The simulation model was tested by the field data from past of the fleet of total 63 aircrafts using a total stock of similar 73 components. The components had a history of total 23 failures. The components were repairable and each component has a record of its own individual failure-repair-failure history. The most failed component has a history of 3 failures and repairs

The failure times of all the components are plotted in the Fig. 2 by red dots. For illustrative reasons the censored times (the last time a component has been stated as a running) are also plotted in the same plot by blue dots. If the component has failed once but not twice the count value of the censored time would be middle between of 1 and 2 that is 1.5. By the same logic the twice failed but currently running components will have count value of 2.5 and so on. The never failed currently running components have the count value of 0.5.

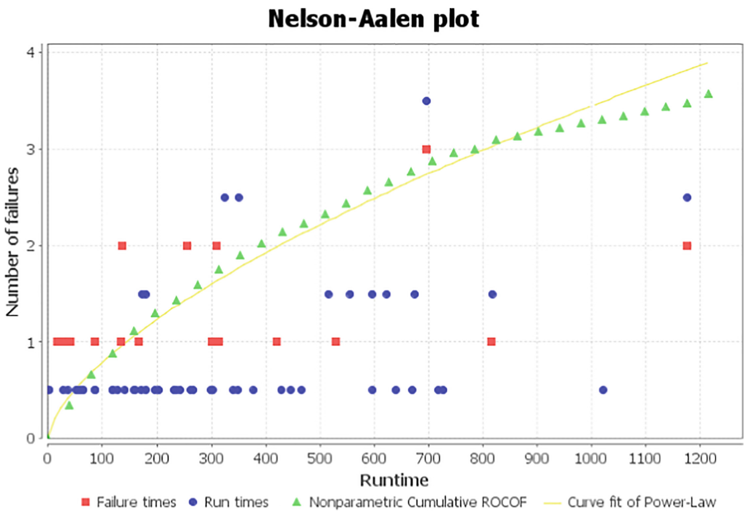


Fig. 2. Nelson-Aalen plot for the components of the study. Red dots are failure times, blue dots censored times, green dots are non-parametric estimate of Λ and the parametric estimate is yellow solid line.

In the Fig. 2 the non-parametric estimate of Λ is by presented green dots generated by the Eq. 2. The parameters for the parametric estimate have been found by fitting the non-parametric estimate on the Eq. 1. With the test data, the resulting parameters are $\alpha = 144$ and $\beta = 0.64$.

Since $\beta < 1$ the component is called "happy" meaning that it improves by every count (Rausand and Hoyland 2004). The happy trend can be due to the improving effect of repair procedure or because the component is still in "running in" stage in the time span of our data. If the reason is "running in" period, the final results should be taken with high scepticism since the run in period will end in some point in the future. If the reason is improving effect of the repair, then the repair has been improved the component this far but it may not be reasonable to assume that the repair process can improve the component forever. If in general the component is happy then the results should be treated with caution. On the other hand if the component happens to be a sad component, that is when $\beta > 1$, that ensures that the component operates at the wear out period and the results will be more conservative.

The cumulative failure probability $F(t)$ of the Eq. 11 for the whole component population have been derived from the data by using Kaplan-Meier estimate described in (Kaplan and Meier 1958), that includes not only a failure times but also a censored times. The failure count related cumulative failure probabilities $F(\Lambda, t)$ have been derived by the methods described in the Methods section and the result is presented in the Fig. 3.

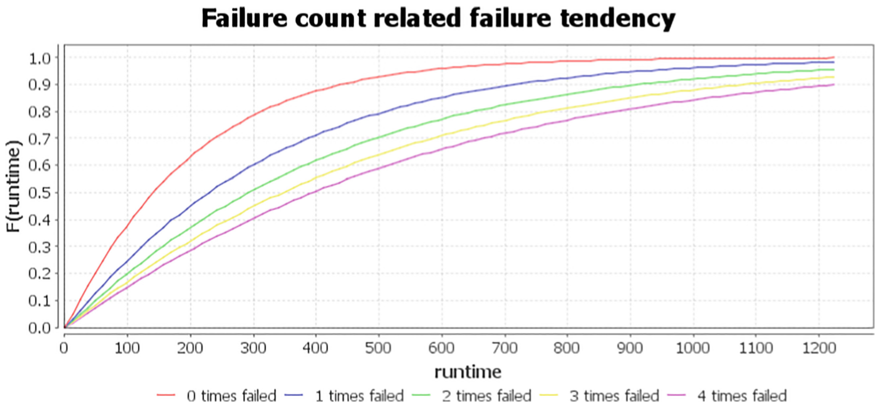


Fig. 3. Failure count related cumulative failure probabilities.

Since $\gamma(\Lambda)$ and $\lambda(\Lambda)$ of $F(\Lambda, t)$ described in Sect. 2.5 are not independent, they both cannot be derived from single $MTNF(\Lambda)$. The problem is solved here by having either shape or scale parameter fixed to the value derived from the entire data. In order to make decision between the two, one must understand the effect of the parameters to the distribution. Shape parameter γ effects to the trend of the failure; for example, when $\gamma = 1$ then the failures are exponentially distributed and failure probability is not the component runtime related. When $\gamma < 1$ then component is improving over the runtime and if $\gamma > 1$ then component reliability is declaring over the runtime. When one recognises that the repair process is capable of changing trend then locking λ may be justified. Otherwise γ should be locked and λ adjusted since scale parameter λ just scales the distribution.

As a result, the simulation did produce fleet availability and stock size development as a function of calendar time, for further analysis for fleet management. The example of the results is presented in Fig. 4. The results of the figure are the averaged values of large number independent fleet lifespan simulations, which guarantees the smoothness of the curves.

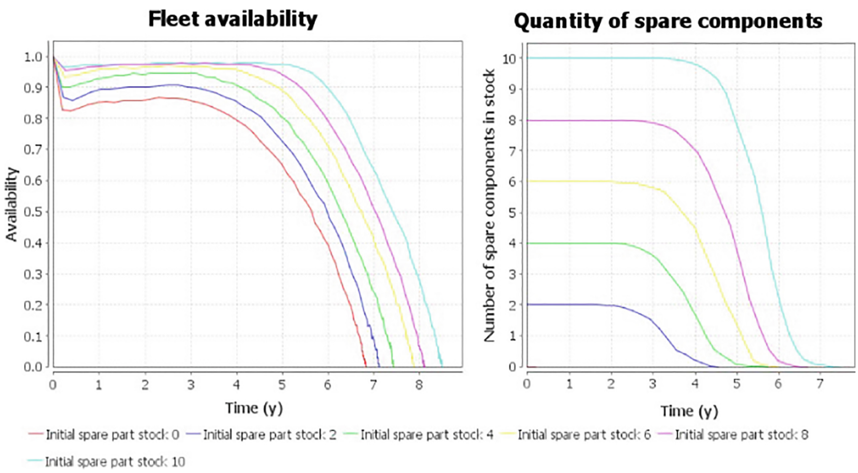


Fig. 4. Availability of the fleet with different amount of initial spare components stock sizes (left) and stock size involvement (right) as a function of calendar time.

4 Conclusions

In this study the problem of estimating the future stock need of spare components of obsolescing feet was solved by computer simulation. The simulation included failure times, other delay times, and the practice of managing the components in the organisation managing the fleet.

Failure times are the most significant events in terms of timespan and thus having the greatest effect on the simulation result. For this reason, the underlying failure probabilities for failure times are estimated here from the real failure history data of the component. The novelty in our approach is that the failure probability distributions are not just a function of times $F(t)$, that is the traditional approach, but also as a function of failure count $F(\Lambda, t)$, thus including the effect of the actual repair process related to the component and in this way mimicking more realistically the actual nature of component failure. The method captures improving or degenerative effect of repair process, which is present more or less in all practical repair processes.

The actual simulation result of this study is not directly applicable in general. On the other hand, the method presented here can be exploited in situations where one desires to include the actual effect of repair process on failure probability distribution function and in this way to achieve more specific results. These probability density functions can be further used for direct parametric analysis or they can be used further for simulation. The limiting factor of the approach is that during the data collection besides the failure time also the number of the failure count need to be recorded. The methods of this study were applied here on one specific fleet and on one specific type of component. In order to further validate the applicability of the methods more study is needed with different data from several different configuration setups.

References

Hagmark, P.-E., Laitinen, J.: Non-parametric bootstrap confidence limits for age-dependent failure tendency using incomplete data. In: Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012), vol. VII, pp. 277–286 (2012)

Kaplan, E., Meier, P.: Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. **53**(282), 457–481 (1958)

Rausand, M., Hoyland, A.: System Reliability Theory: Models, Statistical Methods, and Applications. Wiley, Hoboken (2004)

Estimating MTTF of a Component Based on Spare Parts Consumption Data

R. Jiang(✉)

Faculty of Automotive and Mechanical Engineering,
Changsha University of Science and Technology, Changsha, China
jiang@csust.edu.cn

Abstract. In this paper we consider the situation where the OEM of a repairable system wants to estimate the mean time to failure (MTTF) of a given component of the system and the component's lifetime distribution so as to forecast the demand of spare parts and to optimize the maintenance policy of the component. The OEM does not have exact field failure time data of the component but has the installed base information of the systems and the spare parts consumption information of the component. Due to lack of failure time data, the failure-time-based approach is no longer applicable. To overcome this difficulty, a novel approach is proposed to estimate MTTF and lifetime distribution of the component based on its spare parts consumption data. The proposed approach is based on the assumption of Weibull renewal process and a modification of the asymptotic renewal function. It provides a much more accurate estimate of MTTF than the one obtained from the exponential distribution assumption. A numerical example is included to illustrate the appropriateness and usefulness of the proposed approach.

1 Introduction

Generally, it is difficult for the original equipment manufacturer (OEM) of an expensive product (e.g., wind turbine) to collect the exact field failure time data of its products. However, the OEM may be the unique provider of some spare parts of the product, and hence has the consumption information of those spare parts. This situation is generally true for new products sold under warranty. The OEM hopes to estimate the mean time to failure (MTTF) of a certain component of the system and the component's lifetime distribution so as to forecast the demand of spare parts and to optimize the preventive maintenance policy of the component.

Due to lack of failure time data, the failure-time-based approach is no longer applicable and a novel approach is needed. This paper aims to develop such an approach based on the installed base information of the product and spare parts consumption information of the component.

The traditional method to estimate MTTF is based on the exponential distribution assumption and requires a long observation history. Sheikh et al. (2000) and Krasich

Supported by the National Natural Science Foundation of China (No. 71771029).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 530–539, 2020.
https://doi.org/10.1007/978-3-030-48021-9_59

(2009) discuss the uses of MTTF and MTBF in a variety of contexts, including logistics for planning of the spare parts. Ghodrati (2011) classifies the spare parts demand forecasting models into five categories, including renewal process model and Bayesian approach. The renewal process model can offer an accurate prediction of the number of required spare parts but there may not be an adequate amount of recorded data for analysis. The Bayesian approach is applicable for the situation where a complete set of data to determine MTTF is not available. This is especially true when a new system has not operated long enough. The main drawback of Bayesian approach is that it requires specifying a prior distribution of the unknown distribution parameters and this is usually not easy. Dekker et al. (2013) deal with the reliability-based spare part demand forecasting approach, where the installed base information of the product is used.

In this paper, a novel approach is proposed to accurately estimate MTTF of a component based on its spare parts consumption information and the assumption of Weibull renewal process. The proposed approach first estimates the empirical renewal function (RF), which uses the installed base information. Then, an approximation that relates MTTF to the empirical RF is developed. The approximation is obtained through modifying the well-known asymptotic RF. Another approach is also proposed to estimate the Weibull shape parameter. A numerical example is included to illustrate the appropriateness and usefulness of the proposed approaches.

2 Proposed Approach

2.1 Empirical Renewal Function

Consider a fleet of n identical systems with the same age, and each system contains m copies of a certain component. The component is non-repairable and a run-to-failure policy is implemented. That is, the failure process of the component is a renewal process. The place of run-to-failure policy will be taken by a preventive replacement policy once the component's lifetime distribution becomes known.

Suppose that these systems totally consume k spare parts in time interval (0, t). The empirical RF can be estimated as

$$x(t) = \frac{k}{nm}. \tag{1}$$

If the component lifetime follows the exponential distribution, the estimate of MTTF is given by

$$\mu_{\exp} = \frac{nmt}{k} = \frac{t}{x(t)}. \tag{2}$$

If the component lifetime does not follow the exponential distribution and t is small, Eq. (2) will give a considerable overestimate of MTTF. To illustrate this point, we consider the Weibull distribution with shape parameter β and scale parameter η. Let $M(t)$ denote RF at t. For small t, we have

$$M(t) \approx F(t) \approx (t/\eta)^{\beta},\ M(t) < (t/\eta)^{\beta}. \tag{3}$$

For $\beta > 1$ and $t/\eta < 1$, using Eq. (3) to Eq. (2) we have

$$\mu_{\exp} = \frac{t}{M(t)} \approx \frac{t}{(t/\eta)^{\beta}} = \frac{\eta}{(t/\eta)^{\beta-1}} \gg \eta. \tag{4}$$

As a result, it is desired to develop a novel approach that can accurately estimate MTTF based on the observation, (t, $x(t)$), even when t is not large.

2.2 A Modification to Asymptotic RF

The well-known asymptotic RF for large t is given by

$$M(t) = t/\mu - M_0 \tag{5}$$

where μ is MTTF, $M_0 = 0.5(1 - \rho^2)$ and ρ is the coefficient of variation (CV) of the component lifetime. To obtain a good estimate of MTTF for small t we modify Eq. (5) as

$$M(t) \approx t/\mu - a(t) \tag{6}$$

where $a(t)$ meets the following relations:

$$a(t) < t/\mu,\ \lim_{t \to \infty} a(t) \to M_0. \tag{7}$$

From Eq. (6), MTTF can be estimated as

$$\mu_e = t/[x + a(x)]. \tag{8}$$

The problem is to specify the expression of $a(x)$ for a given distribution. In this paper, we confine our attention to the Weibull distribution. The starting point is to calculate the exact values of $a(x)$ for a given set of β values. This needs to calculate the exact values of the Weibull RF using a numerical integration method. Table 1 shows the values of the Weibull RF for β = 1.5(0.5)3.5. From Eq. (8), the exact values of $a(x)$ can be evaluated as

$$A(x) = t/\mu - x \tag{9}$$

where x is the exact RF.

Table 1. Renewal function of the Weibull distribution

$t/\eta\backslash\beta$	1.5	2	2.5	3	3.5
0.05	0.011118	0.002497	0.000559	0.000125	2.8E-05
0.1	0.031312	0.009963	0.003158	0.001	0.000316
0.2	0.087532	0.039457	0.017758	0.007971	0.003572
0.3	0.158449	0.087342	0.048318	0.026675	0.014686
0.4	0.239843	0.151838	0.097154	0.062197	0.039713
0.5	0.329015	0.230694	0.164765	0.118263	0.084801
0.6	0.424024	0.321384	0.249966	0.196487	0.154795
0.7	0.523408	0.421317	0.35013	0.29581	0.251571
0.8	0.626048	0.528023	0.461623	0.412404	0.372613
0.9	0.731078	0.639301	0.580371	0.540241	0.510601
1	0.837835	0.753323	0.702479	0.672329	0.654705
1.1	0.945811	0.868679	0.824758	0.802291	0.793629
1.2	1.054625	0.984374	0.945065	0.92581	0.919346
1.3	1.163995	1.099786	1.062393	1.041403	1.029729
1.4	1.273712	1.2146	1.176729	1.150252	1.128583
1.5	1.383629	1.328724	1.288743	1.255235	1.22307
1.6	1.493645	1.442216	1.399416	1.359591	1.32012
1.7	1.603688	1.555211	1.509707	1.465797	1.423811
1.8	1.713715	1.66787	1.620332	1.575003	1.534708
1.9	1.823699	1.780342	1.731675	1.687079	1.650799
2	1.933623	1.892746	1.843804	1.801075	1.769013
2.1	2.043482	2.005162	1.95657	1.915779	1.88652
2.2	2.153273	2.117635	2.069722	2.030171	2.001519
2.3	2.262998	2.23018	2.183005	2.143659	2.113478
2.4	2.37266	2.342794	2.296229	2.256102	2.22295
2.5	2.48226	2.45546	2.40929	2.3677	2.33112

Figure 1 displays the plot of $A(x)$. As seen, $A(x)$ quickly increases from zero for small x and tends to a constant for large x. Therefore, as an approximation of $A(x)$, $a(x)$ should have the following features: $a(0) = 0$ and $\lim_{t\to\infty} a(x) = M_0$. The following function meets these requirements:

$$a(x) = \frac{1}{\lambda\Gamma(1+1/\beta)}\left(1 - e^{-\lambda x^{1/\beta}}\right). \tag{10}$$

Noting the second relation of Eq. (7), for large x, we have

$$a(x) \approx \frac{1}{\lambda\Gamma(1+1/\beta)} = M_0. \tag{11}$$

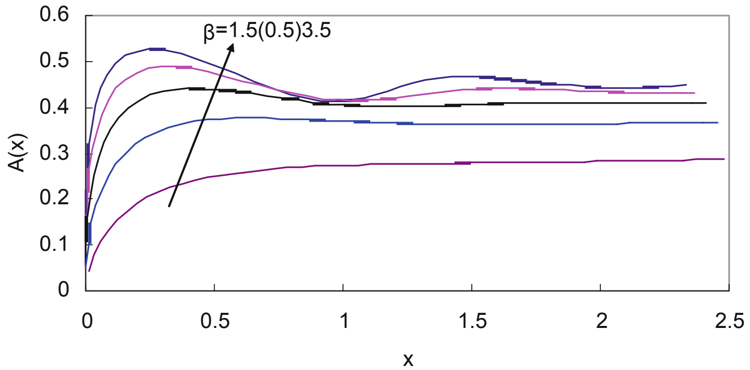


Fig. 1. Plot of A(x)

This yields

$$\lambda = \frac{1}{M_0 \Gamma(1 + 1/\beta)}. \tag{12}$$

As such, Eq. (8) is fully specified. Noting $x \approx (t/\eta)^\beta$ and the second relation of Eq. (3), for small x, applying the Tailor's series expansion to Eq. (10) yields

$$a(x) \approx x^{1/\beta}/\Gamma(1 + 1/\beta) \approx t/\mu,\ a(x) < t/\mu. \tag{13}$$

This implies that the first relation of Eq. (7) can be met and $a(x) \approx A(x)$ for small x and hence Eq. (8) can provide a good estimate of MTTF for small x.

2.3 Accuracy of Equation (8)

The relative error is defined as

$$\varepsilon = \mu_e/\mu - 1. \tag{14}$$

Table 2 shows the relative errors of MTTF estimates in the range of $t/\eta \leq 1$ obtained under the exponential distribution assumption. As seen, the relative errors can be two to three orders of magnitude when t is small.

Table 2. Relative errors of MTTF estimates obtained from t/M(t)

$t/\eta\backslash\beta$	1.5	2	2.5	3	3.5
0.05	3.98167	21.1823	98.1063	442.121	1980.6
0.1	2.53768	10.1189	34.0763	109.823	349.348
0.2	1.53102	4.61486	11.4758	26.7931	61.029
0.3	1.09733	2.8048	5.87769	11.4581	21.629
0.4	0.84743	1.91819	3.56073	6.12404	10.1575
0.5	0.68341	1.40087	2.36155	3.68336	5.53135
0.6	0.56746	1.06805	1.65892	2.38261	3.29368
0.7	0.48147	0.84045	1.21464	1.62132	2.08229
0.8	0.41552	0.67831	0.91972	1.14883	1.3783
0.9	0.36368	0.55945	0.7178	0.8454	0.95252
1	0.32214	0.47046	0.57689	0.6476	0.69196

Figure 2 shows the plots of relative errors of MTTF estimates obtained from Eq. (8). As seen, the maximum relative error is smaller than 40%. This implies that the MTTF estimate obtained from the proposed approach is much more accurate than the estimates obtained under the exponential distribution assumption. However, Eq. (8) overestimates MTTF to some extent when RF is much smaller than one and β is large. An improvement on Eq. (8) is proposed as follows.

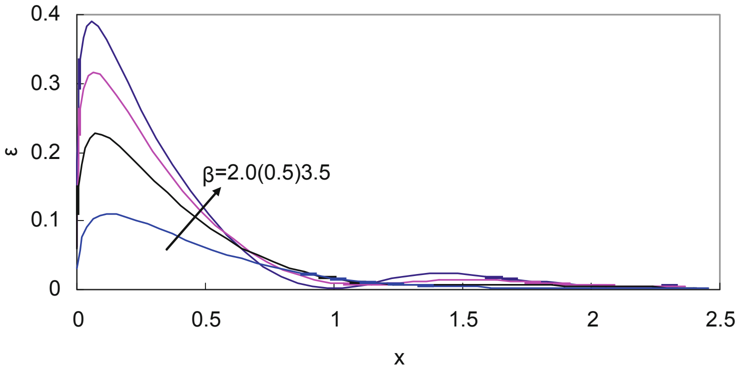


Fig. 2. Plots of relative errors for β = 2.0(0.5)3.5

2.4 Improvement on Equation (8)

According to Fig. 2, the relative errors can be fitted to the following model

$$\varepsilon = \alpha g(x; u, v) \tag{15}$$

where α is a function of β and $g(x; u, v)$ is the gamma pdf with shape parameter u and scale parameter v, which are also functions of β. For the RF data shown in Table 1 and associated with $t/\eta \leq 1$, using the least square method we obtained the estimates of the parameters shown in Table 3.

Table 3. Parameters of Eq. (15)

β	α	u	v
2	0.0690	1.401	0.3268
2.5	0.1143	1.304	0.2873
3	0.1414	1.278	0.2610
3.5	0.1560	1.257	0.2451

After a careful analysis for the data shown in Table 3, it is found that the data of α, v and $z = 1/(vu^p)$ can be well represented by the following model

$$y = b_0 + b_1/\beta. \tag{16}$$

Table 4 shows the regression coefficients. As such, for a given β, we have

$$\alpha = 0.2820 - 0.4238/\beta,\ v = 0.1340 + 0.3844/\beta,$$
$$z = 4.3874 - 4.9139/\beta,\ u = 1/(vz)^{0.7331}. \tag{17}$$

Table 4. Parameters of Eq. (16)

y	b_0	b_1	p
α	0.2820	−0.4238	
v	0.1340	0.3844	
z	4.3874	−4.9139	1.3640

From Eqs. (8) and (14), the modified estimate of MTTF is given by

$$MTTF = \mu_e/(1+\varepsilon) = \frac{t}{[x + a(x)](1+\varepsilon)} \tag{18}$$

where $a(x)$ is given by Eqs. (10) and (12), and ε is given by Eq. (15).

Table 5 shows the maximum absolute errors (MAE) and maximum relative errors (MRE) of Eq. (18). As seen, MRE in the range of $t/\eta \leq 2.5$ is smaller than 2.1% and MAE in the range of $t/\eta \leq 1$ is smaller than 1.1%. The proposed approach can be easily implemented using an Excel spreadsheet program.

Table 5. Maximum errors after modification

t/η	β	2	2.5	3	3.5
(0.05, 2.5)	MAE	0.0023	0.0051	0.0106	0.0188
	MRE	0.0026	0.0058	0.0118	0.0209
(0.05, 1.0)	MAE	0.0023	0.0051	0.0094	0.0087
	MRE	0.0026	0.0058	0.0105	0.0097

3 Illustration

The proposed approach can be applied in several different cases, including

(a) There is only one observation of the empirical RF for a fleet of nominally identical systems,
(b) There are several observations of the empirical RF at different time points for a fleet, and
(c) There are several fleets and each is only one observation of the empirical RF.

For case (a), we need a prior value of β and MTTF can be estimated using the proposed approach. For case (b), for a given value of β, we can estimate several MTTFs using the proposed approach and the value of β can be estimated by minimizing the CV of MTTFs. For case (c), we propose the following approach to estimate the value of β.

For a set of observations of $\{t, mn, k\}$, we can obtain a set of estimates of $\mu_{\exp}$ from Eq. (2) and a set of estimates of μ_e from Eq. (18). Let δ denote the correlation coefficient between $\{\mu_{\exp}\}$ and $\{\mu_e\}$, which is a function of β. When β is too small, δ is close to 1, otherwise, δ is too small, even negative. As such, the curve of δ versus β has an inflection point, where β is neither too small nor too large. We take the value of β that corresponds to the inflection point as the estimate of β. To find the inflection point, we use the least square method to fit the data set (β, δ) to the following relation

$$\delta = (1 + a)W(\beta - 1, b, c) - a. \quad (19)$$

The estimate of β is given by

$$\beta = 1 + c(1 - 1/b)^{1/b}. \quad (20)$$

The scale parameter associated with each observation can be estimated as

$$\eta = \mu_e / \Gamma(1 + 1/\beta). \tag{21}$$

To illustrate, we consider a certain component of four fleets of systems, which operate in different environments. Assume that the shape parameter is $\beta = 3$ and the scale parameters are different and shown in the second column of Table 6. The fleet sizes and observation intervals are shown in the 3rd and 4th columns. Simulation is used to generate the values of k, which are shown in the 5th column.

For $\beta = 1.25(0.25)4.00$, we find the values of δ, which are shown in Table 7. The inflection point of curve of δ versus β is at $\beta = 2.926$, which is very close to its true value. The last two columns show the relative errors (RE) of the MTTF estimates obtained under the exponential distribution assumption and the proposed approach, respectively. As seen, the exponential assumption leads to huge errors and the proposed approach gives accurate estimates of MTTF. In the meantime, the Weibull distribution of each fleet is obtained with the scale parameter being given by Eq. (21).

Table 6. Observed RF and estimated MTTF

Fleet	η, days	mn	t, days	k	μ, days	μ_{exp}, days	μ_e, days	RE of μ_{exp}, %	RE of μ_e, %
1	1100	100	720	26	982	2769	1003	181.9	2.1
2	1300	100	920	31	1161	2968	1194	166.6	2.8
3	1175	150	335	4	1049	12563	1105	1097.3	5.3
4	990	200	455	20	884	4550	918	414.7	3.8

Table 7. Observed RF and estimated MTTF

β	1.25	1.5	1.75	2	2.25	2.5
δ	0.9976	0.9948	0.9865	0.9601	0.8854	0.7072
β	2.75	3	3.25	3.5	3.75	4
δ	0.4016	0.0719	−0.1750	−0.3341	−0.4354	−0.5021

4 Conclusions

In this paper, we have proposed a novel approach to estimate MTTF of a component based on the information of spare part consumption. Theoretic analysis and simulation have illustrated that the proposed approach can provide an accurate estimate of MTTF. Different from the classic approach to estimate MTTF, the proposed approach does not need the information of time to failure, and hence it can considerably simplify the data preparation process for reliability analysis of the components of a complex system. The MTTF estimate can be dynamically updated as more spare part consumption information is obtained.

A main finding of the paper is that MTTF obtained under the exponential distribution assumption is a considerable overestimate when the observed time interval is short. As such, one must be cautious in using this approach to estimate MTTF since mechanical parts do not conform to the exponential distribution (Sheikh et al. 2000).

Though the run-to-failure policy is assumed, a preventive replacement policy can be implemented once the component's lifetime distribution is obtained.

A topic for future research is to extend the proposed approach to other life distributions such as the lognormal distribution and gamma distribution.

References

Dekker, R., Pinçe, Ç., Zuidwijk, R., Jalil, M.N.: On the use of installed base information for spare parts logistics: a review of ideas and industry practice. Int. J. Prod. Econ. **143**(2), 536–545 (2013)

Ghodrati, B.: Efficient product support – optimum and realistic spare parts forecasting, pp. 225–269. Springer, London (2011)

Krasich, M.: How to estimate and use MTTF/MTBF would the real MTBF please stand up? Reliability and Maintainability Symposium, pp. 353–359 (2009)

Sheikh, A.K., Younas, M., Raouf, A.: Reliability based spare parts forecasting and procurement strategies. In: Maintenance, Modeling and Optimization, pp. 81–110. Springer (2000)

Clarifications of Dangerous Detected Failures of Low Demand Safety Instrumented Systems in the Oil & Gas Industry

Pengyu Zhu(✉) and Jayantha P. Liyanage

University of Stavanger, Stavanger, Norway
{pengyu.zhu, j.p.liyanage}@uis.no

Abstract. Low demand safety instrumented systems are widely used to help prevent and mitigate major accidents in the oil & gas (O&G) industry. The paper studies a special type of failures of such systems called dangerous detected (DD) failures that are often underestimated with respect to failure effects. The study aims to address the necessity of classifying and including various DD failures in risk and reliability analyses of low demand systems, and to suggest some ways to implement normative evaluations of DD failures to improve current work practices in the O&G industry. It is argued that there is a considerable potential to make the decision and works practices related to low demand safety instrumented systems more efficient and cost-effective without compromising safety risk.

1 Introduction

In the oil & gas (O&G) industry, low demand safety instrumented systems (SIS) are designated to prevent and mitigate hazardous events during operations. Due to the special operating characteristics of low demand systems, it is difficult to reveal failures of such systems during normal operations. With the absence of special measurement and monitoring tools, most failures of low demand SIS remain hidden and unrevealed. Among all possible failure types of SIS, the type of failures that can especially affects the operability and performance of related safety functions are always of great concern, which are defined as dangerous failures according to IEC 61508/61511 standard series (IEC 61508 2010; IEC 61511 2016).

Depending on the detectability of different dangerous failures, dangerous failures are classified as dangerous undetected (DU) and dangerous detected (DD) failures. According to IEC 61511 (2016), DD failures are those dangerous failures that detected by various means before they affect functions when needed, and DU failures are failures that cannot be detected by diagnostic tests. In the O&G industry, DD failures is a less discussed subject compared to DU failures. In general terms, DDs are considered relatively 'safe' if they are detected right after they occur and effective responsive actions can be taken (IEC 61508 2010; IEC 61511 2016). However, this is not always the case in practice. There is a tendency, mostly unintentionally, to underestimate the impacts of DD failures, which can lead to an incorrect risk perception and a weakened risk management plan. Hence, even though the IEC 61508/61511 defines the criteria of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 540–547, 2020.
https://doi.org/10.1007/978-3-030-48021-9_60

DD failures as 'safe', there is an emerging need for a normative approach to classify and evaluate various DD failures to improve current practices related to low demand SIS.

2 Methodology

The study follows the basic definitions and interpretations of dangerous failures based on IEC 61508/61511 standard series. The study was carried out on the Norwegian continental shelf (NCS) in close collaboration with industry and domain experts. It specially focuses on emergency shutdown (ESD) systems, which is a typical low demand SIS. In total 21 in-person meetings and interviews were conducted, involving experienced experts from operator companies, service providers, academic institutes, and authorities, to map different DD failures and to develop a practical approach to help evaluate such failures. The proposed approach thus was revised and verified in an iterative manner based on failure and maintenance data collected from industrial sources and expert knowledge.

3 Different Perceptions of Dangerous Detected Failures

The term dangerous detected (DD) failure was used by IEC 61508/61511 standard series. Different standards have different terms for such failures. IEC 60300-3-11 (2009) uses 'evident' to describe revealed or 'detected' failures, and uses 'safe/environmental' to denote 'dangerous' failures that have impacts on human lives and the environment. Since this paper studies safety instrumented systems, 'dangerous detected failures' from IEC 61508/61511 standard series is used in the following discussions.

In the IEC 61508/61511 standard series, the definition and evaluation of DD failures have been explained clearly. The impacts of DD failures are often considered minimal when diagnostic frequency is high (IEC 61511 2016). However, treatments of DD failures are not consistent within different parts of this series. IEC 61508 Part 4 (2010) includes DDs during the calculation of safe failure fraction (SFF) and excludes DDs from the estimation of probability of failure on demand (PFD). In the IEC 61508 Part 6 (2010), both DDs and DUs are taken into account in the calculation of PFD. In the IEC 61511 Part 1 (2016), both DDs and DUs are considered as contributing factors in the estimation of failure rate. The Norwegian oil & gas association published a new version NorOG-070 guideline (Norwegian Oil & Gas Association 2018) to implement IEC 61508/61511 in the Norwegian petroleum industry. The guideline considers DD failures safe and does not include such failures in the estimation and updating of PFD and safety integrity level (SIL).

As it seems, the interpretations and practices involving DD failures lack consistency between different international standards and industry guidelines. DD failures need to be studied from case to case so that their actual impacts are understood. This is especially the situation for low demand SIS, creating a timely need to reconsider DD failures.

4 Clarifications Related to DD Failures

In the Norwegian petroleum industry, dangerous failures of critical low demand systems have been recorded in risk level in Norwegian petroleum activities (RNNP) reports (Petroleum safety authority Norway 2018b). However, only DU failures are reported in practice. Based on the RNNP data, the observed failures (purely DU failures) and predicted failures (based on the statistical approach proposed by Hauge and Lundteigen (2008)) of riser ESD systems are plotted in Fig. 1.

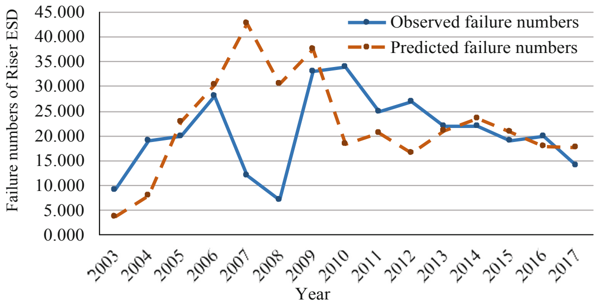


Fig. 1. Observed and predicted DU failures of riser ESD systems on the NCS

Due to the critical impacts of dangerous failures, conservative predictions are appreciated by the industry. However, it is seen from Fig. 1 that the prediction is almost half-time optimistic and half-time conservative. It is argued here that the exclusion of DD failures have contributed to the underestimation of failures.

The correct understanding and treatments of DD failures is important for planning corrective and preventive maintenance and performing quantitative risk & reliability analyses. It is argued here that the accuracy of predictions can be enhanced in comparison to current practice, by including DD failure data in the assessment process. As discussed in Sect. 3, current industry standards and guidelines do not provide a clear and normative guideline to achieve the goal. This section aims to clarify some most important aspects of DD failures related to low demand SIS.

4.1 Detection of DD Failures

According to IEC 61511 (2016), DD failures are related to dangerous failures that are revealed by diagnostic tests. However, based on practices in the Norwegian petroleum industry, detections of DD failures have occurred in other occasions. The paper summarizes the most common situations that DD failures are revealed, the frequency of occurrence, and the detectability of DD failures under different situations, as shown in Table 1.

The frequencies of different situations are different, as seen in Table 1. Depending on detection approaches and the availability of relevant data, the detectability of DD failures (i.e. detection coverage) under different situations varies from low to very high.

Table 1. Detection coverage of DD failures under different situations

Failure type	Detection situations	Frequency of occurrence	Detection coverage
DD	Visual inspection	High	Very low
	Traditional function test (incl. leak test)	Low	Low-medium
	Diagnostic test (SIF initiated)	Low	High-very high
	Diagnostic test (SIF not-initiated)	Very high	Low
	Real demands	Very low	Very low

On occasions when safety functions are actually activated, such as full-scale and partial stroke tests of ESD valves, the detectability of DD failures is normally higher compared to occasions when the systems are not activated, as seen in Table 1.

In practice, few DD failures can be revealed with visual inspections, traditional function tests, or real demands due to the technical limitations of traditional testing programs. Most DD failures are revealed by diagnostic testing, with an detection coverage of up to 60% with online condition monitoring (IEC 61508 2010). As implied by IEC 61511 (2016), high testing frequency is often assumed for diagnostic testing with online condition monitoring. However, Zhu and Liyanage (2019) pointed out that the window to collect truly valuable data is very limited due to the special operating mode of such systems in spite of the high sampling frequency of online diagnostic testing. It is suggested to specify if SIF is initiated during a diagnostic test. For low demand systems, the actual frequency of collecting valuable data sets largely depends on the frequency that the systems are fully or partially operated. Detectability of DD failures under those conditions is higher compared to other situations, as marked in Table 1. Most data that are collected during normal operations, when SIF is not initiated, do not reflect the real condition and performance of the system, and thus the detection coverage is low.

4.2 Actual Effects of DD Failures

In this paper, failure effects are discussed within the scope of safety to human lives and the environment, and the production interference is not discussed. In traditional terms, dangerous and safe are used to describe the effects or potential effects of failures without consideration of in-time detection and repair. When timely detections and effective repairs can be applied, dangerous failures are no longer considered dangerous in terms of actual effect (IEC 61511 2016). This paper uses 'safe' and 'dangerous' to describe the actual failure effects of various DD failures, considering detection and maintenance aspects.

Dangerous detected (DD) failures may occur as degraded or disabled failures (IEC 61511 2016). According to IEC 61511 (2016) definition, degraded failures do not deactivate safety instrumented functions (SIF) immediately but decrease reliabilities of SIFs; disabled failures mean completely deactivations of SIFs. According to the IEC standards, degraded failures only occur when fault tolerance design is implemented

(IEC 61511 2016). However, in practice, a dangerous degradation of a low demand system without fault tolerance can still occur without disabling the SIF, for example, a slow valve operation caused by increased valve surface friction. The definition of degraded failure is hereby extended to cover both reduced fault tolerance and degraded system performance. The terms ‘degraded’ and ‘disabled’ are thus used in this paper to avoid misunderstandings regarding definitions from IEC standards.

Depending on the time efficiency of failure detection and the effectiveness of repair, the actual effects of ‘degraded’ and ‘disabled’ DD failures vary. The actual effects of various DD failures were studied, and the results are shown in Table 2.

Table 2. Actual effects of DD failures, w.r.t. condition of SIF and repair.

Failure type	Availability of SIF	Effective repair after detection	Actual effects
DD	‘Degraded’	Yes	‘Safe’
		No	‘Dangerous’
	‘Disabled’	Yes	‘Dangerous’
		No	‘Dangerous’

‘Degraded’ DD failures, if repaired in time, will not greatly threat safety, and are thus considered ‘safe’. However, degradations can become unacceptable if the repair/restoration is not done effectively and in time. This kind of ‘degraded’ DD failures are thus considered ‘dangerous’ failures. When a detected DD failure is defined as ‘disabled’, the failure has already fully developed, which means that a SIF failure could have occurred if a real demand was presented. In another word, for ‘disabled’ DD failures, the ‘dangerous’ situation has already been developed regardless of whether the failures can be effectively repaired or not.

4.3 Inclusion of DD Failures in RNNP Reports

The risk level in Norwegian petroleum activities (RNNP) report is used by the authority to help operators identify challenges and manage risks in case of major accidents (Petroleum Safety Authority Norway 2018a). In practice, the requirements of RNNP largely affect how companies collect and understand various failures. It is thus very important that the RNNP report is setting up a practical and holistic scheme for companies to follow.

In the RNNP method, both degraded and safety critical (disabled) failures are reported, but it does not specify whether and how dangerous detected (DD) should be handled (ISO 14224 2016; Petroleum Safety Authority Norway 2018a). Since part of DD failures are as dangerous as DU failures, as seen in Table 2, ‘dangerous’ DD failures are here suggested to be included in the RNNP reports in addition to DU failures.

Related to low demand SIS, detections of dangerous failures are not limited to functional tests and internal leakage tests, as indicated by RNNP method (Petroleum safety authority Norway 2017). Based on industry practices, dangerous failures can be revealed by all situations that are identified in Table 1, and all detected 'dangerous' failures, including DUs and DDs, are suggested to be reported. A new data collection template is suggested, as shown in Table 3.

Table 3. Data collection form for RNNP report, using Riser ESD systems as an example

System	Failure detection methods	Total number of tests	Number of failures	
			DUs	'Dangerous' DDs
Riser ESD system	Visual inspection Traditional function tests (incl. leak test) Diagnostic testing Real demand			

It should be noted that the observed failures are predictably to increase as a result of the inclusion of 'dangerous' DD failures with the suggested practice. Nevertheless, the possibly increased reported failures and reduced reliability should not be considered as negative, as it contributes to a more enhanced failure analysis approach, and it helps operators as well as authorities to understand their systems better and gain more control during the management of these critical systems.

4.4 Inclusion of 'Dangerous' DD Failure in PFD Estimation

Probability of failure on demand (PFD) is used to measure the safety unreliability of a low-demand safety instrumented system. As explained in Sect. 3, the role of DD failures is ambiguous in the estimation of PFD. Based on the discussions in Sect. 4.2 and Sect. 4.3, the inclusion of part of DD failures is recommended, but evaluations of failures need to be done from case to case, as shown in Table 4.

Table 4. Inclusion of DD failures in PFD estimation

Failure classification	Detection method estimation	DD failures included in PFD
'Dangerous' DD	Visual inspection	Degraded DD failures without effective repairs; All disabled DD failures
	Traditional function test (incl. leak test)	Degraded DD failures without effective repairs
	Diagnostic test (SIF initiated)	Degraded DD failures without effective repairs; All disabled DD failures
	Diagnostic test (SIF not-initiated)	Degraded DD failures without effective repairs
	Real demands	–

In general, two types of DD failures are suggested to be included in PFD calculation, including degraded DD failures that are not repaired in time, and all disabled DD failures, as explained in Sect. 4.2.

It should be noted that, during a traditional function test or a real demand, 'disabled' dangerous failures can also be detected, but such failures are defined as DU failures instead of DD failures. For DD failures that can be detected by diagnostic test during normal operations (SIF not-initiated), such as leakage through valve seat (Juvik et al. 2002), degradation process is monitored so that a 'disabled' situation can be prevented. However, the detection coverage is low, as seen in Table 1. In this situation, those detected degraded failures that are not repaired in time will become 'dangerous', and thus are suggested to be included in PFD estimation. For the revealed 'degraded' DD failures during real demands, the operability of SIFs will not be affected, and are thus considered 'safe'. This part of 'degraded' failures are thus not included in the PFD estimation.

5 Conclusions

In the Norwegian petroleum industry, IEC 61508/61511 standard series and the NorOG-070 guideline are used to manage low demand safety instrumented systems. The study challenges some of the current practices on evaluating dangerous failures and estimating PFD. The paper addresses the importance and proposes some practices to understand, classify, and treat various DD failures, taking into account of the time efficiency of detection and repairs upon detections. The suggested classification of 'dangerous' DD failure in this paper also helps clarify the ambiguity that exists during the estimation of PFD in the current approach.

Low demand SIS are commonly used as technical barrier elements or independent protection layers in the O&G industry. The paper rethinks the foundations of several commonly adopted approaches related to such systems. The re-evaluation of DD failures, in many cases, is believed to contribute to an improved understanding of safety and reliability of low demand SIS and a conservative risk management culture. The arguments are also supported by the operating experiences and data from diagnostic testing implementations with online condition monitoring on the NCS.

It should be noted that the 'dangerous' failure rate is likely to arise in the short-to mid-term as a result of the inclusion of DD failures. The paper interprets this effect as an improved understanding of real risk picture. The paper also underlines the timely need to improve the current auditing and supervisory practice, for instance using the suggested approach as an additional process to have better confidence related to unwanted safety events. Further studies within this domain can focus on digital technologies, as well as different roles and interfaces between the operators, domain experts, and authorities.

References

Hauge, S., Lundteigen, M.A.: Guidelines for follow-up of Safety Instrumented Systems (SIS) in the operating phase. SINTEF, Trondheim, Norway, Report A, 8788 (2008)

ISO 14224: ISO/TC 67, ISO 14224:2016, Petroleum, Petrochemical and Natural Gas Industries – Collection and Exchange of Reliability and Maintenance Data for Equipment, 3 edn. International Organization for Standardization (2016)

IEC 60300-3-11: IEC 60300-3-11, Dependability Management - Part 3–11: Application Guide - Reliability Centered Maintenance. International Electrotechnical Commission, Geneva (2009)

IEC 61508: Functional Safety of Electrical/Electronic/Programmable Electronic Systems, Parts 1–7. IEC 61508. International Electrotechnical Commission, Geneva (2010)

IEC 61511: Functional Safety-Safety Instrumented Systems for the Process Industry Sector, Part 1–3, vol. 1. International Electrotechnical Commission, Geneva (2016)

Juvik, T., Hermansen, T., Carr, R., Hale, S.: Online valve monitoring systems used on off-shore platforms in the North sea. Paper presented at the ASME 2002 21st International Conference on Offshore Mechanics and Arctic Engineering (2002)

Norwegian Oil & Gas Association: Norwegian oil and gas 070-Application of IEC 61508 and IEC 61511 in the Norwegian petroleum industry, Stavanger (2018)

Petroleum Safety Authority Norway: RNNP data reporting on the Norwegian continental shelf (2017). http://www.ptil.no/getfile.php/1346012/PDF/RNNP-datainnsamling/2017%20RNNP%20datainnsamling%20norsk%20sokkel.xlsx

Petroleum Safety Authority Norway: Requirements for reporting performance barriers (2018a). http://www.ptil.no/getfile.php/1346683/PDF/RNNP-datainnsamling/Krav%20til%20rapportering%20barrierer%20rev15_2.pdf

Petroleum Safety Authority Norway: Risk level in Norwegian petroleum activities (RNNP) (2018b). http://www.ptil.no/om-rnnp/category720.html

Zhu, P., Liyanage, J.P.: Application of prognostics and health management to low demand systems: Use of condition data to help determine function test interval. Paper presented at the 2018 International Conference on Industrial Engineering and Engineering Management, Bangkok (2019)

Towards Establishing an Integrated Health Index for Rail Rolling Stock

Pieter Conradie and Johannes Jooste(✉)

Department of Industrial Engineering, Stellenbosch University, Stellenbosch, South Africa
{pieterc, wyhan}@sun.ac.za

Abstract. In recent years, the concept of asset life extension has become popular in lieu of other end-of-life management strategies, such as replacement or decommissioning of assets. This paper identifies a health index as key contributor during life extension decisions and discusses the application of health indices in various industries. It further discusses the building blocks that need to be considered in the development of a health index, and demonstrates the development of a conceptual health index for rail rolling stock for the Passenger Rail Agency of South Africa. It concludes with a discussion how the conceptual health index can be applied, and the future work which can performed to improve the health index.

1 Introduction

There is an increased demand for improved technical and financial performance of railway operators, and it has subsequently resulted in pressure to control costs while maintaining or improving system performance (Nomad Digital 2015). Capital intensive industries are likely to adopt the asset management framework, as described by the ISO 55000 (2014) standard. The standard requires asset management objectives, which are the link between the organisational objectives and the asset management plans. Amongst these, asset management objectives, asset condition (performance or health score), life cycle costs and life expectancy should be considered.

The concept of asset life extension has become increasingly important in the rail industry, specifically for passenger rail rolling stock (RRS). Globally, the rail industry experiences many challenges, from underinvestment, ageing assets, the unlocking of the capacity of the assets, integration of digitisation into rail, passenger satisfaction and the integration of rail with other modes of transport (Partners 2016). In particular, the underinvestment in rail transport necessitates the rethinking of asset management strategies, in order to slow down the ageing of the assets.

2 Theoretical Background

The concept of asset life extension has become increasingly important in recent years, in particular in safety-critical industries. Life extension can offer a broad range of benefits ranging from technical, economic and social, to environmental benefits.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 548–556, 2020.
https://doi.org/10.1007/978-3-030-48021-9_61

However there must be a balance between extending the life of an asset and managing cost, maintenance, technology and being fit for service. The scholarship about health indices (HIs) and life extension is growing, which is an indication that HIs are becoming popular for life extension decision-making. In the absence of HIs, life extension decisions are often made using a selection of indicators which could lead to distorted conclusions.

HIs for non-living entities are undefined. The approach of social sciences is to associate an HI with the conformity to *well-being* and *role performance*. The HI of an asset is therefore an integrated indicator which is designed, to reflect or characterise asset condition and consequential asset performance in terms of the asset's functional context (Heywood and McGrail 2015). Jahromi et al. (2009) describe the HI as a practical tool that combines the results of field inspections, operating observations and site and laboratory testing into an objective and quantitative index, which provides an indication of the overall health of the asset. From these definitions, it is clear that an HI not only refers to the condition of an asset, but it takes into consideration more aspects of the asset.

According to Heywood and McGrail (2015), two main characteristics of an HI are that it must have a purpose and it must be auditable. Therefore, during the development of an HI, special care must be given to ensure that the HI solve a specific problem, and that the calculation of the HI is repeatable. Deloitte and the Canadian Electricity Association (2014) define three steps in building an HI:

Step 1: Define the asset hierarchy and identify critical asset classes within the hierarchy.
Step 2: For each asset class, define the determining factors, the relative importance for each factor, and determine the evaluation methodology.
Step 3: Conduct asset condition assessments, refine the evaluation methodology with stakeholders, and integrate the methodology within the asset management process.

The second step is arguably the most important in the development of an HI (Islam et al. 2017). Azmi et al. (2017) provide a more scientific view on this step, by redefining the step into the identification of the input parameters, and the development of an algorithm or aggregation method (mathematical equation) for calculating the HI.

It is however clear that the relationships between indicators need to be identified and quantified, for it to be successfully integrated into a single index. Basic HIs use the weighted average method (also referred to as the "scoring and ranking" method) to aggregate indicators, and more complex HIs use advanced aggregation methods. Literature about the use of HIs for life extension decisions are available for many industries, and a short description will follow for the most prominent industries.

HIs are widely used in *power transformers* and is a combination of expert observations, various test results, and field inspection data in a single quantitative index. These HIs consider the integration between different routine and diagnostic tests that are not considered by classical condition monitoring techniques (Islam et al. 2017). The calculated HI can identify transformers that have a higher probability of failure and are close to the end of their life, which can help to identify transformers that need more attention or major capital expenditure. HIs are based on aggregation methods such as the linear weighted average (Islam et al. 2017), fuzzy logic (Abu-Elanien et al. 2012), binary logistic regression (Zuo et al. 2016), general regression neural network (Islam et al. 2017), binary cat swarm optimisation (Mohamadeen et al. 2014) and orthogonal wavelet-networks (Ahmed et al. 2015). The *oil and gas industry* relies on condition assessments to monitor different components. In the literature, reference is made to case studies where different techniques are used to model the condition of different equipment, ranging from risk-based integrity management techniques (Jansen and Van 2005), to an integrated technical and economic technique (Copello and Castelli 2013). A process was established for the technical health assessment of subsea oil and gas systems, and the relationship between technical health, remaining useful life and life extension decision-making was established (Vaidya and Rausand 2011). The *built environment* and fixed *transit assets* are two main groups where life extension and condition ratings are applied successfully in industry. In both industries, indicators are integrated into a single condition index, which is similar to the HI discussed previously. The health indicators are either based on the weighted average of measures (Lavy et al. 2010) or the sum of deferred maintenance (in monetary terms) (Uzarski and Grussing 2011). The application of HIs in different industries and the variety of aggregation methods used for developing HIs are noteworthy. It illustrates how most industries use the linear weighted average method to aggregate indicators, and it illustrates the variety of approaches that can be adopted in pursuit of the extension of asset life.

In the rail context, the condition and performance of RRS can be affected by many factors such as maintenance strategies, driver behaviour, environmental conditions, load conditions, commuter behaviour and the effects from the wheel-rail interface and the pantograph-contact wire interface. The problem is however how to measure condition and performance based on the combined effect of these factors. This necessitates a decision support system for understanding which of these factors are important, and what the impact of these factors are on the asset life for deciding on corrective strategies.

From a literature review of leading research databases, no integrated HI could be found for RRS. Some on-board systems are commercially available for the continuous monitoring of RRS, which can measure the condition and performance of selected systems. In addition, these systems can monitor some of the factors influencing the life of RRS (such as driver controls and passenger load), however only limited factors are taken into account (Galar et al. 2013). Although some commercial on-board systems are labelled as “health systems”, these systems do not have the capability to convert the measures into an integrated HI for the RRS.

3 Development of a Conceptual Health Index for Rail Rolling Stock

In this section, an investigative study for the development of an HI for RRS is proposed. The development of the framework is done according to the process of Deloitte and the Canadian Electricity Association (2014). For the purpose of this paper, step one is not elaborated on and RRS is identified as the critical asset class for the development of the conceptual model.

From a comprehensive literature review, 46 indicators are identified in relation to the condition or performance of RRS. These indicators are diverse and from different perspectives, and although some researchers combined some of these measures, no single quantitative HI is proposed. It is acknowledged that there may be more factors that could influence the condition and performance of RRS however for this research paper, only the 46 indicators are consider for exploratory purposes. These 46 indicators consists of leading and lagging indicators. A foundation of leading indicators is implied for HIs, while a basis of lagging indicators implies more of an "asset value" or "asset satisfaction" index. To remain consistent with scholarship terminology reference to HIs is maintained, although the study results and proposed HI may include lagging indicators.

The 46 indicators can broadly be classified into categories such as reliability, availability, maintainability and safety (the classical RAMS indicators), as well as categories for performance, logistic support and general indicators. To identify the most significant indicators with their associated weightings from this list, a two-step approach is followed:

1. Consensus voting is done by a five participant focus group who were purposively selected from the engineering division of PRASA. These participants are subject matter experts in the asset management of PRASA's RRS. During the consensus voting, the list of 46 indicators was reduced to a list containing the 20 most significant indicators (listed in no particular order in Table 1).
2. A questionnaire is completed by an expanded group of 12 participants who was purposively selected to refine the list to the five most significant indicators. Similar to the consensus voting, these participants are subject matter experts knowledgeable in the PRASA RRS fleet. The best-worst scaling (BWS) method is used because of its simplicity, accuracy and unbiasedness. The method necessitates that respondents make trade-offs between items and their benefits, which overcomes the issue of many items having similar importance weights (Rezaei 2015; Cohen 2009).

During the BWS method, respondents were required to complete a questionnaire containing 20 choice sets of six indicators each, with each indicator appearing six times in the questionnaire. It is not possible to have a fully balanced choice set design using 20 indicators, and it is unavoidable that some indicators will appear more than once in combinations. However, the effect of this will be regarded as marginal with no influence on the end result. From each choice set, the respondents chooses one indicator which contributes the most to an HI for RRS (labelled *best*), and choose one indicator which contributes the least to an HI (labelled *worst*). Best-worst (B-W) scores and average B-W scores are calculated for each indicator, and the results are summarised in Table 1.

Table 1. Importance of RRS indicators as prioritised by the PRASA focus group

No	Indicator	Total best	Total worst	B-W score	Average B-W score
1	Mean time to failure	13	5	8	0.1111
2	Maintenance skills pool availability	17	12	5	0.0694
3	Number of temporary equipment cut-outs	11	16	−5	−0.0694
4	Performance of a critical system	23	0	23	0.3194
5	Time to return to safety	13	12	1	0.0139
6	Mean time between 'safety system failure'	5	1	4	0.0556
7	Fault correction time	7	4	3	0.0417
8	Deferred maintenance	4	23	−19	−0.2639
9	Maintenance support performance	13	12	1	0.0139
10	Train availability	42	6	36	0.5000
11	Fault/repair coverage	3	25	−22	−0.3056
12	Mean time to restore	12	7	5	0.0694
13	Repair time	14	10	4	0.0556
14	Mean time to maintain	6	14	−8	−0.1111
15	Failure probability or probability of success	14	13	1	0.0139
16	Mean down time	9	7	2	0.0278
17	Indicators from on-board monitoring systems	9	13	−4	−0.0556
18	Train age	5	46	−41	−0.5694
19	Mean time between hazardous failure	9	7	2	0.0278
20	Spare parts in stock when needed	11	7	4	0.0556

The average B-W scores are calculated by dividing the B-W score by the number of times the indicator appears in the choice sets, and dividing by the number of respondents (which were 12) (Cohen 2009). By ranking the average B-W scores, the most important indicators can be identified. In Fig. 1, the ranked average B-W scores are displayed, and it can be seen that the most important indicator is "train availability" (#10), and the least important indicator is "train age" (#18). For the remainder of the discussion, reference will only be made to the indicator number.

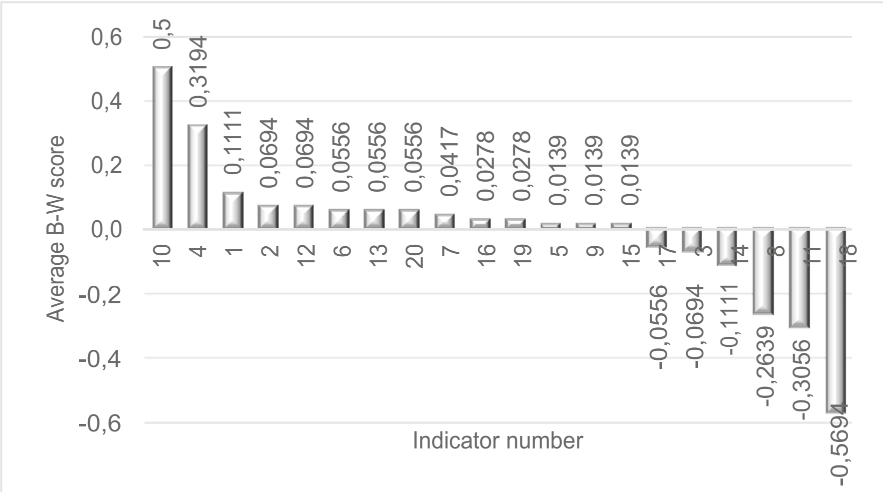


Fig. 1. Average B-W scores of indicators

The linear weighted average method is used as aggregation method, on the basis of its wide application in industry. A generic HI is presented in Eq. 1:

$$HI = \frac{\sum_{i=1}^{n} a_i w_i}{\sum_{i=1}^{n} w_i} \tag{1}$$

where n represents the number of health indicators, a_i - represents the i^{th} indicator and w_i - the associated weight of the indicator. A combination of least square means and Pareto analysis is used to arrive at the selection of five indicators to be included in the HI. Statistically, according to the least square means method, indicators 10 and 4 contribute the most significantly to the HI. However, a Pareto analysis shows that five indicators contribute the majority towards the final HI. Consequently, these five indicators are included for further analysis to account for the overall combined effect of the indicators in the HI, despite the fact that three of the five indicators contribute less significantly, than the other two indicators. These indicators are 1, 2, 4, 10 and 12. By using the B-W scores of these indicators as relative weights in Eq. 1, the conceptual HI is represented in Eq. 2:

$$HI = \frac{0.111.a_1 + 0.069.a_2 + 0.319.a_4 + 0.500.a_{10} + 0.069.a_{12}}{0.111 + 0.069 + 0.319 + 0.500 + 0.069} \tag{2}$$

To compare the HI indicators (a_i), the value of each indicator is normalised. Normalisation is done by dividing the actual value of the indicator by the theoretical maximum or ideal value which such an indicator can achieve. The consequence is that a_i - is represented by proportional values between 0 and 1, with a higher HI value indicating an improved HI. The maximum theoretical value for the HI is therefore 1. The normalisation of the indicators and the application of the HI is described in the next section.

4 Discussion and Conclusion

The application and use of the developed conceptual HI is discussed in this section. The diversity of the indicators contributing to the HI and the inclusion and absence of some indicators are of interest.

Table 2 illustrates the application of the conceptual HI on the fleet of train sets for PRASA in the Western Cape, South Africa. To illustrate normalisation, the actual values and the theoretical maximums or ideal values for each indicator are listed, as well as the unit and the source of data. The indicators in Table 2 are characterised as follows:

1. The maximum value for any indicator is unity.
2. The values of indicators 1, 2, 4 and 10 are expressed as fractions, while the value of indicator 12 is defined as: $1 - \frac{\text{MTTR}}{\text{MTTR}_{\text{max}}}$.
3. The traction motors were chosen as critical system for indicator 4, and a collective mean time to failure (MTTF) was obtained for all the traction motors in the fleet.

Using these values in Eq. 2 produces an HI of 0.517. This means that the fleet is only half as healthy as its potential. It is evident that train availability (#10) has an adverse effect on the HI, and by merely having the required number of train sets in service, an HI of at least 0.468 can be achieved. This reflects the importance of train set availability in the South African context, where operational conditions and unsocial behaviour severely influence the train set availability within the PRASA fleet.

Table 2. Application of the conceptual HI on a fleet of train sets for PRASA

No	Value	Max value	*ai*	Unit	Note
1	8.83	14	0.631	Day	MTTF for fleet
2	304	418	0.508	%skill	Actual vs required skills
4	766.2	1826	0.420	Day	MTTF for traction motors
10	49	88	0.557	#sets	Actual vs required sets
12	5	10	0.500	Min	MTTRestore failures

In the illustrative application, the HI is applied to a fleet of train sets. It can similarly be applied to individual train sets. The purpose thereof will be to rank train sets by their HI score, and train sets with a low HI (poor health) can then be identified for further investigation and analysis. Similar to the application of HIs for transformers, the calculated HI can identify train sets that have a higher probability of failure and are close to the end of their life, which can help to identify train sets that need more attention or major capital expenditure.

In recent times, the rail industry is transformed by information technology and the increased use of digital technology. Digitisation and big data is changing how rail operators collect, analyse and use data (Peterhans and Price 2017). Rail operators will find the most value in an HI when it is calculated using real time data, therefore it is critical to integrate the real time RRS and operational data into the HI. By using algorithms, the continuous monitoring of the HI of each train set can be useful, and any sudden deviation in the HI can be identified for immediate investigation. Similarly the future HI of a train set can be predicted based on the current HI, maintenance practices and operating conditions.

It is concluded that the research provides insights as to how HI can be used in support of RRS decision-making that is purposeful and auditable. The HI is developed using focus groups from the Engineering division of PRASA, and applied to a fleet of train sets. The linear average method allows for the practical aggregation of indicators from readily available historical data. This research sets an exploratory research basis and provides a framework for the development of an HI for RRS and other assets. It further creates the platform for dialog and discussion, and the opportunity for improving the conceptual HI. Future work is proposed to refine the aggregation methods, by considering machine learning and data mining techniques, as well as taking into account the possibility of dependencies between indicators, and applying the HI to individual train sets.

References

Abu-Elanien, E.B., Salama, M.M.A., Ibrahim, M.: Calculation of a health index for oil-immersed transformers rated under 69 kV using fuzzy logic. IEEE Trans. Power Deliv. **27**, 2029–2036 (2012). https://doi.org/10.1109/TPWRD.2012.2205165

Ahmed, M., Elkhatib, M., Salama, M., Shaban, K.B.: Transformer health index estimation using orthogonal wavelet network. In: 2015 IEEE Electrical Power & Energy Conference, pp. 120–124 (2015). https://doi.org/10.1109/epec.2015.7379937

Azmi, A., Jasni, J., Azis, N., Kadir, M.Z.A.A.: Evolution of transformer health index in the form of mathematical equation. Renew. Sustain. Energy Rev. **76**, 687–700 (2017). https://doi.org/10.1016/j.rser.2017.03.094

Cohen, E.: Applying best-worst scaling to wine marketing. Int. J. Wine Bus. Res. **21**, 8–23 (2009). https://doi.org/10.1108/17511060910948008

Copello, S., Castelli, P.: Life extension of a fixed offshore platform structure based on monitoring results. In: Offshore Mediterranean Conference and Exhibition (2013)

Deloitte, Canadian Electricity Association: Asset health indices: a utility industry necessity (2014)

Galar, D., Kumar, U., Villarejo, R., Johansson, C.A.: Hybrid prognosis for railway health assessment: an information fusion approach for PHM deployment. Chem. Eng. Trans. **33**, 769–774 (2013). https://doi.org/10.3303/CET1333129

Partners, H.: Bringing actionable recommendations to revitalise innovation and entrepreneurship in the rail sector (2016)

Heywood, R.J., McGrail, T.: Clarifying the link between data, diagnosis and asset health indices. In: Asset Management Conference 2015, pp. 1–6 (2015). https://doi.org/10.1049/cp.2015.1748

Islam, M.M., Lee, G., Hettiwatte, S.N.: Application of a general regression neural network for health index calculation of power transformers. Int. J. Electr. Power Energy Syst. **93**, 308–315 (2017). https://doi.org/10.1016/j.ijepes.2017.06.008

ISO 55000: Asset management. Overview, principles and terminology. International Organisation Standard (2014)

Jahromi, A., Piercy, R., Cress, S., Service, J., Fan, W.: An approach to power transformer asset management using health index. IEEE Electr. Insul. Mag. **25**, 2 (2009). https://doi.org/10.1109/MEI.2009.4802595

Jansen, M., Van, G.: Life extension of degraded main oil line pipeline sections through improved risk based integrity management. In: Corrosion 2005. NACE International (2005)

Lavy, S., Garcia, J.A., Dixit, M.K.: Establishment of KPIs for facility performance measurement: review of literature. Facilities **28**, 440–464 (2010). https://doi.org/10.1108/02632771011057189

Mohamadeen, K.I., Sharkawy, R.M., Salama, M.M.: Binary cat swarm optimization versus binary particle swarm optimization for transformer health index determination. In: 2nd International Conference on Engineering and Technology, ICET 2014, p. 4 (2015). https://doi.org/10.1109/icengtechnol.2014.7016812

Nomad Digital: Rail industry survey: major challenges facing rail operators, maintainers & owners and the role of ICT (2015)

Peterhans, G., Price, M.: Using big data: there is no time for business as usual. Glob. Railw. Rev. **1**, 8–10 (2017)

Rezaei, J.: Best-worst multi-criteria decision-making method. Omega (U.K.) **53**, 49–57 (2015). https://doi.org/10.1016/j.omega.2014.11.009

Uzarski, D.R., Grussing, M.N.: Building condition assessment metrics: best practices, 1–5 (2011)

Vaidya, P., Rausand, M.: Remaining useful life, technical health, and life extension. Proc. Inst. Mech. Eng. Part O J. Risk Reliab. **225**, 219–231 (2011). https://doi.org/10.1177/1748007810394557

Zuo, W., Yuan, H., Shang, Y., Liu, Y., Chen, T.: Calculation of a health index of oil-paper transformers insulation with binary logistic regression. Math. Probl. Eng. (2016). https://doi.org/10.1155/2016/6069784

Dynamic and Modular Business Models for Maintenance

Mirka Kans[1(✉)] and Anders Ingwald[2]

[1] Department of Mechanical Engineering, Linnaeus University, Växjö, Sweden
mirka.kans@lnu.se

[2] Department of Physics and Electrical Engineering, Linnaeus University, Växjö, Sweden
anders.ingwald@lnu.se

Abstract. With more complex production environments maintenance must be viewed as a value-adding service, and this in turn require new business models and new ways to design service agreements. Many companies have not yet understood the business opportunities that could be achieved by providing maintenance services, and efficient strategies for enabling these opportunities are lacking. Moreover, many believe that you have to choose either or, thus either one provides product-oriented maintenance or value-oriented maintenance. This paper proposes an integrated, dynamic and modular approach to maintenance business model development. Modular-based maintenance offerings classify maintenance services with increasing integration of the offering, and increasing focus on utility for the customer and the customer's customer. Continuing, modular-based maintenance offerings allow for flexibility; one does not have to choose between the product-centred or utility-centred business models. Instead, the offering is packaged based on the available internal resources and key capabilities of the service provider which are matched against specific customer needs. The dynamics of the maintenance offerings are the time and scope dimensions describing the boundaries in which maintenance execution could take place at the customer.

1 Introduction

Changes in the economy forces enterprises to change the business behaviour and business models, focusing on value to the customers rather than selling products. Customer offerings have therefore become more knowledge intensive. Visnjic Kastalli and Van Looy (2013) state that business models based on combined offers, i.e. products combined with services, such as maintenance, have a positive effect on the business, and the extension of products with service offerings show up the highest economical benefits. The key to success is to view the services as a strategic complement, a new business model, and not an add-on product. Digitalisation creates new opportunities within the maintenance area, such as selling value in form of performance or output instead of repair hours and spare parts. In order to succeed, a shift of focus from what is offered to which value the offer gives for the customer is required (from necessary evil to something that enables value creation), and to adopt a holistic view of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 557–565, 2020.
https://doi.org/10.1007/978-3-030-48021-9_62

the value creation process (Kans and Ingwald 2016). Moreover, the maintenance services must be aligned and integrated with the customers' business strategy as well as the service providers' (Barbera et al. 2012). This mental shift is required both for the supplier of maintenance and for the customer. In fact, it is not about maintenance management; it is about managing the physical assets throughout their live time, i.e. Asset management. Moreover, it is not a question to choose; i.e. either to provide traditional maintenance or value-based maintenance. Instead, the offering should be based on the customer needs and maturity (Osterwalder and Pigneur 2010). The purpose of this paper is to propose an integrated, dynamic and modular approach to maintenance business model development. Modular-based maintenance offerings classify maintenance services with increasing integration of the offering, and increasing focus on utility for the customer. The dynamics of the maintenance offerings in form of time and scope dimensions describes the boundaries in which maintenance execution could take place at the customer.

2 Maintenance Dynamics and Maintenance Offerings

The dynamics in maintenance could refer to dynamic maintenance planning with respect to when to act and what to perform (planning issue), or prioritising of activities to be performed (scheduling issue). The relative dynamics in scope, i.e. which activities to be carried out during a maintenance related stoppage, differs depending on the strategy that is applied. Opportunistic maintenance (OM) for instance combines corrective, preventive and condition-based maintenance strategies for optimal outcomes by replacing components in the same system during a shutdown, or replacing components in other machines if the whole production line is shut down due to maintenance (Zhou et al. 2015). The dynamics in the time dimension is affected by factors such as degradation rate (see Sun et al. 2012), and level (echelon, subsystem, system or production level) in which maintenance planning takes (Mollaverdi and Mirabadi 2015). Machines might experience different degradation rates for similar components due to environmental factors. This result in problems with creating optimal maintenance plans both on echelon level and on the production system level. Moreover, the traditional approach of applying different maintenance strategies for single machines affects the possibility to apply an optimised production wide maintenance strategy (Mollaverdi and Mirabadi 2015). The dynamics could also be affected by operational and environmental contexts, e.g. systems situated outdoors such as wind power mills, where weather conditions have to be taken into account when planning the maintenance (Erguido et al. 2017).

The term customer offering refers to the value proposition of a business in order to satisfy a customer (Osterwalder and Pigneur 2010). The maintenance offerings could be categorized into three types, see Ingwald and Kans (2018). **1) Resource-based offerings:** The simplest business agreement is the single purchase characterized by clear boundaries; often there is a fixed price that the buyer and seller agree upon, and the purchase is a one-time isolated occurrence. In resource-based maintenance contracts, the seller (supplier) pays for actual expenses, such as labour costs and spare parts, as well as a profit margin. **2) Performance-based offerings:** Performance-based contracts are regulated based on a predefined performance. The supplier guarantees a certain level of performance and the customer pays for direct and indirect expenses and the increased risk the supplier takes. Performance-based contracts are often in the form of the holistic contracts covering the maintenance needs of a system to ensure uptime or function. **3) Utility-based offerings:** Utility-based contracts guarantee not only the uptime of a product or system, but the customer's total operation. In these contract forms, the supplier takes a greater risk, in that it ensures the business and not a specific system or product's operation, and this risk must be carefully priced and regulated in the contract. Shohet and Straub (2013) show that performance-based contracts are more cost effective than traditional maintenance contracts. According to Bakshi et al. (2015) the customer preference for contract type is affected by the perceived reliability. Customers are often willing to enter into performance-based contracts for mature products, but choose a resource-based contract for new products because there is no reliability data for the new product. Lieckens et al. (2015) note that the optimal pricing of performance-based contracts depends on the price sensitivity of the customer and the sensitivity to stoppages in the production. Sinkkonen et al. (2013) argues that both the customer and supplier reach advantages in so-called partnering based contracts characterised by common goals and transparency. This is because maintenance is the core competence of the supplier; the customer does not have the financial and/or technical premise to carry out effective maintenance, while the supplier is able to invest in for instance condition-monitoring technology to improve their performance, has accumulated expertise in maintenance, and can also achieve economies of scale.

Customers can be divided into different groups, so called customer segments, (Osterwalder and Pigneur 2010). Segmentation could be made based on geographical or product-specific characteristics, or based on type of customer-supplier relationship. Osterwalder and Pigneur suggest that a customer group constitutes a separate entity when:

- they require different types of relationships
- they are accessed through different distribution channels
- their needs require and justify a distinct offering
- they are willing to pay for different aspects of the offer
- they are profitable in substantially different ways

Customer relationships could be in form of self-services or automated services or personal, i.e. the customer is served by a physical person, or as co-creation (Osterwalder and Pigneur 2010). If the person is associated with specific customers, the service is dedicated. Co-creation occurs when the customer is invited to participate in the product development process. The distribution of the offering could be through own channels, third parties, or a mix of both (Osterwalder and Pigneur 2010). The offerings could be in form of products, services, or a mix of both (Kans and Ingwald 2016). A product is a physical entity that is a value-enhancing resource for the customer, where production and delivery are often separated. The service is an intangible offering that is created and delivered simultaneously. Integrated offerings are a mix of products and services, and could be distributed in form of leasing contracts or life cycle solutions. Companies act in a business environment with ever-increasing complexity (Olve et al. 2013). Traditional chains of value creation are replaced with networks of actors, or business ecosystems, where the same actor could possess different roles, such as producer, distributor, service provider or supplier, depending on the context.

3 Modular-Based Maintenance Offerings

In order to meet customer demands, the offerings have to be apprehended as different by the customer. This could be reached by customer specific contracts with detailed specifications on row level, or by combining a basic package with optional add-ons. The result is often complex and administration heavy aftersales business. Modular maintenance offerings is a way to meet customer demands in form of choices and customization while reducing the variation, and thereby reduce internal administration, as well as resource utilization. Figure 1a describes three dimensions to consider when developing modular maintenance offerings: time, scope and maturity of the seller as well as the buyer. Time and scope are described according to their dynamic behavior, see Fig. 1b. Time rigidness occurs when the maintenance plan must be fixed and cannot change, for instance when maintenance has to be coordinated with external resources or require heavy administration, such as major revisions. Scope rigidity occurs when the tasks to be made during the maintenance cannot be changed, while dynamics in scope allows for changes in the tasks to be performed during the maintenance intervention. A highly dynamic maintenance plan allows for late decisions regarding the maintenance intervention. Often, the planning is based on condition monitoring and real life events. The maturity dimension describes the customers' maturity in several aspects: procurement maturity, i.e. the ability to handle procurement and the availability of internal processes for handling procurement, maintenance maturity, i.e. existing skills and maintenance organization, and technology maturity, i.e. the current automation and digitalization of the company (Kujala et al. 2011).

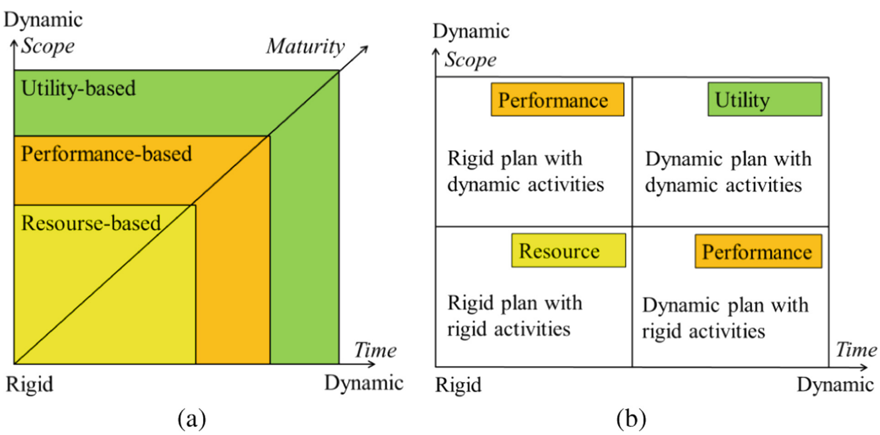


Fig. 1. a Modular-based maintenance offerings. b Dynamics in the maintenance offering.

The core value proposition, see Fig. 1a, is maintenance as a **resource**. The supplier provides maintenance planning and execution according to a basic contract, with a pricing based on the fixed maintenance plan. Core in-house maintenance knowledge and skills of the supplier forms the basis for these types of contracts. The next level considers maintenance from a **performance** perspective, which is connected to the functionality of a specific system. Key competencies for these types of contracts are, in addition to the ones for resource-based contracts, condition monitoring and diagnosis skills and technologies. The supplier might extend the core in-house competencies with external capabilities, such as expertise in oil analysis or spare parts logistics. **Utility-based** contracts stress the benefits of the maintenance offering from the customer's perspective. The benefits are found in the core business of the customer, expressed e.g. as productivity or production efficiency, rather than on the system performance. In addition to key competencies described above, the ability to coordinate the production and maintenance planning is required (an asset management perspective is adapted rather than a maintenance perspective). Long term and close relationships with the customer is needed and could also be seen as a key resource: without willingness to share operational data and production plans, the optimized planning could not be achieved (Ng et al. 2013). Financial incentives for all partners is important, and profit-sharing mechanisms is one way to achieve this (Sinkkonen et al. 2013). Modular maintenance offerings provide an integrated classification of maintenance services. The one and same maintenance service provider can therefore offer contracts on all three levels to different customer segments. Even if the different offerings are perceived as separate business propositions by the customer, the different contracts are logically based on one of three levels (modules) of the utility-based offering; from the outside the offerings seems quite different, but internally they are all parts of the utility-based offering. A customer segment could be based on aspects such as the customers' needs to require and justify a distinct offer, the distribution channel, the need of different types of relationships, different ways to gain a, or willingness to pay for different aspects of the offer (Osterwalder and Pigneur 2010). For maintenance services, the segmentation could be based on the customer requirements and demands (connected to dynamics in time and scope), as well as the ability to handle a certain type of contract (the maturity). The maturity dimension also takes into account the suppliers' maturity, such as marketing approach and maturity of the product that maintenance services are intended for (Kujala et al., 2011; Bakshi et al., 2015). Three generic customer segments are proposed for the modular maintenance offerings taking into account the above mentioned factors, see Table 1.

Table 1. Generic customer segments and their characteristics.

Dimensions and factors	Resource-based	Performance-based	Utility-based
1. Technology: a) Perceived complexity of technology[1]	a) Low complexity	a) High complexity	a) High complexity
2. Maintenance: a) Maintenance dynamics b) Customer's maintenance organization[1], c) Skill level of the maintenance organization[1]	a) Time and scope rigidness b) No own organization c) Low skill level	a) Dynamics in time or scope b) No own organization/own organization c) Low to high skill level	a) Time and scope dynamics b) Own organization c) High skill level
3. Business: a) Customer's core business[1], b) Customer relationship[3]	a) Non-business critical/Business critical b) Non-complex, value chain	a) Non-business critical b) Complex, network	a) Business critical b) Complex, ecosystem
4. Procurement: a) Business practices of the customer[1,3], b) Revenues and pricing[2,3], c) Supplier relationship[2], d) Distribution/marketing channels[1,2]	a) Product-oriented b) Transactional or fixed price contracts c) Provider: Personal, self-service, or automatic d) Third party	a) Performance-oriented b) Leasing or performance-based contracts c) Enabler: Personal or dedicated personal d) Third party/own channels	a) Utility-oriented b) Long term contracts c) Partner: Dedicated personal, d) Co-creation Own channels

[1]Kujala et al., 2011, [2]Osterwalder and Pigneur, 2010, [3]Olve et al, 2013

3.1 Application Example, the Power Generation Industry

The concept of modular maintenance offerings is in this section illustrated with a real case example. The case company, here called TechComp, is an original equipment manufacturer for the power generating industry. The product is technically advanced with a long lifetime. Customers are located globally and spans from small publicly owned plants to large sites. Traditional customer segmentation based on industry, size and geographical location is used. The maintenance is mainly performed as overhauls with long intervals, but a number of additional maintenance services are offered, such as condition monitoring, expert support and spare parts management.

TechComp is a traditional engineering company where the product and maintenance is sold separately; the latter in form of medium term contracts, and the different maintenance services offered can be flexibly combined. The maintenance contracts are individually agreed upon for every customer – the number of contract variants is about as many as customers. TechComp has a reputation of providing high quality and reliable products, but as competitors slowly reach the same technical level, TechComp has to find other means for surviving on the market. On the positive side, TechComp is well-known, highly reputed and major actor in this sector. They possess deep technical know-how as well as knowledge regarding the product, its usage and health (condition monitoring technology is already implemented on the product). Thus, TechComp has the ability to develop their maintenance services into combined utility-based offerings. However, the view on service has to change, both internally at TechComp and amongst the customers. Internally, service is today seen as a byproduct of sales of products. Instead, products and services should be seen as parts of a value-generating customer offering. Externally, this is also a challenge, as it means changes in customer behavior. Customers must allow an external part to increase their responsibility of the production and performance, and give them access to business sensitive information. Utility-based business models imply that TechComp exposes themselves to greater business-related risks than today, and this requires mechanisms for risk sharing. Today, cost-plus models are used, which do not involve any sharing of risk (or benefits) between the seller and the customer. Rules and regulations in different countries also represent a challenge. The aftersales administration is today huge. Consequently, the number of customer contract should therefore decrease. Instead, a few offerings are preferred, such as one basic service agreement (resource based) and two holistic agreements (performance or utility based), with few variants, see Table 2.

Service outside of these contracts could be offered, but since it will be a customized solution, pricing will be thereafter. The new offerings are mainly developed for new sales (see Bakshi et al. 2015), but existing service customers can gradually be offered to switch to these new service contracts. The basic service agreement is a low-risk offering because TechComp is in control of and highly knowable in these maintenance activities. The holistic agreements requires repackaging of already available services and clearly selling operating support instead of maintenance. Operating support is proactive (dynamic), while maintenance is preventive or reactive (rigid)! The goal is to ensure continuous and safe operation with as few disturbances as possible. The real technical risk is not as dominant as the business risk; Tech-Comp already has access to certain operating data as well as data about system health. For new systems, data access could be agreed upon, or managed by Tech-Comp taking overall responsibility for the product and its surrounding systems. Profit-sharing mechanisms could be considered, see e.g. Sinkkonen et al. (2013).

Table 2. Proposed modular-based maintenance offerings for TechComp.

Contract	Description	Customer segment
Basic service agreement	Standardized service at fixed intervals during the contract period, and support service. Only preventive activities are regulated at fixed price with relevant price adjustments in time. Spare parts according to the fixed maintenance plan are included. TechComp retains the ownership of spare parts.	Customers lacking maintenance competences (low skill level, small or non-existing maintenance organization, requests traditional preventive maintenance), Power generation non-operational critical, Customers requiring continuous payment model based on fixed price (traditional contract)
Holistic performance agreement	Agreement for a certain operating profile regulated on the system performance. Condition monitoring allows for adjustments in the maintenance plan within a certain time frame (about 3 months) both with respect to timing and scope. TechComp takes a greater business related risk, which must carefully be priced and regulated in the contract.	High maintenance competence and possibilities to allow dynamic scheduling of maintenance, Power generation not core business, but where the core business is affected by power generation, Customers open for data sharing
Holistic utility agreement	Overall power generation agreement regulated on system output. Long term contract form where the full operations is handled by TechComp. The pricing must cover business related risks as well as technology investments. Adjustments in the plan (time/scope) a decision for the supplier within the frame of the holistic contract; the contract only defines required production output.	High maintenance competence and coordinated operations allowing for dynamic planning, Power generation core business and affects customers' customer, Customers open for co-creation, partnering and data sharing

4 Conclusions

Great business potential exists in offering products combined with services, and in offering maintenance based on the customer's real needs. One obstacle is to find suitable ways to package the maintenance offering in a smart way, i.e. that the customer perceive uniqueness and value in the offering while the administration of contracts remains unchanged or becomes more efficient, and without large additional input of resources. Instead, the suppliers' internal resources should be utilized in a more efficient way. The concept of modular-based and dynamic maintenance offerings proposed in this paper is a promising approach to address these problems. Modular maintenance offerings classify maintenance services with increasing integration of the offering, and increasing focus on utility for the customer and the customers' customer while the value proposition is based on the core competencies of the service provider in form of key activities and resources. Future work will include the further development of the modular-based business model to include aspects of contract forms, key performance indicators and follow-up mechanisms. The distribution channels and customer segmentation are also areas of further development. In addition, the sustainability perspective, such as circular economy and environmental impact, should be addressed.

References

Bakshi, N., Kim, S.-H., Savva, N.: Signaling new product reliability with after-sales service contracts. Manag. Sci. **61**(8), 1812–1829 (2015)

Barberá, L., Crespo, A., Viveros, P., Stegmaier, R.: Advanced model for maintenance management in a continuous improvement cycle: integration into the business strategy. Int. J. Syst. Assurance Eng. Manag. **3**(1), 47–63 (2012)

Erguido, A., Crespo Marquez, A., Castellano, E., Gomez Fernandez, J.F.: A dynamic opportunistic maintenance model to maximize energy-based availability while reducing the life cycle cost of wind farms. Renew. Energy **114**, 843–856 (2017)

Ingwald, A., Kans, M.: Modular-based framework of key performance indicators regulating maintenance contracts. In: 12th World Congress on Engineering Asset Management eBook proceedings (2018, in press)

Kans, M., Ingwald, A.: A framework for business model development for reaching service management 4.0. J. Maintenance Eng. **1**, 398–407 (2016)

Kujala, S., Kujala, J., Turkulainen, V., Artto, K., Aaltonen, P., Wikström, K.: Factors influencing the choice of solution-specific business models. Int. J. Project Manag. **29**(8), 960–970 (2011)

Lieckens, K.T., Colen, P.J., Lambrecht, M.R.: Network and contract optimization for maintenance services with remanufacturing. Comput. Oper. Res. **54**, 232–244 (2015)

Mollaverdi, N., Mirabadi, S.: Dynamic maintenance in series – parallel manufacturing systems with consideration of buffer inventory. In: Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management, Dubai, United Arab Emirates (UAE), 3–5 March 2015 (2015)

Ng, I.C.L., Xin Ding, D., Yip, N.: Outcome-based contracts as new business model: the role of partnership and value-driven relational assets. Ind. Market. Manag. **42**(5), 730–743 (2013)

Olve, N.-G., Cöster, M., Iveroth, E., Petri, C.-J., Westelius, A.: Prissättning - Affärsekologier, affärsmodeller, prismodeller. Lund, Studentlitteratur (2013)

Osterwalder, A., Pigneur, Y.: Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers [Electronic Resource]. Wiley, Hoboken (2010)

Shohet, I.M., Straub, A.: Performance-based-maintenance: a comparative study between the Netherlands and Israel. Int. J. Strateg. Prop. Manag. **17**(2), 199–209 (2013)

Sinkkonen, T., Marttonen, S., Tynninen, L., Kärri, T.: Modelling costs in maintenance networks. J. Qual. Maintenance Eng. **19**(3), 330–344 (2013)

Sun, J., Li, L., Xi, L.: Modified two-stage degradation model for dynamic maintenance threshold calculation considering uncertainty. IEEE Trans. Autom. Sci. Eng. **9**(1), 209–212 (2012)

Visnjic Kastalli, I., Van Looy, B.: Servitization: disentangling the impact of service business model innovation on manufacturing firm performance. J. Oper. Manag. **31**(4), 169–180 (2013)

Zhou, B., Yu, J., Shao, J., Trentesaux, D.: Bottleneck-based opportunistic maintenance model for series production systems. J. Qual. Maintenance Eng. **21**(1), 70–88 (2015)

A Risk Indicator in Asset Management to Optimize Maintenance Periods

Vicente González-Prida[1,2(✉)], Jesús Zamora[1,2,3], Antonio Guillén[4], Joel Adams[5], Khairy Kobbacy[6], Carmen Martín[7], Antonio de la Fuente[4], and Adolfo Crespo[4]

[1] University of Seville, Seville, Spain
{jpzb,jpzb}@fsof.uned.es
[2] UNED, Madrid, Spain
[3] Department of Logic and Philosophy of Science, UNED, Madrid, Spain
[4] Department of Industrial Management, University of Seville, Seville, Spain
ajguillen@us.es, antoniodela84@gmail.com,
adolfo@etsi.us.es
[5] Department of Engineering, University of Cambridge, Cambridge, UK
ja579@cam.ac.uk
[6] Formerly with University of Taibah, Medina, Saudi Arabia
kobbacyk@gmail.com
[7] Mechanics, Materials, Structure and Process Division, University of Toulouse, Toulouse, France
Carmen.Martin@enit.fr

Abstract. Different methodologies are nowadays employed to identify failure events in industrial process, allowing the decision makers to choose appropriate technical and organizational safety measures. The treatment of data in order to prevent dangerous events may affect significantly the diverse analyses and is reflected in the results. Quantification risk analysis is therefore one of the most critical areas in asset management (AM) as stated in the ISO 55000. In the same way, intelligent risk management should be one critical challenge of the Industry 4.0, since nowadays and by using new technologies, it is possible to gather large amounts of data extrapolated from the physical assets.

With all the above, this paper is intended to understand uncertainty, trying to reduce the risk of dangerous events by the treatment of big data. Particularly, a time window is obtained showing minimum and maximum thresholds for the best time to apply a preventive maintenance task, together with other interesting statistics.

1 Introduction

In reliability assessments, uncertainty reduces the results' validity as it increases the probability of obtaining inaccurate calculations, making consequently wrong decisions. Among other reasons, risk analysis is concerned with uncertainty related to the outcomes of carrying out some activity that are considered important in a decision-making setting [1]. Therefore, risk becomes relevant for both the system reliability evaluation [2] and the selection of the appropriate maintenance policy [3].

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 566–573, 2020.
https://doi.org/10.1007/978-3-030-48021-9_63

According to the Society for Risk Analysis (SRA) [4], risk analysis is built on two main pillars: (i) risk knowledge related to an activity in the real world and (ii) risk knowledge on concepts, theories, frameworks, approaches, principles, methods and models to understand, assess, characterize, communicate, manage and govern risk. This paper is mainly focused in the second pillar, intended to characterize risk under the new concepts arrived from Industry 4.0. More in particular, the goal here is to understand uncertainty in the risk assessment by using approaches provided by technologies like the big data. As a prerequisite, a proper treatment of uncertainty will be mandatory in order to achieve a broad, informative and balanced picture of risk [5].

The common definition of risk (associated with failure) is the probability that a failure will occur and the consequences of that failure. It is not intended to deal with risk as an expected value, but as a probability of undesirable consequences [1]. According to [6], Risk is basically expressed as follows (i referred to event i):

$$R = \sum_{i=1} Pfi \times Cfi \tag{1}$$

Where:

- R is the risk,
- Pfi is the probability of failure
- Cfi is the consequences of the undesired event.

Thus, the goal of this paper is to review risk assessment under the new scope of technologies from the concept of Industry 4.0. These technologies will be a support to improve the knowledge about uncertainty [7].

2 Risk Indicator in Asset Management

Risk is one of the main aspects in the Asset Management (AM) approach. Reference [8] introduces how the organizations should determine the actions needed for addressing risks for its AM system. An asset's risk is a useful indicator for determining optimal time of repair/replacement for assets in order to yield minimal operational cost of maintenance. For a successful asset management practice, asset intensive organizations must understand the risk profile associated with their asset portfolio and how this will change over time [9]. Unfortunately, in many risk-based asset management approaches, the only thing that is known to change in the risk profile of the asset is the likelihood (or probability) of failure.

The criticality (or failure consequences) of asset is assumed to be fixed and has been considered as, more or less, a static quantity that is not updated with sufficient frequency as the operating environment changes. While addressing risks, the organization should determine the risk assessment criteria within the asset management decision making process.

As seen in Fig. 1, risk is defined as the combination of failure probability and the consequences of failure (criticality). The scenarios below describe what an asset risk profile (incorporating maintenance interventions) will look like when criticality is considered to be static versus when it is considered to be dynamic. Scenario 1 shows the intervention time when change in criticality is not taken into consideration. But from scenario 2, increase in criticality would result in corresponding increase in risk. Therefore, in order to maintain a maximum risk of r2(£), a new intervention time tn has to be adopted instead of the original time t2 in the initial asset management plan. This will result in savings of (rn – r2) £ in the event that the asset fails before t2.

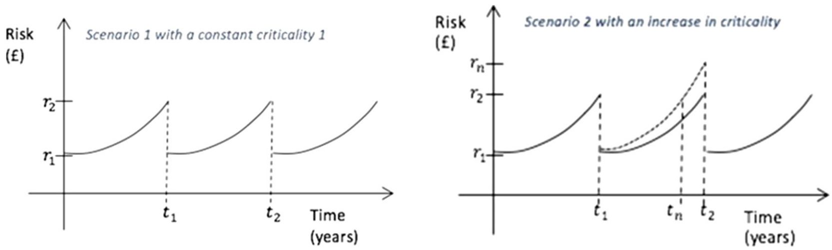


Fig. 1. Scenario 1 with a constant criticality/Scenario 2 with an increase in criticality

It is important to highlight that a true picture of the risk profile of an organization's asset portfolio is required to enhance value, generated from the asset, to the organization. As seen in the scenarios above, there are considerable cost savings to be made by making informed choices on the timing of repairs/replacements that strikes the right balance between value-versus-cost. Another immediate benefit is better risk management, as the organization now have a better picture of the risk and can determine a maximum tolerable risk they can cope with.

Taking this discussion further (considering the oil and gas industry and studying the OREDA data more closely [10]) a rising trend in the subsea infrastructure can be noticed. This is applicable for both Greenfield and Brownfield developments. Maintenance in the subsea sector is very unique and different from an onshore or offshore topside facility. Very expensive ships/rigs are required for installation, inspection, maintenance and pigging activities. However these costs are incurred occasionally and the daily operation cost is relatively low as the daily manning requirements for a subsea installation are limited and could be located in a remote location of more convenience and safety.

The remote locations are able to meet both the control and operational requirements. On close evaluation the OPEX cost for subsea oil and gas facilities reveal huge spikes from their normal low OPEX costs. This is a reflection of the high ship/rig cost associated with subsea maintenance and repair works. However it is critical to know that the cost of the rigs on the spot market varies substantially based on the weather and seasons of operation along with the future operations coming up in the geographical location.

Parameter changes along the asset life cycle impacts directly on its reliability and failure probability and, consequently, in the risk assumed for such a failure [9]. A risk indicator will be needed in maintenance management processes with the objective of preserving the asset operation, maximizing operational performance and economic profitability. With this purpose, references [11] and [12] present a Preventive Replacement by Time model with the objective to quantify the period of maintenance execution time in which the lowest cost is generated (minimum cost per unit of time). The following are the mathematical expressions that allow the calculation of the time period that generates the minimum cost of maintenance of preventive replacement by time (2):

$$\frac{C(ti) = (Cf \times F(ti)) + (Cp \times (1 - F(ti)))}{ti(1 - F(ti)) + \int_0^{ti} \frac{ti \times f(t)}{F(ti)} dt} = \frac{monetary \cdot unit}{time} \tag{2}$$

Where:

- ti: time until the Time To Failure i (TTFi)
- C(ti): average cost of the maintenance task per unit of time (see Fig. 2)

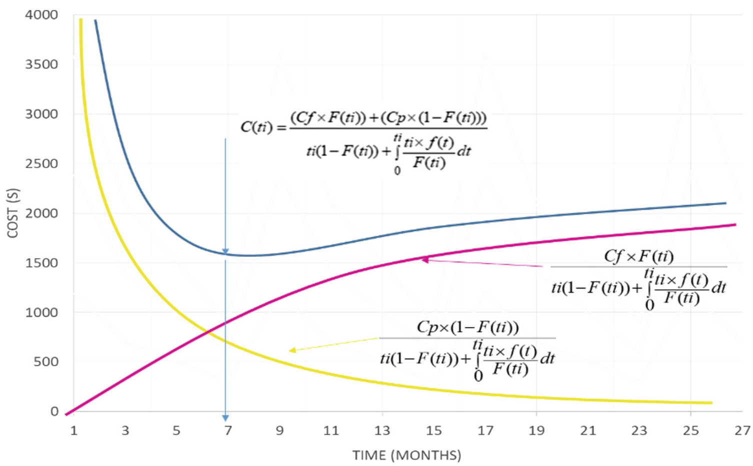


Fig. 2. Minimum expected cost per unit of time for MP at constant age

- F(ti): probability of failure occurrence (%)
- f(ti): failure probability density function
- Cf: costs for corrective maintenance due to failure (materials, labor, lost profits, safety, environment, etc.)
- Cp: costs for preventive/planned maintenance (materials, labor, lost profits, safety, environment, etc.)

The risk indicator (with the above mentioned notation: C(ti)) quantifies the influence of both magnitudes: failure probability and failure consequence, which is useful for maintenance optimization [13]. This indicator is useful to quantify the time for a preventive replacement at a lowest cost per unit of time [14].

3 Connection to Industry 4.0

Due to the current state of scientific and technological development, the transformation of the industry is facing a new challenge: The transition from discrete technological solutions to a global conception where assets, plant, processes and engineering systems are conceived, designed and operated as an integrated complex unit. Reference [7] proposes to develop the transition from current isolated production assets, to a Smart Factory adapting a consolidated Asset Management Framework [15] (integrated into ISO 55000) to the complexity of a "System of Systems" by stages like: identification, prioritization, risk management, planning, scheduling, execution, control, and improvement, supported by system engineering techniques and agile/concurrent project management.

One of these stages refers to the resources optimization where a basic tool is the Risk-Cost Optimization considering also the new opportunities that Cyber-physical Systems bring. The use of self-compare behavior and the machine to machine data exchange will allow taking advantage of dynamic planning and scheduling forecasting with optimal allocation of resources for maintenance and operation. The capacity of Cyber-physical Systems to identify misuse or unbalanced workloads will also affect the production plans or operators training plans. In addition to this, health prognosis and dynamic planning will be possible because of data from new systems. This implementation should embrace the Agile philosophy as an iterative, incremental and evolutionary process, where risk and uncertainty are reducing in each iteration.

Consequently, the appropriate use of available data together with statistics and simulation tools as well as the conjunction of new technologies from Industry 4.0, may provide more valid forecasts. The addition of Big Data and Artificial Intelligent (AI) can go beyond changing decision-making processes to changing entire business models and assets management with winner take-all dynamics and organizations that wait for the Industry 4.0 technologies to settle risk being left behind [13]. The implementation of technologies like Big Data and AI, may establish when necessary, inspections to be automatically scheduled with machine learning algorithms, making a judgment call on whether the deployment of smart devices (e.g. drones or robots) could be sufficient at first, or whether human intervention is needed right away. A key challenge is planning maintenance at system level and addressing groups of components. In fact, smart-based surveillance can replace the time-intensive and risky manual inspection of industrial assets. This reduces assets downtime because inspections by smart devices can happen while assets remain running, helping keep costs down [13].

Nowadays, and regarding the a.m. consolidated Asset Management Framework [15], the use of tools like Life-Cycle Cost analysis for decision making regarding long-term renewal and replacement of assets is quite extended. The disadvantage in many cases is the large amount of variables that must be handled generating scenarios of high uncertainty. This eventuality can be reduced as indicated by the use of Big Data and Data Analytics technologies, improving the estimation accuracy. At that moment, it is where, for instance, a proper definition of the Assets Health and its measure by an index or a risk indicator comes into play as a support tool [16]. This example is also a simple application that justifies the proposed arguments regarding risk and how industry 4.0 may help enhancing the quality of any asset function along its life cycle.

4 Results Discussion and Alternative Modelling Approaches

The authors recognise that in high production systems availability can be of much higher concern as the value of lost production can surpass the maintenance cost. However, in order to adhere to the definition of asset risk, one can overcome this difficulty by adding lost production cost to the breakdown cost. Alternatively, an approach to minimise downtime can be used. In order to illustrate the goal of expressing risk in terms of cost (consequences) linked to system reliability (probabilities), an example has been used considering a Weibull distribution for modelling system reliability and, by assuming different values for its failure rate, it is possible to obtain a better knowledge of the subsequent risk, achieving a greater sensitivity of risk assessment in order to obtain relevant information about the potential costs to maintain the system at a specific time [17]. This risk assessment can also be enhanced with real-time information flows, since machine learning and the treatment of large amounts of data improve the predictive accuracy of defaults, assets degradation, production interruptions connected with maintenance of cost information.

However the authors are aware of alternative approaches that can help in risk modelling though they were mainly developed for preventive maintenance modelling and optimization. Such approaches which have sound mathematical underpinnings make specific assumptions about the nature of the system in hand. Obviously both the numerical and analytical approaches have their pros and cons. As shown here there is mix of analytical and numerical methods in these approaches.

To overcome the shortcomings of the NHPP to model deteriorating repairable systems subject to preventive maintenance, [18] a Proportional Hazards Model (PHM) was developed for preventive maintenance scheduling. A heuristic approach for implementing the semi-parametric PHM to schedule the next preventive maintenance interval on the basis of the equipment full condition history is developed. Two PHMs are fitted for the life following corrective and preventive maintenance using relevant explanatory variables. These models are then used within a simulation framework to schedule the next preventive maintenance interval. Example for applying the model on pumps used in continuous process industry was presented with selected explanatory variables such as age of equipment, average preventive maintenance interval and times since last preventive and corrective work. The major underlying assumption for using PHM is that lives following preventive and corrective maintenance are statistically independent, conditional upon the explanatory variables. While the authors claim that this is a reasonable assumption nonetheless they acknowledged that this assumption requires justification.

Reference [19] uses Cox's Proportional Intensities Model, PIM, (1972) to further develop the model presented by [20]. They modelled the system's failure pattern using a NHPP, linking the corresponding intensity function to a linear predictor which is a function of observable explanatory variables and of the history of the failure pattern. The advantage of this model is its flexibility and it doesn't require the assumption of system renewal following preventive maintenance. However, [19] raised the concern that in their model the optimal preventive maintenance interval is based on a variable planning horizon and hence not necessarily valid for fixed or infinite horizons.

5 Conclusions

Industry 4.0 provides advances in technologies like Big Data treatment or Artificial Intelligent that will enable organizations to leverage rapid growth in the volume of data in order to optimize decision making processes in real time as well as the risk management among other fields. Besides considering a methodology to better decide the scheduling of a replacement activity, taking into account a minimization in maintenance costs for an assumed system, this paper suggests how advanced technologies can improve engineering efficiency, prevent faults, increase safety, etc. by automating risky activities, reducing maintenance costs with better spares supply, demand planning, task scheduling together with many other benefits. The implementation of such technologies becomes an important tool for the success of the maintenance function and, in some cases, it complements methodologies for auditing the resources allocation of critical maintenance and risk prevention activities.

Following the example, alternative modelling approaches to decide on optimal preventive maintenance policy for such deteriorating repairable systems have been discussed. It is concluded that, while the approach presented in this paper has no theoretical underpinning, it also has no restriction on untested assumptions that are usually required in the analytical modelling approaches. Nevertheless, industry decision-makers may apply Big Data and AI tools in order to update and increase the accuracy in the prediction of assets behaviour, with visibility on component availability and risk management.

References

1. Nilsen, T., Aven, T.: Models and model uncertainty in the context of risk analysis. Reliab. Eng. Syst. Saf. **79**(3), 309–317 (2003)
2. Persona, A., Sgarbossa, F., Pham, H.: Systemability function to optimisation reliability in random environment. Int. J. Math. Oper. Res. **1**(3), 397–416 (2009)
3. Persona, A., Pham, H., Sgarbossa, F.: Age replacement policy in a random environment using systemability. Int. J. Syst. Sci. **41**(11), 1383–1397 (2010)
4. Aven, T., Andersen, H.B., Cox, T., López, E., Greenberg, M., Guikema, S., Kröger, W., Renn, O., Zio, E.: Risk Analysis Foundations. Society for Risk Analysis (SRA) (2017)
5. Flage, R., Aven, T., Zio, E., Baraldi, P.: Concerns, challenges, and directions of development for the issue of representing uncertainty in risk assessment. Risk Anal. **34**, 1196–1207 (2014). https://doi.org/10.1111/risa.12247

6. ISO 31000: Risk management. Principles and guidelines. International Standards Organizations (2010)
7. Villar-Fidalgo, L., Crespo Márquez, A., González Prida, V., De la Fuente, A., Martínez-Galán, P., Guillén, A.: Cyber physical systems implementation for asset management improvement: a framework for the transition. In: Haugen, et al. (eds.) Safety and Reliability – Safe Societies in a Changing World, pp. 3063–3069. Taylor & Francis Group, London (2018). ISBN 978-0-8153-8682-7
8. ISO 55002 (2013). Asset management - Management systems – Guidelines for the application of ISO 55001. International Standards Organizations
9. Adams, J., Srinivasan, R., Parlikad, A.K., González-Prida, V., Crespo, A.M.: Towards dynamic criticality-based maintenance strategy for industrial assets. IFAC-PapersOnLine **49** (28), 103–107 (2016). https://doi.org/10.1016/j.ifacol.2016.11.018
10. OREDA Handbook 2015. Offshore and Onshore Reliability Data, 6th edn. Volume 1&2 (2015)
11. Gonzalez-Prida, V., Viveros, P., Crespo, A., Martin, C.: Multi-criteria decision tool applied to a system reliability for the priorization of spare parts. Reliab.: Theory Appl. **9**(#02 (33)) (2014). ISSN 1932-2321
12. Campbell, J., Jardine, A.: Maintenance Excellence. Marcel-Dekker, London (2001)
13. Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Henke, N., Trench, M.: Artificial Intelligence: The Next Digital Frontier? McKinsey Global Institute (2017)
14. Woodhouse, J.: Managing Industrial Risk. Chapman Hill Inc., London (1993)
15. Crespo Márquez, A.: The Maintenance Management Framework: Models and Methods for Complex Systems Maintenance Springer, Heidelberg (2007)
16. De La Fuente, A., González-Prida, V., Guillén, A., Crespo, A., Sola, A., Gómez, J., Moreu, P.: Strategic view of an assets health index for making long-term decisions in different industries. In: Haugen, et al. (eds.) Safety and Reliability – Safe Societies in a Changing World, pp. 1151–1156. Taylor & Francis Group, London (2018). ISBN 978-0-8153-8682-7
17. González-Prida, V., Shambhu, J., Guillen, A., Adams, J., Peres, F., Crespo, A.: An approach to risk quantification based on pseudo-random failure rates. In: 3rd IFAC AMEST on Workshop Maintenance Technologies for Performance Enhancement, 19–21 October 2016 Biarritz, France (2016)
18. Kobbacy, K.A.H., Fawzi, B.B., Percy, D.F., Ascher, H.E.: A full history proportional hazards model for preventive maintenance scheduling. Qual. Reliab. Eng. Int. **13**, 197–198 (1997)
19. Percy, D.F., Kobbacy, K.A.H.: Using proportional-intensities models to schedule preventive-maintenance intervals. IMA. J. Math. Appl. Bus. Ind. **9**, 289–302 (1998)
20. Ascher, H.E., Kobbacy, K.A.H.: Modelling preventive maintenance for deteriorating repairable systems. IMA. J. Math. Appl. Bus. Ind. **6**, 85–99 (1995)

Topology-Based Model of Survivability in Network Utilities

Pablo Martínez-Galán[1(✉)], Juan F. Gómez[1], Adolfo Crespo[1], Antonio de la Fuente[1], Vicente González-Prida[2], and Antonio Guillén[1]

[1] Department of Industrial Management, University of Seville, Seville, Spain
pablomgf93@gmail.com, juan.gomez@iies.es, {adolfo, ajguillen}@us.es, antoniodela84@gmail.com
[2] University of Seville and UNED, Seville, Spain
vicente.gonzalezprida@gdels.com

Abstract. Our society-increased dependence on utilities performance fosters the need for this infrastructures resilience. Especially in networks utilities, costs of incidents can enormously propagate throughout the network, affecting numerous customers and impacting business results, image and reputation. At present, measures of resilience or survivability are crucial within utilities. In this work, we first present a review about existing approaches to tackle network utilities survivability from a holistic point of view; focusing on graph theory principles, a new model is proposed to evaluate this feature. The model allows the comparison of different networks' survivability and may serve as a valuable tool to support networks services management along their life cycle.

1 Introduction

The dependence of humans on the proper functioning of networks has been increasing over the centuries. At the same time as this dependence has been growing, so has the risk and the consequences of the failures. Given the dependence on the networks and the costs related to network failures it is essential the survivability, which is the capability of a system to fulfil its mission, in a timely manner, in the presence of threats such as attacks or large-scale natural disasters. Design of survivable networks is a complex task because current models are simulation models which carry great cost and time. Mainly because of the dependencies that exist inside the networks, which are the ones that have the most effect on cascading failures (Zhou *et al.* 2015). To tackle this, simple models are required in order to help in the decision making of an organization.

The objective of the paper is to find a simple mathematical model that locates the weakest points of the network. Therefore, those that we should reinforce more. The aim of protect these points is to avoid all kinds of threats, and in case of failures, that the network is restructured to continue fulfilling its objective.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 574–582, 2020.
https://doi.org/10.1007/978-3-030-48021-9_64

2 State of the Art

The network survivability is one of the disciplines, which is related to challenge tolerance, which make a network resilient and it has become important over the years. The first articles on network resilience appeared in the mid-1900s (Rohrer and Sterbenz 2011) and are based on fault tolerance, another of the network resilience disciplines. The techniques used to deal with fault tolerance were based on redundancy. Later, with the emergence of new technologies and strong dependencies in the networks, it appeared the necessity to design new techniques that would avoid cascading failures due to dependencies. The first articles on network survivability were written in 1970s. To build a survivable network, diversity is required in addition to the redundancy required by fault tolerance.

Keeping in mind that we want a simple and practical model that does not take much time, we will see through the revised survivability literature the characteristics of the models that most closely match our requirements. Figure 1 shows the path that we have followed to build our model.

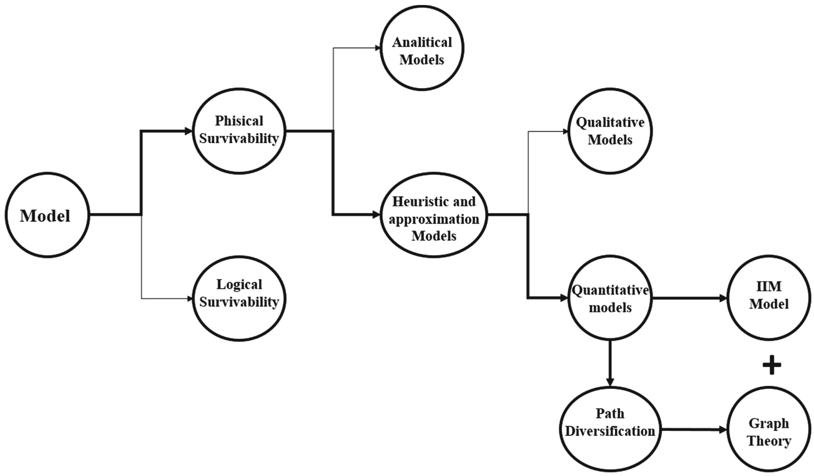


Fig. 1. Different models in survivability networks literature

1. The survivability models can be divided into physical and logical models (Soni et al. 1999). Referring first to the faults that occur physically in the nodes or links of the networks and the second to the capacity they have to carry. In our case we want a model that considers the physical faults of the network. Therefore, we choose a physical survivability model in which we will take into account both the external threats and the degradation of the network components.
2. Secondly, we can divide the survivability models into analytical and approximation models. In the analytical models the time required, and the number of constraints grows astronomically with the size of the network. As an example in (Dharmaraja et al. 2016) the survivability of the network, with respect to reliable hardware and

channel availability is explored using Markov chains and Markov reward model. In contrast, the heuristic and approximation models of recent years have proved to be promising, with good approximations, low cost and fast computational times. In (Feng et al. 2018) they have used Monte Carlo Simulation to enhance the easy propagation of Common-Cause Failures across the complex system, instead of an analytical approach, which currently is impossible.

3. In (Azni *et al.* 2015) another interesting classification is made. They divide the survivability models in quantitative and qualitative. It is also questioned what parameters can measure the survivability networks and what are the limitations of current models. These issues are fundamental when developing our model. We want to build a quantitative model that gives us information of the state of the network. The fact that there are no direct measures for the survivability network implies that we have to find an effective measure to quantify the survivability.
4. In (Rohrer et al. 2009) the path diversification is introduced, a concept closely linked to the survivability network. Based on this concept in (Rohrer and Sterbenz 2011) the authors analyse the measurement of the path diversity as a measure of survivability. The results show that it is a good measure of survivability, above any other property based on topology. We can calculate the diversity in the network using the graph theory (Habibi and Viet 2012). We developed our own algorithm in order to calculate the number of independent paths between a pair of nodes.
5. Once we have the measure of survivability, we will calculate the impact of one node on another to evaluate the survivability of a node depending on the state of the others. The Input-Output Inoperability Model (IIM), calculates the impact between nodes according to the degradation of one node when the other is out of service (Haimes and Jiang 2001). We propose to calculate the impact of one node on another based on the reduction of diversity that one node causes in another.

3 Bases of the Model

We propose a Model based on the Input-Output Inoperability Model (IIM) (Haimes and Jiang 2001) and the Path Diversity (Rohrer et al. 2009).

3.1 IIM Model

The IIM is an evolution of the original Nobel Prize Leontief input-output model (Leontief 1951). It focuses on the spread of inoperability among the network. They introduce the term of 'inoperability' as degradation of the normal production. The IIM is oriented towards undesired failures. Instead, we want to build a model that not only considers undesired events, but also the risk of failure of each node. The inoperability corresponding to a perturbation c is given by (Setola et al. 2009):

$$x = (I - A)^{-1} \cdot c = S \cdot c \tag{1}$$

Where x and c are vectors representing, respectively, the inoperability level and the reduction in the operability associated to the n examined nodes. A is a $n \times n$ matrix called influence matrix. It models the direct effects due to first-order dependencies between the nodes. The coefficients a_{01} represent the inoperability of node i when node j is completely inoperable. Matrix S also takes into account, the amplifications introduced by second-order and higher-order dependencies.

In (Setola and De Porcellinis 2008) the influence matrix is calculated using the knowledge provided by the experts and technicians. We tried to find a numerical influence matrix based on the network properties to use the IIM, but this matrix is very restrictive and only worked in small networks. In addition, as the size increased, it caused more problems. For all cases to converge the IIM model, it is necessary that the sum of the columns is less than 1, that is:

$$\sum\nolimits_{i=1}^{n} a_{ij} < 1 \tag{2}$$

When dealing with small networks (<5–10 nodes) this is not strictly necessary, but for larger network sizes (>10 nodes), it becomes an essential requirement for the model to work. After that, we focused on finding the matrix S, in which the consequences of second and higher dependencies were included. We found a good solution based on topology. We propose a matrix S based on the Path diversity.

3.2 Path Diversity

We propose to use the Path Diversity to calculate the number of independent paths between a node-pair. The more independent paths arriving at a node, the more secure it will be. A path P is a vector containing all links and intermediate nodes traversed by the path. The length of this path P is the combined total number of elements. The Path Diversity for two arbitrary paths P_a and P_b is defined as:

$$D(P_a, P_b) = 1 - \frac{|P_b \cap P_a|}{|P_a|} \tag{3}$$

where $|P_a| \leq |P_b|$. Notice that as consequence $D(P_a, P_b) = D(P_b, P_a)$. The Path Diversity has a value of 1 if the paths are completely disjoint and a value of 0 if they are the same path.

3.3 Proposed Model

Once we have seen the bases on which the model is based and the result we want to reach, we show the steps of the model. The programming of the model has been carried out in Wolfram Mathematica. Starting from all the existing paths between two nodes, the first step is to calculate the number of independent paths between each pair of nodes, based on the path diversity explained above. Next, we show the process:

1. Create the Path Diversity Matrix (*PD*), in which we represent the number of independent paths between each pair of nodes. It is a symmetric matrix because the

number of independent paths from node A to B is equal to the number of independent paths from node B to node A. As we said before this is because we work with undirected graphs. we calculate the number of independent paths between node i and j as:

$$PD_{ij} = IP + \sum\nolimits_{k=1}^{IP} \frac{D^k_{max}}{IP} \tag{4}$$

where IP is the integer number of independent paths between nodes i and j and D^k_{max} which is a plus of diversity due to the paths dependent on the aforementioned.

2. After calculating the Path Diversity Matrix (PD) we have a matrix that indicates the level of connectivity between nodes, but we have no information on the impact that each node causes on the others.
3. For the calculation of the impact that a node provokes on the others we remove one node and we look at the effect that it causes in the network. We measure the impact as the reduction in the number of independent paths in each node. We can calculate the percentage in which independent paths has been reduced as:

$$IM_{ij} = \frac{PD_i - PD^j{}_i}{PD_i} \tag{5}$$

Where PD_i is the total number of independent paths arriving at a node i. It is calculated as the sum of the rows or columns of the PD matrix. $PD^j{}_i$ is the total number of independent paths arriving at node i when node j is completely inoperable. IM_{ij} means the percentage of independent paths lost by the node i when the node j is totally inoperable.

Doing this for each node, we get the percentage in which the number of independent paths that reaches each node is reduced when node j is completely inoperable. Therefore, we have the j-th column of the $n \times n$ impact matrix denoted by IM_{ij}. We perform the same process with the rest of nodes and complete the impact matrix.

The impact matrix has a high redundancy in its rows because independent paths are composed of many nodes. In the case of working only with the impact that one node produces in the rest of the network, the matrix could be included in the model and there would be no redundancy. In contrast, if we eliminate two or more nodes from the same independent path, we are eliminating the independent path several times and therefore increasing the impact. In this second case we would have to normalize the matrix by rows as shown in Eq. 6. Based on the number of degraded nodes and their degradation $F_i\,t$, we will combine both matrices to calculate the impact of one node i on another j.

$$IM'_{ij} = \frac{IM_{ij}}{\sum_{j=1}^{j=n} IM_{ij}} \tag{6}$$

We have calculated the coefficients of the impact matrix as shown in (8) and (9).

$$If\ F_i(t) \le \overline{F(t)},\ IM''_{ij} = IM'_{ij} \tag{7}$$

$$If\ F_i(t) > \overline{F(t)},\ IM''_{ij} = \left(F_i(t) - \overline{F(t)}\right) \cdot IM_{ij} \cdot FC + \overline{F(t)} \cdot IM'_{ij} \tag{8}$$

4. The model can be expressed as:

$$x(t) = IM'' \cdot (F(t) + e(t)) \tag{9}$$

where $F(t)$ represents the cumulative density function, which indicates the degradation of each node and e represents the impact on each node due to external failures or perturbations.

5. Finally, we calculate the risk as the number independent paths remaining (PR):

$$PR_i(t) = PD_i \cdot (1 - x_i) \tag{10}$$

After calculating the PR of all nodes, the most critical node will be the one with the lowest value.

3.4 Study Case

Once the model has been explained, it is applied to the network of Fig. 2 to calculate the impact matrix. It is a network with 17 nodes distributed differently. There are more centralized nodes such as 6, other intermediate nodes such as 5, 11 or 7 and others in the periphery such as 9 or 17. It is assumed that in the initial moment the degradation of the nodes is zero. In the successive points the different stages of the model are applied.

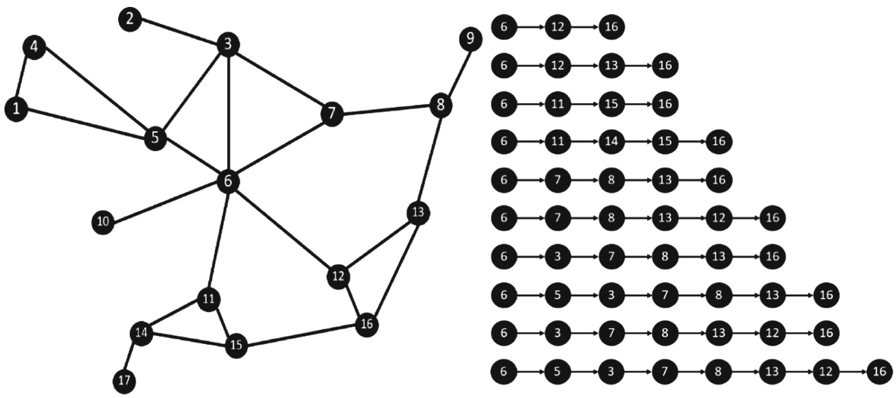


Fig. 2. Study case and all existing paths between nodes 6 and 16.

1. The first step of the model is to calculate the Path Diversity Matrix (*PD*). As an example, the number of independent paths between node 6 and 16 will be calculated, and for the rest of the pair of nodes it will be done in the same way.

First, we calculate the path diversity between all the existing paths between nodes 6 and 16 through Eq. 3. Second, we calculate the shorter independent paths, in this case $P_)$, $P_[$ and $_\backslash$ of Fig. 3.

After obtaining this set of paths, the rest of the paths will be grouped with the independent paths with which their diversity is lower, that is, with those that have more elements in common. The paths are grouped as follows, $P_)$ and $P_]$ the first group, $P_[$ and $P_\wedge$ the second group and the rest of the paths form the third group. The maximum diversity existing between one path and the shortest path of group k is taken as D^I_{J9K}. Logically, the number of groups is equal to the number of independent paths. Finally, calculate the number of independent paths between 6 and 16 as:

$$PD_{6-16} = IP + \sum_{k=1}^{k=IP} \frac{D^k_{max}}{IP} = 3 + \frac{0.33}{3} + \frac{0.2}{3} + \frac{0.28}{3} = 3.27$$

Doing the same with the rest of pairs of nodes we obtain the path diversity matrix, from which we will calculate the impact matrix.

2. To obtain the impact matrix, we eliminate one node and calculate the Path Diversity matrix again. Once we have done this we apply the Eq. 6 and we calculate the coefficients of one column of the impact matrix. We do the same with all nodes of the network and we obtain the complete impact matrix. In Table 1 a part of the impact matrix is shown. Each coefficient represents the reduction of path diversity in node i when node j is completely inoperable.

Table 1. Impact matrix (%)

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	100	5	21	21	97	38	10	14	6	5	9	7	11	13	11	11	6
2	8	100	100	8	22	41	12	17	6	6	10	7	14	14	12	12	6
3	6	3	100	6	21	57	29	28	5	4	17	15	26	13	19	21	5
4	21	5	21	100	97	38	10	14	6	5	9	7	11	13	11	11	6
5	8	4	39	8	100	54	26	27	5	4	16	13	24	12	19	22	5
6	6	4	23	6	19	100	25	24	5	3	21	18	25	13	25	28	4
7	6	4	23	6	19	57	100	30	4	5	18	15	30	13	18	20	5
8	6	5	20	6	19	54	37	100	3	5	19	16	39	13	18	21	5
9	7	6	18	7	21	41	21	100	100	6	13	9	24	13	12	13	6
10	8	6	20	8	21	100	15	19	6	100	12	8	16	14	16	17	6
11	6	5	18	6	18	45	20	24	5	4	100	16	23	15	41	37	4
12	5	5	18	5	17	51	27	28	4	5	26	100	30	14	25	30	4
13	5	4	18	5	18	54	28	28	4	5	26	23	100	13	23	29	5
14	6	5	18	6	18	45	19	22	5	5	43	14	22	100	44	33	3
15	6	5	17	6	18	51	20	23	5	5	40	15	22	15	100	34	4
16	5	5	17	5	17	53	25	27	4	5	27	22	28	13	24	100	4
17	7	6	16	7	21	33	11	16	6	6	25	7	11	100	28	20	100
ρ_i	116	76	408	116	462	811	334	440	79	76	331	210	356	299	345	357	77

The sum by columns of the impact matrix, which represents the overall impact of node j on the network, is shown at the bottom of the Table 1. The impact matrix allows us to know those nodes whose degradation will provoke a greater impact on the network (node 6) and those whose impact on the network is almost irrelevant (nodes 2, 9, 10, 17). Therefore is a very valuable information for the study of the network survivability.

4 Conclusions

This section presents the main conclusions obtained, as well as the possible future research lines that arise as a result of the work carried out. The presented model is applicable to any type and size of network.

1. The model faces the lack of heuristic models for the calculation of network survivability. These models allow you to quickly and easily calculate the points with the greatest risk in the network.
2. Although holistic models generally work with economic or empirical data, such as historical data or expert opinions, this model uses topology properties. In this way, subjectivity is avoided, creating a model with good approximation and with the greatest possible objectivity.

In addition to the model presented, with which a good approximation is achieved in any type of network, several lines of research are still open:

1. A new model in which not only the degradation of the nodes is taken into account, but also the degradation of the links.
2. Obtaining impact matrices not only based on the topology of the network, but also on the capacity that nodes and links can accommodate.
3. Adaptation of the model to the different case studies in an experimental manner. Due to the great casuistry that exists in networks, the possibility of generating an experimental model that is a function of several properties of the network is being studied.

References

Azni, A.H., et al.: Systematic review for network survivability analysis in MANETS. Procedia Soc. Behav. Sci. **195**, 1872–1881 (2015). https://doi.org/10.1016/j.sbspro.2015.06.424

Bhandari, R.: Optimal physical diversity algorithms and survivable networks. In: Proceedings Second IEEE Symposium on Computer and Communications, pp. 433–441 (1997). https://doi.org/10.1109/iscc.1997.616037

Dharmaraja, S., Vinayak, R., Trivedi, K.S.: Reliability and survivability of vehicular ad hoc networks: an analytical approach. Reliab. Eng. Syst. Saf. **153**, 28–38 (2016). https://doi.org/10.1016/j.ress.2016.04.004

Feng, G., George-Williams, H., Patelli, E., Coolen, F.P.A., Beer, M.: An efficient reliability analysis on complex non-repairable systems with common-cause failures. In: Safety and Reliability – Safe Societies in a Changing World, pp. 2531–2537 (2018)

Habibi, D., Viet, Q.: Graph theory for survivability design in communication networks. New Front. Graph Theory 421–434 (2012). https://doi.org/10.5772/36450

Haimes, Y.Y., Jiang, P.: Leontief-based model of risk in complex interconnected infrastructures. J. Infrastruct. Syst. **7**(1), 1–12 (2001)

Leontief, W.: The structure of American economy, 1919–1939: an empirical application of equilibrium analysis (1951)

Rohrer, J.P., Jabbar, A., Sterbenz, J.P.G.: Path diversification: a multipath resilience mechanism. In: Proceedings of the 2009 7th International Workshop on the Design of Reliable Communication Networks, DRCN 2009, pp. 343–351 (2009). https://doi.org/10.1109/drcn.2009.5339988

Rohrer, J.P., Sterbenz, J.P.G.: Predicting topology survivability using path diversity. In: 2011 3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), pp. 1–7 (2011)

Setola, R., De Porcellinis, S.: A methodology to estimate input-output inoperability model parameters. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). LNCS, vol. 5141, pp. 149–160 (2008). https://doi.org/10.1007/978-3-540-89173-4_13

Setola, R., De Porcellinis, S., Sforna, M.: Critical infrastructure dependency assessment using the input-output inoperability model. Int. J. Crit. Infrastruct. Prot. **2**(4), 170–178 (2009). https://doi.org/10.1016/j.ijcip.2009.09.002

Soni, S., Gupta, R., Pirkul, H.: Survivable network design: the state of the art. Inf. Syst. Front. **1**(3), 303–315 (1999). http://www.scopus.com/inward/record.url?eid=2-s2.0-0033268277&partnerID=tZOtx3y1

Zhou, J., et al.: A new model of network cascading failures with dependent nodes. In: Proceedings of Annual Reliability and Maintainability Symposium, May 2015 (2015). https://doi.org/10.1109/rams.2015.7105077

Degradation Management at Nuclear Power Plants

John Einar Hulsund(✉) and Miki Sirola

Institute for Energy Technology, Kjeller, Norway
{John.Einar.Hulsund,Miki.Sirola}@ife.no

Abstract. Maintenance at nuclear power plants, and in many other industries, is usually performed by time-based schedules or after specifications from equipment manufacturers and by monitoring the condition of equipment. Research at IFE addresses condition-based maintenance with prognostic estimation of the degradation of process components and calculation of remaining useful life (RUL) based on measurements and selection of health indicators for the components. Prognostic models have been developed at IFE through case studies for air and sea filters at nuclear power plants and choke valves from offshore industry. Currently many of the nuclear power plant worldwide are reaching the lifetime they were designed and licensed for and many of them have started a process to extend the operational lifetime of these plants. In this research equipment health indicators from many various sources will be combined to analyse the condition of components to estimate the future condition of the plant using a combination of machine learning and physical prognostic models.

1 Introduction

In condition-based maintenance (CBM), the health condition of equipment decides when to carry out maintenance, as opposed to time-based maintenance carried out at regular intervals. A proactive maintenance approach helps in eliminating failures of equipment.

CBM is a decision-making strategy based on real-time diagnosis of failures and prognosis of future equipment health (Peng et al. 2010). Prognostic methodologies have entered the literature later, but it has become more and more as a focus field as well. A step towards prognostic maintenance has been taken (Peng et al. 2010). Prognostic models can be classified into such categories such as physical model, knowledge-based model, data-driven model and combinational model including features of both physical and data-driven models.

Large amounts of data stored over many years can be used to estimate the longterm degradation development of a plant. To handle the large, and sometimes diverse and complex data sets, new methods like distributed computing and machine learning is used. Our approach is to collect health indicators, like maintenance logs, technical reports, measurement, and to calculate the frequency of process conditions, transients and incidents from process data using machine learning. This report summarizes methods earlier used for prognostics that will be combined with methods based on

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 583–590, 2020.
https://doi.org/10.1007/978-3-030-48021-9_65

machine learning to detect failures and transients at a plant. It is not always easy to get data from real plant with transients and failures, and we have therefore used a simulator to produce some of the failure data.

2 Related Work

We have collected some literature about machine learning methods to classify events in nuclear power plants component degradation management. CBM recommends maintenance decisions based on information collected in condition monitoring (Jardine et al. 2006). The three main steps are data acquisition, data processing and maintenance decision making. There are two important aspects in a maintenance program: diagnostics and prognostics. The literature in diagnostics is much larger than the literature in prognostics, which is much smaller area (Jardine et al. 2006).

Fault detection and diagnosis methods in nuclear power plants have been reviewed in (Ma and Jiang 2011). An elder review of artificial intelligence methods in detection and identification of component faults in nuclear power plants (Reifman 1997) collects methodologies in this field including knowledge, reasoning, inference mechanisms and diagnostic approach.

Nuclear component degradation is identified by time-frequency ridge pattern in (Park 2006). Time-frequency analysis (TFA) is applied for identification of operational status of various nuclear power plant components. TFA is used especially in analysis of vibration signals in pipes where some chemical corrosion may occur.

Artificial intelligence and machine learning has been used a lot in the transient analysis of nuclear power plants, see e.g. (Na 2004), (Medeiros and Schirru 2008), (Santosh et al. 2007), (Roverso 2000a; b), also on earlier times (Bartal et al. 1995), (Cheon and Chang 1993). Reference (Bartal et al. 1995) discuss about proximity measure and Multilayer perceptron (MLP) classifier. In (Cheon and Chang 1993) backpropagation network (BPN) is applied in connectionist expert system to identify transients in nuclear power plants.

In (Na 2004) probabilistic neural network (PNN) has been used in classification of severe accident progression scenarios, where the initiating event can be such as loss of coolant accident (LOCA), total loss of feedwater (TLOSFW), station blackout (SBO) or steam generator tube rupture (SGTR). Also, a fuzzy neural network (FNN) is used in this study.

Particle swarm optimization algorithm (PSO) is used in (Medeiros and Schirru 2008) for optimizing a distance-based discrimination transient classification method. Also, such artificial intelligence techniques such as expert systems, neuro-fuzzy and genetic algorithms are commonly used. Four presented approaches in (Roverso 2000a, b) are radial basis function (RFB) neural networks and cascade-RBF, self-organizing map neural networks and recurrent neural networks.

Machine learning algorithms have been used in data classification in nuclear power plants e.g. in (Rocco and Zio 2007) and (Roverso 2000a, b). In (Rocco and Zio 2007) support vector machine (SVM) approach to classification of transients in nuclear power plants is used. In (Roverso 2000a, b) combined use of wavelets and recurrent neural networks improves solution to transient classification problem. Reference (Dietterich

1998) reviews approximate statistical tests and in (MacKay 1992) Bayesian ideas for supervised adaptive classifiers are presented. Regression models are used as well.

Machine learning has also been used in accident scenario analysis, see (Na 2008), and in risk analysis (Guikema 2009) (Rasmussen 1997). Support vector machines (SVM) are used to identify the break location in LOCA and predict the break size using support vector classification (SVC) and support vector regression (SVR). A genetic algorithm is used in optimization.

3 Prognostic Methods and Models for Degradation Managements

A paper from 2014 (Saarela et al. 2014) describes methods to create prognostic models for air filters for the exhaust ventilation air from nuclear power plants. The data contains measurements of the differential pressure over the air filters for air originating at two different locations in the reactor building. One of the locations is the reactor hall, and the other is a laboratory for radiation experiments. The remaining of this chapter describes analysis from the laboratory data.

The differential pressure **dP(t)** is modelled as an aggregation of three phenomena, each occurring at different time scales. The modelled phenomena are:

1. Gradual accumulation of aerosols. This is the normal behaviour due to normal operation conditions.
2. Sporadic aerosol emissions due to, e.g., some maintenance operations are modelled as stepwise increments in the differential pressure.
3. Seasonal variation caused by changes in relative air humidity. This was applied to only the data for differential pressure for air from the laboratory, since the laboratory had an unsealed door where unfiltered humid air could enter into the room.

The data were quite noisy due to measurement uncertainties and where filtered using a median filter. The increasing differential pressures over the air filters were modelled as a gamma process (van Noortwijk 2009). The threshold for when to change filters was set to the date where the filters were actually replaced. Figure 1 shows the estimated differential pressure development **v(t)** and where the estimation crosses the threshold for when the filters should be replaced.

The models predicts that air filters should be changed $\sim$400 days in advance.

4 Machine Learning Methods to Classify Events at NPPs

There is a comprehensive set of methodologies available in statistical learning (Hastie et al. 2008) that can be utilized in the classification of data. The basic elements of supervised learning, linear methods for regression, linear methods for classification, and model assessment and selection are to be studied first.

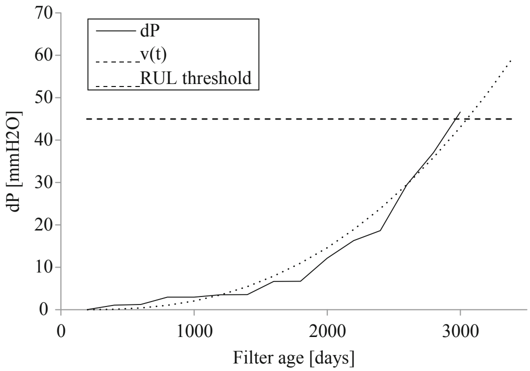


Fig. 1. Estimated deterioration function, v(t), for the differential pressure over the air filters at the laboratory.

4.1 Methods Used in Case Examples

In our test and case examples we have used the following methodologies. In clustering the data without any prior knowledge about it we have used K-means clustering (Hartigan and Wong 1979). It seems to work nicely in two-dimensional and three-dimensional scatter plots in almost all our examples. The number of clusters in each case is defined by selecting a suitable number according to the data shape in the scatter plots, so that the classification looks as natural as possible. It is also possible to use e.g. “elbow method” is selecting the number of clusters (Ketchen and Shook 1996).

In classification of the data by using prior knowledge of a training set we have used several classification methods. First the training set has been classified with various criteria, for example by using K-means clustering. Then according to a fit produced by the training set we have predicted the corresponding classification of the test set by using Nearest-neighbourhood (KNN) classification (Altman 1992) (Cover and Hart 1967), Support vector machine (SVM) classification (Cortes and Vapnik 1995), or Multilayer perceptron (MLP) classification (Rosenblatt 1961). Also, Self-organizing map (SOM) classification (Kohonen 1984) has been tried out.

In dimensionality reduction we have used Principal component analysis (PCA) method (Jolliffe 2002) (Abdi and Williams 2010). The PCA methods makes a projection to desired amount of fewer dimensions by using biggest variance as a criterion. We have also made some visualizations from basic statistical properties (Croxton et al. 1968).

In realization we have used the Jupyter tool as programming environment. All code is written in Python 3 programming language. The data is from Loviisa nuclear power plant and HAMBO simulator (simulating Forsmark nuclear power plant) of OECD Halden reactor project. We have also degradation data including water chemistry variables from the Halden reactor, but we have not analysed it yet.

4.2 Examples of Identifying Failure Data

One goal in data analysis is to detect anomalies in the data. We try to classify the data so that the anomalies are clearly separating from the normal data. Here we call the part of data including anomalies failure data.

We have done clustering and classification of data and tried to identify failure data from normal data. Here are some examples with Loviisa nuclear power plant data and HAMBO simulator data. In the HAMBO simulator the failure is a simulated air leak in the turbine section. In Loviisa power plant there is no real failure, but in the end of data set there is beginning of a slow shutdown for the next outage period where the electrical power is decreasing, which we have used as a 'simulated failure' in our analysis for this example. We have used also PCA analysis in differentiating the normal data and failure data.

In Fig. 2 a scatter plot of two variables, a temperature and a flow in the main condensate system of the Loviisa plant (Unit 1), is presented. The data is normalized and shuffled before dividing it into a training set and a test set. 100 samples from the total 392 have been taken to the test set and the rest of the data is left to the training set. The data representing the failure mode is clearly separated to one cluster pointed with an arrow in the figure.

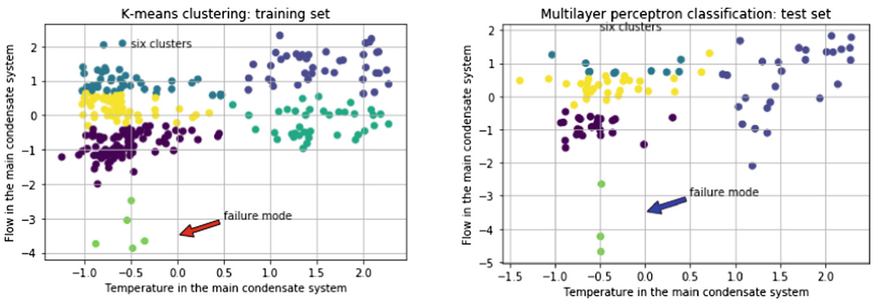


Fig. 2. K-means clustering of the training set and multilayer perceptron classification of the test set. The data is divided in six clusters in both cases. The cluster representing the failure mode is coloured with light green colour.

In the HAMBO simulator an air leak scenario is analysed, see Fig. 3. A scatter plot of three variables, a pump mass flow, a pipe mass flow and a temperature in the turbine section of Forsmark plant (Unit 3), is presented. The data is shuffled and normalized before dividing it into training set and test set. About 18% of the data is used as the test data in the analysis and the rest is left to the training data. As the normal data is much more stable than the transient data, it is clearly clustered to a rather small area in the three-dimensional space pointed with an arrow in the figure.

Clustering data gives information and illustrates the data structure of the data set. PCA analysis is also a good method expressing information about the data structure, but it has not been used in this example. In Fig. 2 the failure data is clearly separable into an own cluster, which is an interesting result. In Fig. 3 the simulator air leak

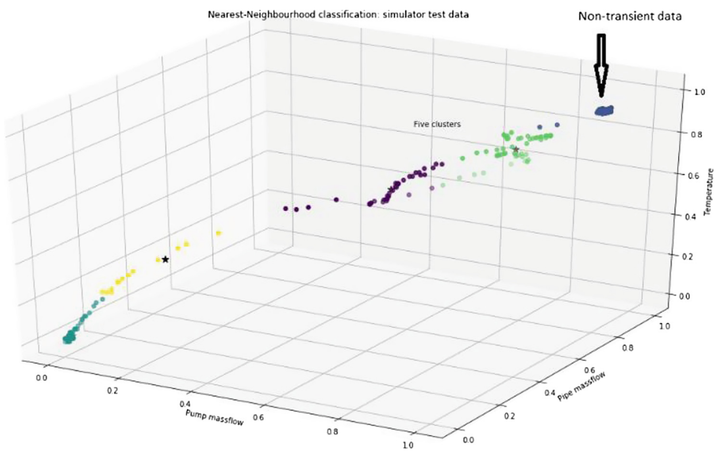


Fig. 3. Nearest neighbourhood classification of the test set, where the training set is clustered into five clusters with K-means clustering algorithm. The test set has the same number of classes. The non-transient data cluster is here quite clearly separated and coloured with blue colour. All the other clusters represent different phases in the transient caused by the failure.

constitutes most of the whole data structure and the data is clustered into different phases of the scenario. As there is not much happening in the normal data before the leak, this part clearly constitutes one cluster of dense data in a small range.

We have presented two cases to detect anomalies in the data by clustering and classification. These anomalies are identified as transients in some critical physical variables. Next step is to dig into the long-term degradation issue and try to find out how the aging of nuclear power plant components can be identified as anomalies in the data. What kind of variables are interesting in this respect and how the behaviour of the process changes when certain components are approaching their end of the remaining useful lifetime? We are looking for answers in for instance this kind of questions.

5 Conclusion

Long term degradation management in nuclear power plants project aims in improvements in condition-based maintenance and better estimations of remaining useful lifetime of the plant components. We have shortly summarized in a literature review the current state-of-the-art. We have presented some earlier developed prognostic methods and models for degradation management, and reviewed machine learning methods and tools to classify events in nuclear power plants.

Examples of separating failure data from normal data with data analysis methods have been presented. Our goal here is to link the theoretical and analytical framework and the practical data mining to find specific decision settings including suitable criteria and characteristics. By following frequency of failures and faults we aim to constitute a 'health-index' that could be used to predict the development of the component security and reliability, as well as in estimations of replacement needs.

References

Abdi, H., Williams, L.J.: Principal component analysis. Wiley Interdisc. Rev. Comput. Stat. **2**(4), 433–459 (2010)

Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. **46**(3), 175–185 (1992)

Bartal, Y., Lin, J., Uhrig, R.: Nuclear power plant transient diagnostics using artificial neural networks that allow “don’t know” classifications. Nucl. Technol. **110**(3), 436–449 (1995)

Cheon, S., Chang, S.: Application of neural networks to a connectionist expert system for transient identification in nuclear power plants. Nucl. Technol. **102**(2), 177–191 (1993)

Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)

Cover, T.M., Hart, P.E.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theory **13**(1), 21–27 (1967)

Croxton, F.E., Cowden, D.J., Klein, S.: Applied general statistics. Teoksessa. Pitman, s.l. p. 625 (1968)

Dietterich, T.: Approximate statistical tests comparing supervised classification learning algorithms. Neural Comput. **10**(7), 1895–1923 (1998)

Guikema, S.: Natural disaster risk analysis for critical infrastructure systems: an approach based on statistical learning theory. Reliab. Eng. Syst. Saf. **94**(4), 855–860 (2009)

Hartigan, J.A., Wong, M.A.: Algorithm AS 136: a k-means clustering algorithm. J. Roy. Stat. Soc.: Ser. C (Appl. Stat.) **28**(1), 100–108 (1979)

Hastie, T., Tibshirani, R., Jerome, F.: The Elements of Statistical Learning – Data Mining, Inference and Prediction, 2nd ed. Springer, s.l. (2008)

Jardine, A., Lin, D., Banjevic, D.: A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech. Syst. Sig. Process. **20**(7), 1483–1510 (2006)

Jolliffe, I.T.: Principal Component Analysis, Series. Teoksessa: Springer Series in Statistics, p. 28. Springer, New York (2002)

Ketchen, D.J.J., Shook, C.L.: The application of cluster analysis in strategic management research: an analysis and critique. Strateg. Manag. J. **17**(6), 441–458 (1996)

Kohonen, T.: Self-Organization and Associative Memory. Springer, Berlin (1984)

MacKay, D.: The evidence framework to applied classification networks. Neural Comput. **4**(5), 720–736 (1992)

Ma, J., Jiang, J.: Applications of fault detection and diagnosis methods in nuclear power plants: a review. Prog. Nucl. Energy **53**(3), 255–266 (2011)

Medeiros, J., Schirru, R.: Identification of nuclear power plant transients using the particle swarm optimization algorithm. Ann. Nucl. Energy **35**(4), 576–582 (2008)

Na, M.: Prediction of major transient scenarios for severe accidents in nuclear power plants (2004). https://doi.org/10.1109/TNS.2004.825090

Na, M.: Detection and diagnostics of loss of coolant accidents using support vector machines (2008). https://doi.org/10.1109/TNS.2007.911136

Peng, Y., Dong, M., Zuo, M.: Current status of machine prognostics in condition-based maintenance: a review. Int. J. Adv. Manuf. Technol. **50**(1–4), 297–313 (2010)

Rasmussen, J.: Risk management in a dynamic society: a modelling problem. Saf. Sci. **27**(2–3), 183–213 (1997)

Reifman, J.: Survey of artificial intelligence methods for detection and identification of component faults in nuclear power plants. Nucl. Technol. **119**(1), 76–97 (1997)

Rocco, C., Zio, E.: A support vector machine integrated system for the classification of operation anomalies in nuclear components and systems. Reliab. Eng. Syst. Saf. **92**(5), 593–600 (2007)

Rosenblatt, F.: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, Washington DC (1961)

Roverso, D.: Multivariate temporal classification by windowed wavelet decomposition and recurrent neural networks. In: International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human-Machine Interface Technologies (NPIC&HMIT 2000), Washinton, USA (2000a)

Roverso, D.: Soft computing tools for transient classification. Inf. Sci. **127**(3–4), 137–156 (2000b)

Saarela, O., Hulsund, J.E., Taipale, A., Hegle, M.: Remaining useful life estimation for air filters at a nuclear power plant. In: 2014 Second European Conference of the Prognostics and Health Management Society, Nantes, France (2014)

Santosh, T.: Application of artificial neural networks to nuclear power plant transient diagnosis. Reliab. Eng. Syst. Saf. **92**(10), 1468–1472 (2007)

van Noortwijk, J.M.: A survey of the application of gamma processes in maintenance. Reliab. Eng. Syst. Saf. **94**(1), 2–21 (2009)

The Dynamic Risk Simulator – A Decision Support Tool for Rotating Equipment Integrity Management

Shyam Krishna Ramanan[1(✉)], Jawad Raza[2], and Jayantha Prasanna Liyanage[3]

[1] University of Stavanger, Stavanger, Norway
shyam.krishna89@gmail.com
[2] Apply Sørco AS, Stavanger, Norway
[3] Center for Industrial Asset Management, University of Stavanger, Stavanger, Norway

Abstract. Asset operation and maintenance has over the last few years become a dynamic and a continuous optimization process in offshore oil & gas industry. It has always been challenging in visualizing availability risks associated with some equipment in such a dynamic environment within finite time horizon due to various practical reasons. An application was therefore developed to model and simulate the lifetime of single-unit repairable systems subjected to age-related failures. The basis of the application, the Dynamic Risk Simulator (DRS), comes from proven statistical methods as well as practical input from maintenance engineers and from engineering processes at Apply Sørco AS. The application enables visualizing equipment availability risks dynamically when varying parameters within selected preventive maintenance strategies by simulating the life of the asset in finite time horizon taking into account its age, preventive maintenance strategy and other factors which reduce/extend its useful life. The application also provides quantitative estimates on the cost of maintenance, optimum maintenance interval and equipment availability in finite time horizon. Two principal maintenance strategies were considered in the modelling namely, the minimal-repair periodic-overhaul strategy and the run-to-failure strategy. The purpose of such an application is to support decisions in a relatively non-complicated manner in appraising opportunities related to reviewing and updating existing maintenance strategies. The application was partially validated with data from an offshore asset. The limitation of DRS is that it cannot replace human judgement with regards to taking the final call on whether or not to postpone maintenance. DRS provides quantitative and qualitative results and is reliant on the experience and insight of industry experts to take the most appropriate course of action based on the provided input.

Keywords: Asset management · Simulation · Maintenance · Optimisation · Cost estimation

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 591–599, 2020.
https://doi.org/10.1007/978-3-030-48021-9_66

1 Introduction

For repairable systems, there are several maintenance strategies developed over the years to cater to a wide variety of boundary conditions from budget to equipment availability to reducing down time and maintenance costs. These are included in the works of Barlow and Hunter (1960), Malik (1979), Wang (2002) and Pham (2003) among others. In this paper, the lifetime of a single-unit repairable system is modelled, simulated and partially validated for two types of rotating equipment – a pump and a motor, subjected to age-related failures. A generic simulation tool, the Dynamic Risk Simulator, is described and has been used to qualitatively visualise the financial, safety and environmental risks associated with varying periodic preventive maintenance intervals for the pump and the motor for variants of two maintenance strategies - the minimal-repair-periodic-overhaul strategy and the run-to-failure strategy. The purpose of the Dynamic Risk Simulator tool is to serve as decision-support to help the decision maker to value his/her options on proceeding with or delaying planned preventive maintenance.

2 The Dynamic Risk Simulator

A screenshot of the Dynamic Risk Simulator tool is included in Fig. 1 Screenshot of the Dynamic Risk Simulator tool. The tool provides estimates on equipment availability over finite time-horizon and total cost of maintenance for two maintenance strategies, a variant of the minimal-repair-periodic-overhaul maintenance strategy and the run-to-failure maintenance strategy.

A brief explanation on the two maintenance strategies modelled in the Dynamic Risk Simulator. Minimal repair is defined as restoring the system to the state it was immediately prior to its failure (Rausand and Høyland 2004). Minimal-repair-periodic-overhaul maintenance strategy is a combination of minimal repair, restoring the equipment to its functioning state immediately prior to failure in the shortest time possible towards minimising interruptions to operations and periodically over-hauling the equipment to an as-good-as-new state at each preventive maintenance interval. This strategy has been described in detail in Pham (2003). Other research conducted on this strategy can be found in the works of Tahara and Nishida (1975), Nakagawa (1984) and Block et al. (1985). The Dynamic Risk Simulator builds on this model assuming non-negligible repair times at each periodic preventive maintenance interval. Although the equipment is restored to as-good-as-new state at each periodic preventive maintenance interval, the Dynamic Risk Simulator simulates equipment aging/wear by modelling the time until the first failure of the equipment as a Weibull distribution.

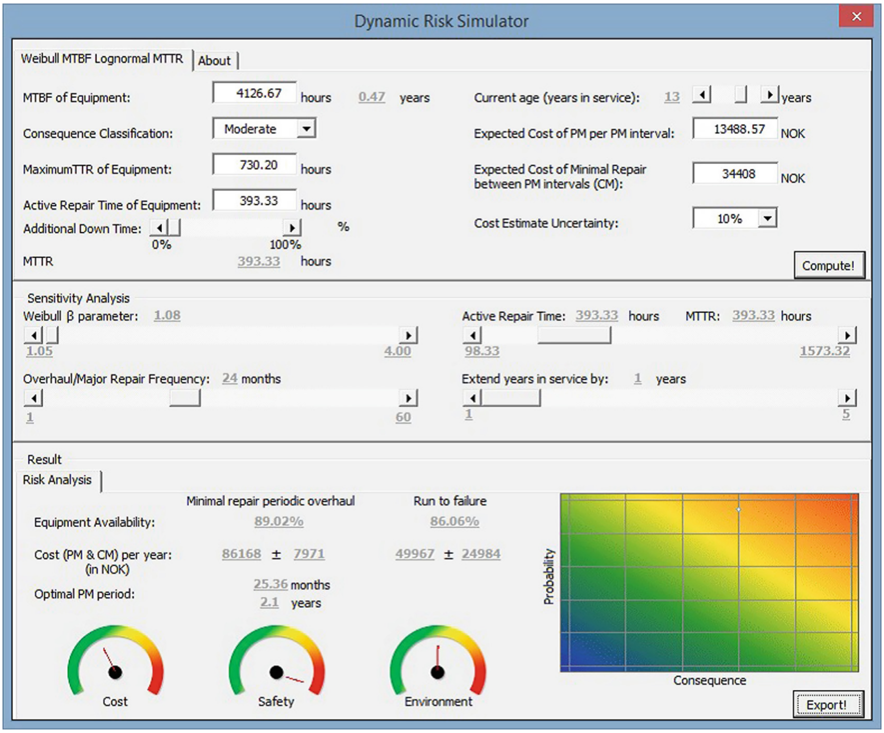


Fig. 1. Screenshot of the dynamic risk simulator tool

Equation 2.1 is based on the equations for the expected cost per maintenance interval for the minimal-repair-periodic-replacement maintenance strategy included in Pham (2003). The mathematical function in Eq. 2.1 in layman terms is the expected number of failures in a given time interval multiplied by the expected cost of CM, added to the expected cost of each PM within the same time interval, the time interval being the PM interval for the minimal-repair-periodic-overhaul maintenance strategy. Minimising this function yields a mathematical optimum PM interval for an equipment for this model (Fig. 2).

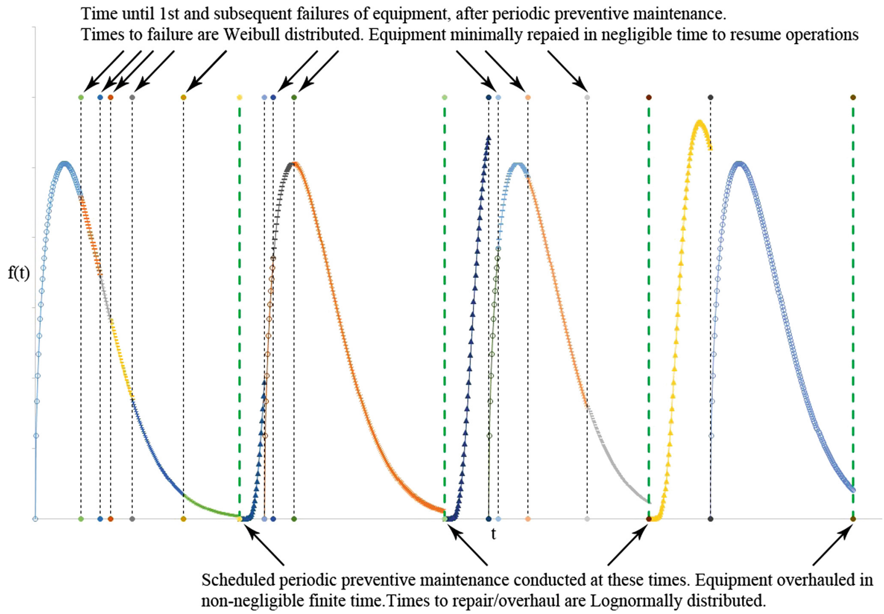


Fig. 2. PDF of the simulated times between failures and simulated times to perform repairs for the variant of the minimal-repair-periodic-overhaul maintenance strategy with non-negligible PM repair times as implemented in the Dynamic Risk Simulator. Figure for illustrative purposes only.

$$\text{Expected total cost of PM and CM, per PM interval } u = \frac{(\mu_{CM}) \int (\lambda(t)dt) + \mu_{PM}}{\mu_{CM} \alpha u./0 + \mu_{PM} u/0} \tag{2.1}$$

The optimal PM interval for the this maintenance strategy is given in Eq. 2.2

$$u_{optimal} = \left(\frac{\mu_{PM}}{\mu_{CM} \alpha \beta(-1)} \right)^{0} \tag{2.2}$$

Where $u_{optimal}$ is the optimal PM interval, μ_{PM} is the *expected cost of preventive repair* and μ_{CM} is the *expected cost of minimal-repair in-between preventive repair intervals* (i.e. corrective repair), α and β are the scale and shape parameters respectively of the two-parameter Weibull distribution which models the time to first failure of the equipment. λ(t) is the rate of occurrence of failures (or failure-rate or hazard-rate function) for the two-parameter Weibull distribution.

In the Dynamic Risk Simulator the failure times are modelled as Weibull distribution, the repair times are modelled as lognormal distribution and the corrective and preventive maintenance costs are each modelled by Normal distributions with expected costs μ_{CM} and μ_{PM} respectively with their respective standard deviations σ_{CM} and σ_{PM}. Between 10,000 and 100,000 data-points belonging to each of these distributions are generated using the Mersenne Twister algorithm in the Dynamic Risk Simulator and are used to run dynamic simulation scenarios on the total cost of maintenance, equipment availability and optimum preventive maintenance interval.

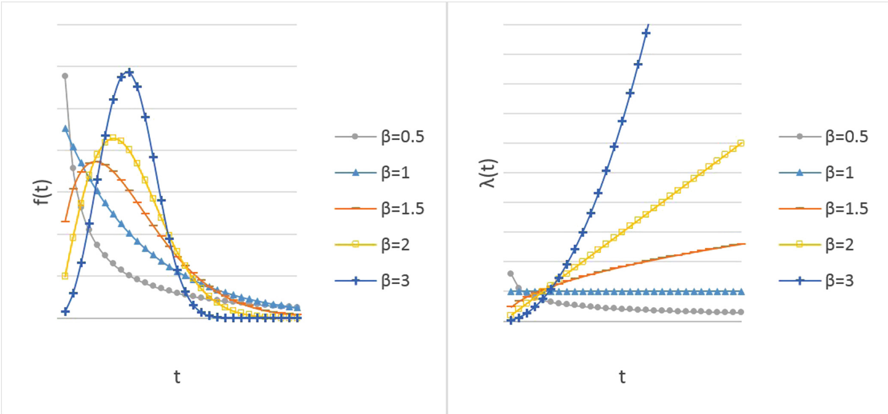


Fig. 3. Left: Probability Distribution Function of Weibull distribution for the same value for α and various values of β. Right: Failure-rate (hazard-rate) of Weibull distribution for the same value for α and various values of β. Decreasing failure-rate for $\beta < 1$; constant failure-rate for $\beta = 1$; increasing failure-rates for $\beta > 1$.

The Weibull distribution (Weibull 1951) was chosen since it can emulate a wide variety of failure time distributions by varying its shape (β) and scale parameters (α), Fig. 3 Left: Probability Distribution Function of Weibull distribution for the same value for α and various values of β. Right: Failure-rate (hazard-rate) of Weibull distribution for the same value for α and various values of β. Decreasing failure-rate for $\beta < 1$; constant failure-rate for $\beta = 1$; increasing failure-rates for $\beta > 1$. In particular, a Weibull distribution with shape parameter $\beta > 1$ simulates equipment degradation with time (or aging or wear with time). This in turn allows longer periodic preventive maintenance intervals for newer equipment and dictates shorter periodic preventive maintenance intervals for older equipment. Preventive maintenance interval for an equipment is hence not a constant in this model but varies depending on equipment age, cost of corrective maintenance (CM), cost of preventive maintenance (PM) and other factors. Details on the two-parameter Weibull distribution, goodness-of-fit tests, Weibull plot and others can be referenced in Rinne (2009).

Other factors considered in the development of the Dynamic Risk Simulator include single-unit repairable system, homogeneous data-type (mean time between failures, mean time to repair), finite time horizon, optimum preventive maintenance interval based on minimising the sum of preventive and corrective maintenance costs in a given time period. Factors excluded from the Dynamic Risk Simulator model – modes of equipment failure, imperfect repair (Malik 1979; Nakagawa 1979a, b; Nakagawa and Yasui 1987), multi-unit systems (series, parallel, k-out-of-n systems, Markov chains, petri-nets), other optimisation criteria (maximise equipment availability, minimise operations downtime).

3 Partial Validation Results

3.1 Water Injection Booster Pump

The lifetime data for a water injection booster pump on-board an offshore platform located in the Norwegian Sea was made available by Apply Sørco (Table 1).

Through backward simulation of the known maintenance strategy, age, PM and CM costs for another pump of the same model working in the same plant under same operating loads, the Weibull β parameter for the pump model was estimated to be 1.24 using the Dynamic Risk Simulator. Using this value for the Weibull β parameter, the Dynamic Risk Simulator estimated the availability of the given pump at 98% for the minimal-repair-periodic-overhaul maintenance strategy and at 85% for the run-to-failure strategy. The Dynamic Risk Simulator estimated maintenance cost for pump per year as 55976 $\pm$ 4629 NOK (Eq. 2.1) for the minimal-repair-periodic-overhaul strategy and this lies between the PM and CM costs data provided for the analysis and seems to be within expected values for an initial estimate.

Table 1. Data for a water injection booster pump

Mean Time Between Failures (MTBF)		12112	Hours
Mean Time To Repair (MTTR)		163	Hours
Maximum Time To Repair		286	Hours
Preventive Maintenance (PM)	September 2007	1422	NOK
	Cost of Condition Monitoring per week for 23 pumps	4555	NOK
	Cost of Condition Monitoring per week per pump	198	NOK
	Cost of Condition Monitoring per PM interval	20597	NOK
	Total Cost of PM per PM interval	22020	NOK per 2 years
	PM interval	2	Years
Corrective Maintenance (CM)	June 2003	80455	NOK
	August 2005	43551	NOK
	Average Cost of CM per PM interval	62003	NOK per 2 years

The Dynamic Risk Simulator estimated maintenance cost for pump per year as 44085 $\pm$ 20574 NOK for the run-to-failure strategy which has a very large uncertainty associated with the costs. One plausible explanation could be the difference between the MTBF which is 1.38 years and the age of the equipment which is 13 years. Under the run-to-failure strategy, no preventive maintenance is assumed to have been done on the pump and the pump is repaired at failure, hence simulating the life of the pump over 13 years with its MTBF 1.38 years and a Weibull beta of 1.24 could result in fewer or more failures and hence lower or higher corrective repair costs resulting in large uncertainty in overall maintenance costs per year.

3.2 Motor Driving a Water Injection Booster Pump

The lifetime data for the motor driving the water injection booster pump mentioned in the previous section, on-board an offshore platform located in the Norwegian Sea was made available by Apply Sørco (Table 2 – Data for a motor driving a water injection booster pump).

Table 2. Data for a motor driving a water injection booster pump

Mean Time Between Failures (MTBF)		4126	Hours
Mean Time To Repair (MTTR)		93	Hours
Maximum Time To Repair		730	Hours
Preventive Maintenance (PM)	2013	11726	NOK
	2015	15250	NOK
	Average cost of PM per PM interval	13488	NOK
	PM interval	2	Years
Corrective Maintenance (CM)	June 2003	5060	NOK
	October 2003	12144	NOK
	Total amount for 2003	17204	NOK
	Estimated Cost of CM per PM interval	34408	NOK per 2 years

As with the pump, the Weibull β parameter for the motor model was estimated to be 1.08. The Dynamic Risk Simulator estimated availability of the motor at 89% for the minimal-repair-periodic-overhaul and at 86% for the run-to-failure strategy. The Dynamic Risk Simulator estimated maintenance cost for the motor per year as 86168 $\pm$ 7971 NOK (Eq. 2.1) for the minimal-repair periodic-overhaul strategy, which is considerably higher than the costs data provided. The Dynamic Risk Simulator estimated maintenance cost for the motor per year as 49967 $\pm$ 24984 NOK for the run-to-failure strategy which has a very large uncertainty associated with the costs. As with the pump, one plausible explanation could be the difference between the MTBF which is 0.47 years and the age of the equipment which is 13 years.

3.3 Discussion

Reviewing results from the simulator, different strategies are found to be optimal for each type of equipment. From the total cost estimate, the minimal-repair-periodic-overhaul strategy is best suited for equipment class pumps over the selected time horizon. While for equipment class motors, the simulator predicts that run-to-failure is the preferred strategy over the selected time horizon. For both equipment, based on availability, between the two maintenance strategies, the minimal-repair-periodic-overhaul strategy seems preferable. While, for both equipment, based on the costs (including uncertainty), between the two maintenance strategies, the run-to-failure maintenance strategy seems preferable.

4 End Applications and Conclusion

From a commercial standpoint, the Dynamic Risk Simulator reduces the redundancy of condition monitoring by leveraging on reverse-engineered Weibull shape and scale parameters from one equipment and providing customised preventive maintenance intervals for all equipment of the same type subjected to the same operating conditions. This eliminates the need for capital intensive sensors and monitoring systems and the constant monitoring by maintenance engineering teams.

The Dynamic Risk Simulator could minimise total maintenance costs by providing customised optimum preventive maintenance interval taking into account the equipment's age and other factors. For equipment which experience wear due to age, newer equipment could work well with longer periodic preventive maintenance intervals while older equipment would require shorter periodic preventive maintenance intervals. Taking into account the equipment's age, the Dynamic Risk Simulator could make do with repair teams on standby which periodically perform preventive maintenance at too frequent or too long intervals, both being non-optimal to equipment. Depending on the distribution of equipment age in a given plant, decisions could be arrived at using the Dynamic Risk Simulator on whether repair teams could be shared resources or whether it makes commercial sense to have one on permanent standby.

The decision of whether to proceed with or to delay planned maintenance arises both when activity levels are high and when activity levels are low, but for different reasons. When activity levels are high, the downtime due to maintenance translates to lost revenue which could justify decisions on delaying planned maintenance. When activity levels are low, the fixed costs associated with planned maintenance might not outweigh the benefit of an operational ready equipment which need not be necessarily put into operation. By simulating the maintenance strategies and comparing the equipment availability in a finite time-horizon, maintenance managers can simulate the cost impact and marginal improvement/detriment to equipment availability should he/she decide to delay/advance planned maintenance or increase/decrease preventive maintenance interval.

The Dynamic Risk Simulator provides qualitative and quantitative estimates on the cost of maintenance, optimal PM interval and equipment availability for finite time horizon for two maintenance strategies. The Dynamic Risk Simulator cannot replace human judgement with regards to taking the final call on whether or not to postpone maintenance. The simulation tool is reliant on the experience and insight of industry experts to take the most appropriate course of action. Given the limited partial validation test results, further validation of the Dynamic Risk Simulator is required with equipment data before it can be considered for integration into existing maintenance management systems.

References

Barlow, R., Hunter, L.: Optimum preventive maintenance policies. Oper. Res. **8**(1), 90–100 (1960)

Block, H., Borges, W., Savits, T.: Age-dependent minimal repair. J. Appl. Probab. **22**(2), 370 (1985)

Malik, M.A.K.: Reliable preventive maintenance scheduling. AIIE Trans. **11**(3), 221–228 (1979)

Nakagawa, T.: Optimum policies when preventive maintenance is imperfect. IEEE Trans. Reliab. **R-28**(4), 331–332 (1979a)

Nakagawa, T.: Imperfect preventive-maintenance. IEEE Trans. Reliab. **R-28**(5), 402 (1979b)

Nakagawa, T.: Optimal policy of continuous and discrete replacement with minimal repair at failure. Naval Res. Logist. Q. **31**(4), 543–550 (1984)

Nakagawa, T., Yasui, K.: Optimum policies for a system with imperfect maintenance. IEEE Trans. Reliab. **R-36**(5), 631–633 (1987)

Pham, H.: Handbook of Reliability Engineering, p. 353. Springer, London (2003)

Rausand, M., Høyland, A.: System Reliability Theory: Models, Statistical Methods and Applications. Wiley-Interscience, Hoboken (2004)

Rinne, H.: The Weibull Distribution: A Handbook. CRC Press, Boca Raton (2009)

Tahara, A., Nishida, T.: Optimal replacement policy for minimal repair model. J. Oper. Res. Soc. Jpn. **18**(3–4), 113–124 (1975)

Wang, H.: A survey of maintenance policies of deteriorating systems. Eur. J. Oper. Res. **139**(3), 469–489 (2002)

Weibull, W.: A statistical distribution function of wide applicability. J. Appl. Mech. **18**(3), 293–297 (1951)

Augmented Reality Technology for Predictive Maintenance Education: A Pilot Case Study

Idriss El-Thalji[1](✉), Mustafa Serkan Abdüsselam[2], Santiago Echeverri Duque[1], and J. P. Liyanage[1]

[1] Industrial Asset Management, University of Stavanger, Stavanger, Norway
idriss.el-thalji@uis.no

[2] Department of Computer Education and Instructional Technology (CEIT), Giresun University, Giresun, Turkey

Abstract. Industry 4.0 is an industrial era where several disruptive technologies e.g. Internet of Things, cloud technology, 3D printed, advanced robotics and materials have merged or integrated to enables a new level of organising and controlling the entire valve chain with the product lifecycle by creating dynamic and real-time understanding of cross-company behaviours. Such collaboration between physical and cyber spaces is rapidly growing and the benefits becoming more feasible. However, the interaction of human (in case automation is not applicable) at physical space with digital contents requires special facilitation where argumentation is the key. Predictive maintenance which is based on real-time measurements e.g. health/process parameters requires high level of expertise to understand and diagnose the present status of machine health. Therefore, the purpose of this paper is to explore, through a pilot project, how could augmented reality technology be utilised to assist novice learners to gain deep understanding of predictive maintenance process and practical skills as well. The case study is developed to assist the student of industrial asset management programme to learn the machine fault simulator and the analysed content of vibration measurements. Such contribution aims to enhance the training technology and accelerate the learning process of novice technician/engineers to be expert.

Keywords: Augmented reality · Predictive maintenance · Engineering education · Training technology · Industry 4.0

1 Introduction

Augmented reality gain gradually a real footing in the industrial applications in early 1990s, after almost three decades of Ivan Sutherland's experiments (Navab 2004). The idea of augmented reality is to enhance the user's understanding and learning by bringing a visualised additional information as seamlessly as possible into the user's view (Bernd Schwald and De-Laval 2003). The application of augmented reality illustrated in several industrial field, plant life cycle management (design, construction, operation and maintenance) (Siltanen et al. 2007), design process (Dunston et al. 2003), facility management (Gheisari and Irizarry 2016), operation process (Träskbäack and Haller 2004).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 600–609, 2020.
https://doi.org/10.1007/978-3-030-48021-9_67

In summary, augmented reality technology offers a wide range of benefits over the whole life cycle phases for any asset. It is extensively utilised in education to assist the students to understand the basics of several topics. Moreover, it helps the industrial stakeholders to design, install, operate and maintain their assets using both training and in field learning (Bernd Schwald and De-Laval 2003). Augmented reality was also applied to help the maintenance personnel to repair the equipment by overlaying several graphs, information and expert suggestions with the actual equipment in real time situation (Henderson and Feiner 2007). Such adaptive augmented reality systems have significant enhancements on the productivity of maintenance personnel (Henderson and Feiner 2007). However, maintenance operations are not limited to only repair/replace operations of specific physical assets. Specially, the new generation of maintenance, which should comply with industry 4.0 vision, is a combination of physical and cyber assets. Thus, the industrial worker might need augmented features/objects that illustrate both physical asset and digital content e.g. spectrum plot in more simplified and/or visualised manner.

Therefore, the purpose of this paper is to explore, through a pilot project, how could augmented reality technology be utilised to assist novice learners to gain deep understanding of predictive maintenance process and practical skills as well. The case study is developed to assist the student of industrial asset management programme to learn the machine fault simulator and the analysed content of vibration measurements. Consequentially, the experimental set up and learning outcomes are described in the following section. Later, the developed augmented reality application is described. Finally, several issues are discussed to draw up clear conclusion of the current maturity level of the developed application and future roadmap.

2 Exploration of AR in Predictive Maintenance: A Case Study

Basically, predictive maintenance is based on using real-time measurements to detect and diagnose the asset health and predict the remaining useful lifetime in case a fault is detected. It means that the maintenance engineer/technician needs to learn how to measure, detect a fault (using time and/or frequency domain) and diagnose the fault features e.g. type, location and severity. Measuring the asset health can be acquired using several type of sensors and can be transmitted via cable to desktop or cloud server, or Wi-Fi and Bluetooth. Detecting the fault or health abnormality can be done using several signal analysis techniques e.g. spectrum analysis. However, diagnosis is quite demanding since a good understanding of the physical asset and its dynamics is required in order to determine the fault type, location and severity. Therefore, the first consideration in designing this case study is the smooth learning process by starting with providing augmented learning content of the machine fault simulator (as the selected asset within the project, shown in Fig. 1), whereas measuring, detection and diagnosis modules are sequentially provided after that module.

The second consideration is related to teaching/learning process where 5E's model is implemented (Bybee et al. 2006). The 5 E's is an instructional model based on the constructivist approach to learning that describes learning in 5 phases: Engage (connect

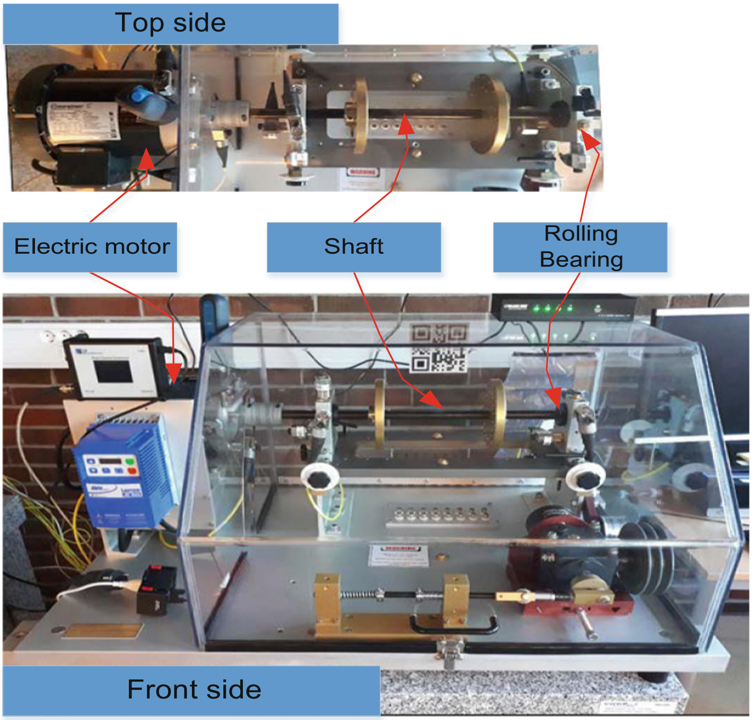


Fig. 1. Machine fault simulator and the experimental set up

with past experience or link to required concept/skill), Explore (explore the environment or manipulate materials), Explain (to verbalize their conceptual understanding or to demonstrate new skills or behaviours), Elaborate (to practice/obtain more skills and behaviours), and Evaluate (to assess students' understanding of key concepts and skill development and teaching effectiveness). As mentioned, the digital worksheet is designed to achieve the learning objectives using 5 E's model as illustrated and described in Table 1.

We have developed a digital worksheet for the students to facilitate their lab exercise toward learning predictive maintenance process. The digital worksheet consists of 29 questions/steps that covers the 5 E's phases. The digital worksheet (as shown in Fig. 2) provides guidelines for the student about what to do and links to supplementary materials. Moreover, it provides space for them to upload pictures, record voice and video from the experimental environment, document their understanding and observations.

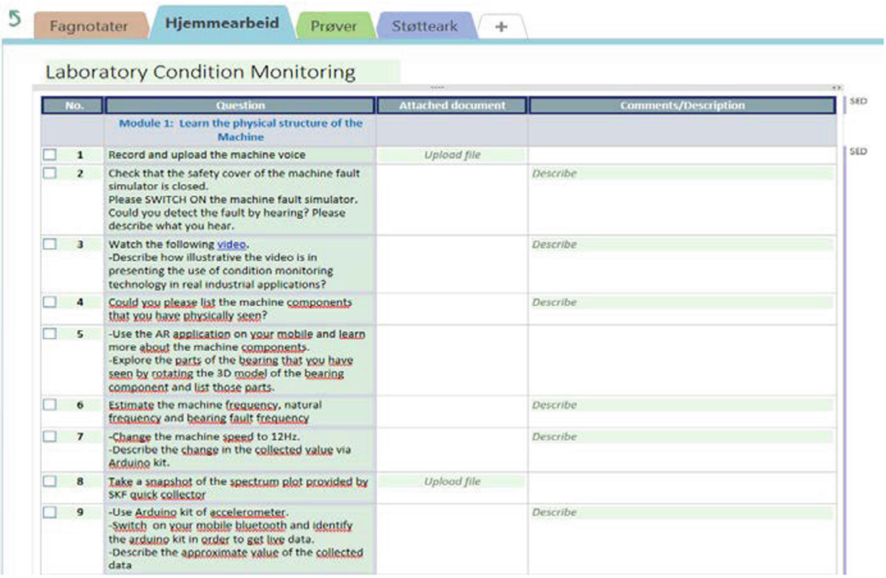
Table 1. The 5 E's model and learning objectives

5E's phases	Learning objectives	Methods
To engage	To engage the student to explore/learn about condition monitoring and predictive maintenance	- industrial video of the application - running the machine fault simulator with defective bearing
To explore	To explore the machine components e.g. motor, rotor, bearing (outer, inner, rolling elements)	
	To explore machine physics and dynamics	- equation of motion - component frequencies equations
	To explore the monitoring life cycle from data to decision	- 7 layers of Industry 4.0
	To explore the sensing technology To explore speed effect on machine dynamics	- Arduino kit for accelerometer measurements
	To explore signal analysis tool To explore speed effect on machine dynamics in terms of spectrum plot	- Quick collector SkF
	To explore professional monitoring systems e.g. channels, cloud solution, data sets, plots	- SKF IMx8 and @ptitude
To explain	To explain the collected data e.g. axis, values, value changes due to speed changes	Data from Bluetooth and Arduino
	To explain the mechanism of wireless accelerometer technology To explain the spectrum plot e.g. axis, peals and their values, value changes due to speed changes	
	To explain the time plot and waveform of healthy data set e.g. shape of the signals	
	To explain the time plot and waveform of faulty data set e.g. the impact signals (physical meaning of the impacts), numbers of impact per second, shape of the signals	
	To explain the spectrum plot and waveform of healthy data set e.g. peaks at fundamental frequency	
	To explain the spectrum plot and waveform of faulty data set e.g. peaks at 1x, fault frequencies, natural frequency zone, amplitudes	

(*continued*)

Table 1. (*continued*)

5E's phases	Learning objectives	Methods
To elaborate	To compare between measure and estimated frequencies e.g. machine frequency, natural frequency, and fault frequencies	
	To compare between time plot of both healthy and faulty cases	
	To compare between spectrum plot of both healthy and faulty cases	
	To compare between waterfall spectrum plot of both healthy and faulty cases	
	To compare between the spectrum plot of the faulty case extracted by quick collector and IMx8	
To evaluate	To assess the student understanding of machine functionality, faults and corresponding dynamic behaviour, measurement techniques, transmission techniques, diagnostic techniques	


Fagnotater Hjemmearbeid Prøver Støtteark +

Laboratory Condition Monitoring

No.	Question	Attached document	Comments/Description
	Module 1: Learn the physical structure of the Machine		
1	Record and upload the machine voice	Upload file	
2	Check that the safety cover of the machine fault simulator is closed. Please SWITCH ON the machine fault simulator. Could you detect the fault by hearing? Please describe what you hear.		Describe
3	Watch the following video. -Describe how illustrative the video is in presenting the use of condition monitoring technology in real industrial applications?		Describe
4	Could you please list the machine components that you have physically seen?		Describe
5	-Use the AR application on your mobile and learn more about the machine components. -Explore the parts of the bearing that you have seen by rotating the 3D model of the bearing component and list those parts.		
6	Estimate the machine frequency, natural frequency and bearing fault frequency		Describe
7	-Change the machine speed to 12Hz. -Describe the change in the collected value via Arduino kit.		Describe
8	Take a snapshot of the spectrum plot provided by SKF quick collector	Upload file	
9	-Use Arduino kit of accelerometer. -Switch on your mobile bluetooth and identify the arduino kit in order to get live data. -Describe the approximate value of the collected data		Describe

Fig. 2. Snapshot of the digital worksheet for predictive maintenance lab exercise.

3 Development of the AR Application and the Pilot Use

Within the experimental exercise that has been developed in the previous section, there are couple of steps where AR technology can be useful to facilitate the student learning by providing augmented objects e.g. 3D objects, texts, arrows. The authors determined three steps within the experimental exercise that AR can be used: (module 1) provide virtual 3D drawing of the physical machine, (module 2) virtual guiding arrows projected on the spectrum plot (that generated by SKF @ptitude software), (module 3) virtual 3-axes projected beside the transmitted data (acceleration measurements of x, y, and z axis). Modules 1, 2 and 3 are respectively named as "Learn module", "Basic module" and "Expert module", as shown in Fig. 3.

Fig. 3. Interface of AR application for predictive maintenance education.

3.1 Learn Module

A simplified 3D drawing of the machine fault simulator is projected into the physical world (the machine by itself). The 3D drawing consists of three main components, as shown in Fig. 4: electric motor, shaft/rotor and bearing. In this case, the machine fault simulator has a class cover where the student see the basic components and don't really need an augmented objects. However, all machines in industry are almost covered with metallic cover and very rarely they have transparent view. In fact, this is one of the most significant advantage of augmented technology where hidden objects e.g. components, pipelines can be virtually visualised. One good example of how AR technology could help to visualise hidden physical items is the bearing item that is used in the experiment. The physical picture (Fig. 5) of the bearing and its housing show nothing about the bearing elements. Moreover, the virtual representation facilitates the student with several features, for example, to zoon-in and rotate the axis, as shown in Fig. 5.

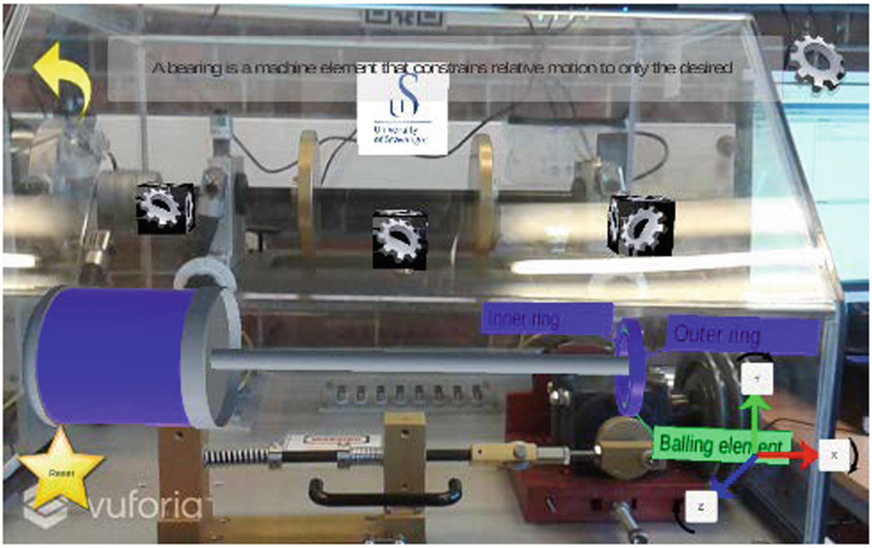


Fig. 4. Augmented 3D drawing of the machine fault simulator

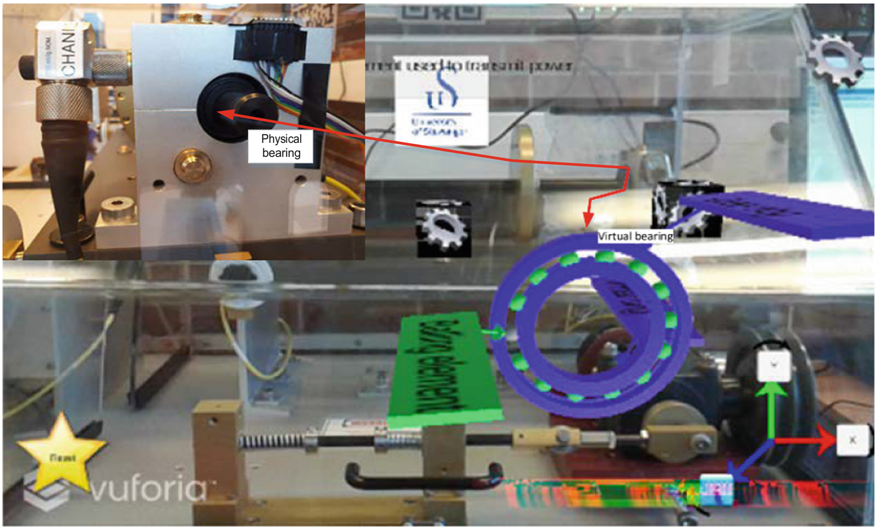


Fig. 5. Augmented bearing: rotated and enlarged.

3.2 Basic Module

After the physical machine was augmented with virtual 3D drawing, the machine simulator was operated and measurements were taken using accelerometers and transmitted via internet into SKF cloud server. Then, the student is able to log/in into SKF could server and get the signal analysed by SKF @ptitude software. Therefore, the purpose of this AR module is to facilitate the student to read/understand the spectrum plot, which is automatically generated by SKF @ptitude software, as shown in Fig. 6. Moreover, there are three augmented objects that are projected on the spectrum plot to facilitate the student to determine what are the critical frequencies that s/he need to look at. Arrow No. 1 indicates the machine frequency, where arrow No. 2 indicates fault frequencies. Finally, arrow No. 3 indicates the natural frequency range of the machine.

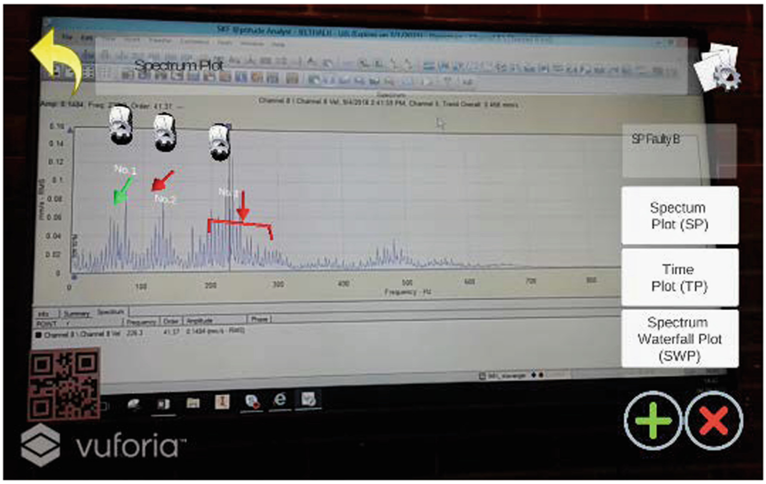


Fig. 6. Augmented spectrum plot

3.3 Expert Module

The expect module is linked to the student activity of building their own sensor and transmission set up. Arduino kit is used where an accelerometer is connected into

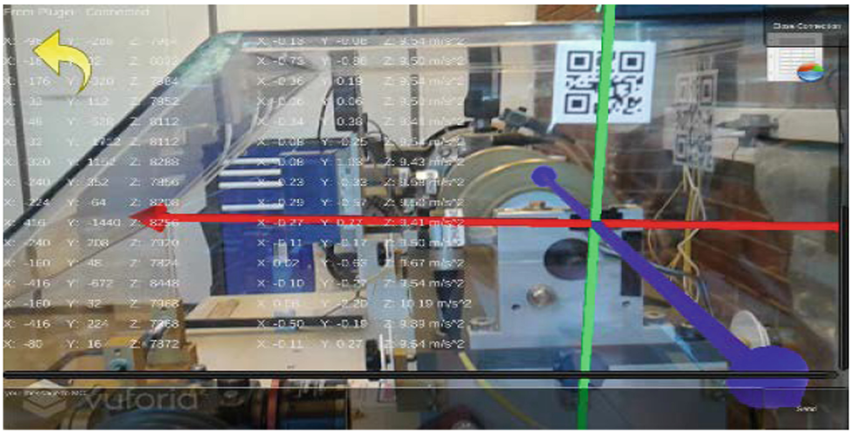


Fig. 7. Augmented acceleration measurements

Arduino Uno microcontroller and then data is transmitted via Bluetooth to student tablet or mobile. The accelerometer is three axial sensor where AR application is used to visualise the three axes directions (Fig. 7).

4 Discussion and Conclusion

It clear that AR technology can facilitate the education and training of predictive maintenance. It can be used to visualise physical content, but it can also be used to visualise digital content e.g. spectrum plot. For sure, the augmented features that simplified the digital content can be later embedded into the digital content. However, this pilot project is clearly showing that the effectiveness of AR application is highly dependent on several issues: (1) user interface and integrity (2) availability and veracity of required visualised information (text, static picture, video, audio, haptic), (3) environmental compatibility e.g. position/space constraints, vibration, weather, temperature, light, noise and humidity), inaccessible (4) system complexity e.g. flexible/movable parts, geometrical dependencies, scale of the "field of view" (5) task cognitive effort level, complexity and ergonomics e.g. requires gloves, glass, helmet, fast and live data, local/remote assistance connectivity, ability to update/report. The aim of this pilot project is to explore the opportunities, challenges and dependence of AR technology for industrial training and education sector. Thus, even though, the pilot application required comprehensive enhancements, this stage of development was useful to technical explore such enhancements.

Acknowledgment. Financial support from the University of Stavanger, Faculty of Science and Technology for ARTecho project and The Department of Mechanical and Structural Engineering and Materials Science for IAM lab are acknowledged.

References

Bernd Schwald, B., De-Laval, B.: An augmented reality system for training and assistance to maintenance in the industrial context. J. WSCG (2003)

Bybee, R.W., Taylor, J.A., Gardner, A., Van Scotter, P., Powell, J.C., Westbrook, A., Landes, N.: The BSCS 5E Instructional Model: Origins, Effectiveness, and Applications. BSCS (2006)

Dunston, P.S., Wang, X., Billinghurst, M., Hampson, B.: Mixed reality benefits for design perception. Nist Spec. Publ. SP 191–196 (2003)

Gheisari, M., Irizarry, J.: Investigating human and technological requirements for successful implementation of a BIM-based mobile augmented reality environment in facility management practices. Facilities **34**, 69–84 (2016)

Henderson, S.J., Feiner, S.K.: Augmented reality for maintenance and repair (ARMAR). Air Force Research Laboratory Human Effectiveness Directorate Warfighter Readiness Research Division Logistics Readiness Branch Wright-Patterson AFB OH 45433-7604 (2007)

Navab, N.: Developing killer apps for industrial augmented reality. In: Rosenblum, L., Macedonia, M. (eds.) Projects in VR IEEE Computer Society (2004)

Siltanen, P., Karhela, T., Woodward, C., Savioja, P.: Augmented reality for plant lifecycle management. In: 13th International Conference on Concurrent Enterprising (ICE2007), 4–6 June 2007, Sophia Antropolis, France (2007)

Träskbäack, M., Haller, M.: Mixed reality training application for an oil refinery. In: Proceedings VRCAI 2004, ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications in Industry, 16–18 June 2004, Nanyang Technological University, Singapore (2004)

Integrated Production and Maintenance Planning for Successful Asset Management Strategy Implementation

Harald Rødseth[1(✉)], Alan Wilson[2], and Per Schjølberg[1]

[1] Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway
{Harald.Rodseth,Per.Schjolberg}@ntnu.no
[2] Carmichael Smith Ltd., Nottingham, UK
Carmichael.smith@gmail.com

Abstract. The balance between maintenance and production activities plays a pivotal role for any asset in an organisation. Maintenance activities can require that the asset must be unavailable for production during maintenance and hence reduce the production volume. However, if the asset is not maintained the technical condition will degrade and can result in unexpected breakdowns and reduced plant capacity. An important question will then be: How can an organisation ensure the right balance of maintenance during the lifetime of the asset? To approach this question, a fundamental understanding in asset management and strategic approach in an organisation is indeed necessary. The standard EN16646 gives further detail of the interrelationship between maintenance and other processes at asset and asset system level. Although a list is provided, it remains to investigate how the relationship between operation and maintenance processes can contribute to an asset management strategy. The objective of this article is therefore to propose an approach in asset management strategy that will ensure a right balance between plant capacity and maintenance activities during the life cycle of an asset.

1 Introduction

The role of asset management can be considered to be the meeting point between the technical and business fields (Hastings 2015). The asset manager must bring to bear the skills of technical knowledge and business knowledge which require a capability to communicate both to the technicians as well as the board members in the organisation.

In asset management, value is one fundamental requirement, whereby assets exist to provide value to the organisation (ISO 2014). For strategic asset management, it is vital to ensure that value is created and kept throughout the life cycle of the asset. When further evaluating the strategic dimensions for the physical asset, the standard EN 16646 is of relevance where important interrelationships between operating and maintaining the asset are presented (CEN 2014).

To ensure high value creation in the asset all the disciplines that affect the technical condition must be integrated and planned for in integrated planning (IPL). To

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 610–617, 2020.
https://doi.org/10.1007/978-3-030-48021-9_68

understand this process, IPL can be understood according Rødseth (2017) to be "*multidisciplinary decision making process that manages technical condition and results in increased production, reduced costs, improved safety. This process is performed in a manner that optimises across multiple planning disciplines through updating of objectives and attended by the power and intention to commit resources and to act as necessary to implement the chosen plan.*"

IPL should foster the asset management strategy (Bai and Liyanage 2012). This has been demonstrated in an earlier study where the maintenance programme for a gas turbine has been evaluated (Rødseth et al. 2016). In this case study, the balance of maintenance activities and plant capacity is of importance where the indicator profit loss indicator (PLI) has been proposed.

Several proposed asset management strategies exist today (Volkova and Kornienko 2014, Wang et al. 2016). Although these strategies comprise vital elements such as performance measures and technology aspects, it remains to investigate more in detail how the balance between plant capacity and maintenance activities is ensured in the strategy.

The objective of this article is to propose an approach for asset management strategy that will improve balance between plant capacity and maintenance activities during the life cycle of an asset.

The further structure in this article is as follows: Sect. 2 presents relevant asset management strategies based on existing literature. Section 3 proposes an asset management strategy for improving the balance between plant capacity and maintenance activities and discussed based on a relevant example of asset management strategy. Finally, concluding remarks are made in Sect. 4.

2 Relevant Asset Management Strategy Concepts

2.1 Balanced Scorecard and Physical Asset Management

Asset management strategy can be denoted as strategic asset management plan (SAMP) and defined as follows: *"Documented information that specifies how organizational objectives are to be converted into asset management objectives, the approach for developing asset management plans, and the role of the asset management system in supporting achievement of the asset management objectives."* (ISO 2014) From this definition, strategy should not only be understood as a plan, but also establishing objectives and using a system for achieving the objectives.

Inspiration from balanced scorecard could provide valuable contribution to the asset management strategy. An argument that supports this view is that performance is one of the key components in a strategy (Guerras-Martín et al. 2014). The value of a balanced scorecard for measuring performance in a strategy has also been discussed thoroughly (Atkinson 2006, De Geuser et al. 2009). Balanced scorecards measure several aspects for an enterprise and form the causality between the measurements within four perspectives (Atkinson 2006):

- *Financial:* To ensure shareholder satisfaction, measures here generally involve profitability and other financial measures such as sales growth.

- *Customer:* This perspective is focusing on "real" customer satisfaction such as delivery time.
- *Internal processes:* Important measures for this perspective should highlight, for example, critical processes in the organisation.
- *Learning:* This perspective underpins the three perspectives with indicators for improving flexibility and investing for future improvements.

Although balanced scorecard has been criticised, e.g. for challenges in "combining" it with other established systems in the organisation, it is still claimed that balanced scorecard can play a role in an organisation (Atkinson 2006).

In existing asset management strategies, the balanced scorecard seems to be of inspiration when establishing the asset management strategy (Wang et al. 2016). Furthermore, hierarchies for KPIs are also applied where production KPIs are at a corporate level whereas maintenance KPIs are at an operational level (Volkova and Kornienko 2014).

When further considering the asset management strategy, the standard EN 16646 for maintenance within physical asset management can be of relevance (CEN 2014). This standard clearly points out that the organisations strategies have a great influence on physical asset management. It details this further by presenting the interaction between the strategic analysis and requirements for physical asset management. This standard describes the interrelation between maintenance and other processes at the asset level. Figure 1 presents the relationship between operation and maintenance processes at the asset. Furthermore, the standard lists different issues that are necessary to consider when evaluating this relationship. For example, observations made by maintenance personnel that can reveal degradation levels higher than expected should be a contribution from maintenance and therefore as an output to the operation for determine appropriate operating profiles. When establishing an asset management strategy, such relationship during the operation phase of the asset should be supported by the asset management strategy.

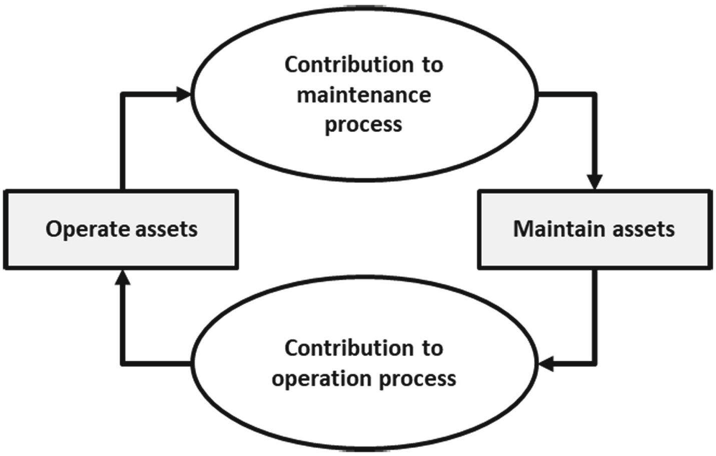


Fig. 1. Relationship between operation and maintenance processes, adapted from (CEN 2014).

2.2 World Class Maintenance as a Strategy

Maintenance has become a significant contributor towards achieving the strategic objectives of organizations in today's competitive markets. The best way for companies to compete in the future is to have an integrated aim and plan for maintenance, to have the most cost-effective products and to have the best integrated supply chain. This can be achieved through adapted processes.

Maintenance processes for serving the production facilities can guarantee high productivity, capacity, availability, safety and profit. The companies must aim to deliver world class maintenance (WCM). WCM is now a term which is often used.

A WCM organisation consistently demonstrates the best practice in maintenance and therefore achieves a competitive advantage over its competitors within maintenance and bottom-line results. Although it is challenging to find a uniform definition of WCM (Mueller), a sound understanding of WCM seems to be the function in a company that follows the maintenance practice that enables a competitive advantage in the organisation (Wireman 1990).

To be at world class, the companies must introduce and implement the correct maintenance processes. The processes must be comprised of planned and unplanned actions carried out to retain a physical asset to the acceptable operating condition. The process is a set of interrelated activities that use input to deliver an intended result (to be world class). It consists of many maintenance processes, which can be grouped in process families according to EN 17007 (CEN 2017):

- *Management processes*: These processes comprise setting the objectivities and the policy to be implemented in order to achieve them, and allocating the necessary resources.
- *Realization processes:* These affect directly the expected result and will ensure the needs expressed by the customer are satisfied.
- *Support processes:* These are important to the other process since they will provide the necessary resources:
  - human resources;
  - financial resources;
  - material resources and their maintenance; and
  - information processing.

2.3 Asset Management Indicators and Digital Twin

One recommended indicator for strategic decisions in asset management is the indicator PLI. This indicator has been demonstrated earlier for a gas turbine (Rødseth et al. 2016). In this case study, the frequency of preventive maintenance activities was considered to be reduced and thus increasing production efficiency and reducing maintenance costs. Figure 2 presents the expected improvements when based on strategic decisions from this case study. The decision support in the case study was to evaluate the opportunities of updating the maintenance programme from four maintenance activities down to three maintenance activities each year for the gas turbine. To justify this investment of updating the maintenance programme it was of interest to calculate the reduced PLI value using net present value (NPV) for 6 years. Due to increased production efficiency it would be expected that the reduced turnover loss would be 10.9 million USD over 6 years due to having three instead of four maintenance activities each year. If the extra costs were included as well such as maintenance costs for this "extra" maintenance activity, the profit loss would be reduced even more.

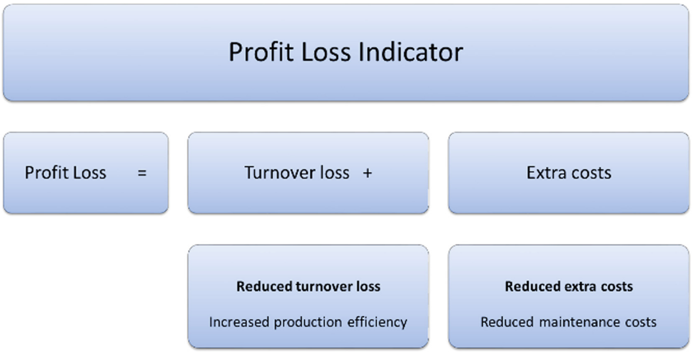


Fig. 2. Expected improvements based on earlier PLI calculations and adapted from (Rødseth et al. 2016).

An opportunity with today's focus in digitalization is to include PLI in a digital twin of the asset. The concept of digital twin seems to stem from both NASA and U.S. Airforce, and concerns mirroring the life cycle of a physical asset with a integrated multiphysics, multiscale and probabilistic simulation technologies where the best available models, sensors, and fleet history, are included (Glaessgen and Stragel 2012). The digital twin will in this context for instance forecast the health of the physical asset and the probability of mission success. The digital twin also seems to be useful for other industry branches such as power plants. For instance, General Electric (GE) has proposed a digital twin framework, see Fig. 3 which is adapted from (General Electric 2016). The digital twin will require specific customer needs in terms of reliability, capacity and emissions specified as KPIs. The digital twin will then have different data inputs, both real-time data in terms of operational data as well as historic maintenance and inspection data. The output from the digital twin will then be an asset performance management where data is transformed into robust analytics with domain expertise. With predictive analytics downtime is reduced and operational life is extended.

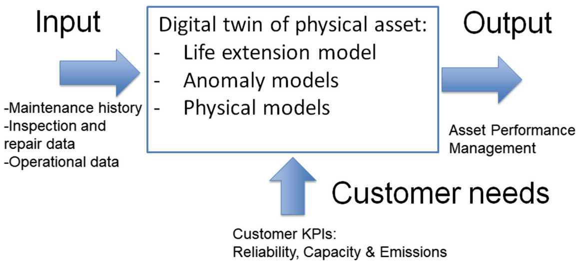


Fig. 3. Digital Twin at GE, adapted from (General Electric 2016).

3 Proposed Asset Management Strategy

3.1 Development of the Asset Management Strategy

In general, an asset management strategy should include an approach where the maintenance process is more balanced with the production process in order to maximize the value creation from the asset. Inspired by the presented the relevant literature in Sect. 2, the authors would propose an approach for asset management strategy that improves the balance between plant capacity and maintenance activities. This approach is divided into four main steps:

- **Step 1:** *Map the maintenance processes*. To improve the maintenance processes so that they can be more balanced with plant capacity, a structured approach for mapping them is recommended. The standard EN 17007 should be applied by the organisation in this step.
- **Step 2:** *Map the relationship between operating and maintaining the asset.* When the maintenance processes have been mapped, it can be possible to map more in the details about the relationship of the maintenance processes and the production processes based on EN 16646. As an example in EN 16646 it is pointed out that production plans will affect the maintenance plans when maintenance should be performed. Based on this contribution to the maintenance process it is then necessary to have analytical models of technical condition to understand the risk picture of postponing future maintenance activities.
- **Step 3:** *Develop relevant KPIs and guidelines for decision support.* In EN 16646 it is further emphasized the importance of performance monitoring in order to control the relationship between operating and maintain the asset. In particular this standard points out "silo" behaviour as one risk that can be avoided by using the KPIs. An example of a relevant KPI could be internal criticality of equipment such as bottlenecks. In addition to the KPIs outlined in EN 16646 the monetary indicator PLI should also be developed in order to understand the financial perspective to have a transparency between the technical effect of changing the maintenance programme in terms of reduced downtime and an indication of future savings for the company. Also, the KPIs should be balanced where e.g. quality of work processes are measured.
- **Step 4:** *Develop a digital twin for smart analytics*. To ensure an asset performance management system that includes sensor data and maintenance history to perform advanced analytics, it is recommended that the company develop a digital twin. With the customer needs communicated with KPIs such as reliability, the expected result would be faster strategic decision making with less "silo" behaviour in the organisation with more analytical models of e.g. the indicator PLI.

3.2 Example of the Asset Management Strategy

A relevant example of an asset management strategy is presented based on the authors experience. A chemicals and oil company in Europe developed six workstreams for implementation within a site transformation towards asset management and cost reduction. All site functions and KPIs were to be aligned with the new strategic objectives centred on reliability, cost, output, quality, and of course compliance. Inclusive demand and supply forums and improvement ideas meetings invited comment from functional heads to technicians, union representatives, and contractors. Auditable reasons for actions were based on the costs and benefits, implementation duration, and risks.

The organization of operations and maintenance was changed to one unit and the operations manager became the accountable asset manager. A single asset management plan was developed for the O&M, Projects and Shutdown activities, and software developed for integrated work scheduling and control planning system, including more focused risk based inspections. All the relevant management, realization and support processes on the site were renewed by the Technical Authorities, and additional training commenced. The result after two years included control of projects and procurement brought inhouse, achievement of the targeted cost per asset replacement value, and 20% reduction in contractor utilisation.

Based on this example it should be possible to relate it to the proposed approach for asset management strategy from Sect. 3.1. If the company has implemented the three first steps, the development of the digital twin should be of highly relevance in their strategic decision making in asset management. Figure 4 shows how the digital twin could have contributed to this asset management strategy based on the example for the chemicals and oil company. With a virtual access to the relevant data this could in future result in faster and better strategic decision making.

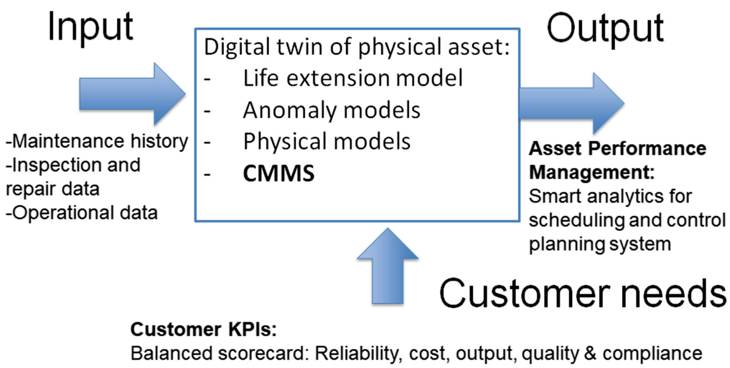


Fig. 4. Digital Twin at GE, adapted from (2016, General Electric).

4 Concluding Remarks

The main objective in this article is to propose an approach for asset management strategy that will ensure a right balance between plant capacity and maintenance activities during the life cycle of an asset. Based on relevant literature and own experiences in asset management strategy, as well as existing European standards in physical asset management and maintenance processes, an approach for asset management strategy divided into four main stages has been recommended. Further research will investigate how this strategy can be related to a detailed roadmap for implementation.

References

Atkinson, H.: Strategy implementation: a role for the balanced scorecard? Manag. Decis. **44**, 1441–1460 (2006)

Bai, Y., Liyanage, J.P.: Framework and systematic functional criteria for integrated work processes in complex assets: a case study on integrated planning in offshore oil and gas production industry. Int. J. Strateg. Eng. Asset Manag. **1**, 49–68 (2012)

CEN-European Committee for Standardization. EN 16646: Maintenance - Maintenance within physical asset management. Brussels (2014)

CEN-European Committee for Standardization. EN 17007: Maintenance process and associated indicators. Brussels (2017)

De Geuser, F., Mooraj, S., Oyon, D.: Does the balanced scorecard add value? Empirical evidence on its effect on performance. Eur. Account. Rev. **18**, 93–122 (2009)

General Electric. GE Digital Twin – Analytic Engine for the Digital Power Plant (2016)

Glaessgen, E.H., Stragel, D.S.: The digital twin paradigm for future NASA and U.S. air force vehicles. Paper for the 53rd Structures, Structural Dynamics, and Materials Conference: Special Session on the Digital Twin, pp. 1–14 (2012)

Guerras-Martín, L.Á., Madhok, A., Montoro-Sánchez, Á.: The evolution of strategic management research: recent trends and current directions. BRQ Bus. Res. Q. **17**, 69–76 (2014)

Hastings, N.A.J.: Physical Asset Management: With an Introduction to ISO55000. Springer, Cham (2015)

Mueller, S.: How to Define World-Class Maintenance. ReliablePlant. https://www.reliableplant.com/Read/29941/world-class-maintenance. Accessed 6 July 2018

ISO. ISO 55000 Asset management - Overview principles and terminology. Switzerland (2014)

Rødseth, H.: Development of Indicators for Maintenance Management within Integrated Planning. NTNU (2017)

Rødseth, H., Schjølberg, P., Kirknes, M., Bernhardsen, T.I.: Increased profit and technical condition through new KPIs in maintenance management. In: Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015). Springer, Cham (2016)

Volkova, I., Kornienko, E.: The approach to the asset management strategy choice in an electric grid company. WIT Trans. Ecol. Environ. **190**, 61–71 (2014)

Wang, J., Chen, Z., Parlikad, A.: Designing performance measures for asset management systems in asset-intensive manufacturing companies: a case study. In: Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015). Springer, Heidelberg (2016)

Wireman, T.: World Class Maintenance Management. Industrial Press Inc., New York (1990)

Dynamic Maintenance Scheduling Based on Cost Analysis and Genetic Algorithm for Offshore Facilities

Guicang Peng(✉) and Tore Markeset

University of Stavanger, Stavanger, Norway
{Guicang.peng, tore.markeset}@uis.no

Abstract. This paper introduces a dynamic Maintenance Work Order (MWO) schedule model for offshore facilities' daily maintenance management. The objective of the MWO schedule model is to improve maintenance performance by reducing MWO overall delay and suspension time and the related costs. More facilities now are equipped with predictive maintenance systems to generate MWOs in short time periods, which means periodical maintenance forecasting and planning strategies are now challenged by a more dynamic context. We examine these challenges and design a model to generate an optimal MWO schedule instantly based on cost analysis and real-time data processed by customized heuristic algorithms.

1 Introduction

Offshore facilities' maintenance activities are often constrained by operation time windows and weather conditions, resources' allocation and spare parts' logistics etc. To obtain efficient maintenance schedules, one should take all constraints into consideration even they are dynamic and interrelated. For offshore facilities equipped with predictive maintenance systems, more MWOs will be generated in shorter time periods based on condition monitoring data, which means periodical maintenance forecasting and planning strategies are now challenged by an even more dynamic context. MWOs thus are subjected to possible delays and suspensions due to service capacity and environmental constraints. These delays and suspensions will increase costs due to extended equipment downtime, higher labor cost, longer time logistics etc. To improve maintenance performance, one should reduce overall delay and suspension of MWO execution by proper prioritizing of MWOs under such scenarios, and it is a typical multistage discrete optimization problem which can only be practically solved by a heuristic algorithm in combination with industry domain knowledge. The rest of the paper will be organized as follows: Sect. 2 will analyze offshore facilities' MWO delays and suspensions and their cost impact as domain knowledge input, then present them mathematically for modeling. Section 3 will introduce the customized heuristic algorithm and the way in which MWO delays and suspensions can be modeled based on it. Section 4 will test the model's performance with a specifically designed dataset.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 618–627, 2020.
https://doi.org/10.1007/978-3-030-48021-9_69

2 MWO Delay and Suspension Analysis

Regarding maintenance services for critical equipment on offshore facilities, the following two constraints are often the norm:

- MWO execution sequential constraint: A rigid working process has to be followed to achieve a high level of service integrity.
- Work center occupation constraint: Often, onsite maintenance services for critical equipment are executed one at a time, and the service will not be suspended by technicians working on shifts until the service is done. The MWO is suspendable between different work centers, but the work center's ongoing job is unlikely to be suspendable. This is mostly because technical complexity requires operation continuality, moreover, logistic will be easier if services are completed in a single offshore trip.

The above two constraints fit well into the manufacturing industry's well-studied Flow Shop Schedule (FSS) problem, where n jobs have to be processed by j machines in identical order; each machine can only process one job at a time and will not be stopped until it is done (Werner 2011). In the following section, we will apply these two constraints, together with FSS methodology, to model the MWO delays and suspensions.

2.1 MWO Delay and Suspension Calculation

To start with a simple example: At the time of scheduling, we assume that we have 3 MWOs (ID: 1, 2, 3) for 3 equipment within the foreseeable future. All 3 MWOs have due dates as of now and need to be processed by 3 work centers in identical order as: (1) MWOs' planning and registration (planning), (2) resources' allocation and spare parts' logistic (logistic), (3) MWO execution and verification (execution). Given that the available man-hours at each work center are one man-hour at a time, the standard man-hours required by each MWO at each work center are given below (Table 1):

Table 1. Required man-hours for each work order at each work center

MWO ID	Standard man-hours needed for each MWO at each work center (hr)			
	Work center 1	Work center 2	Work center 3	Total
1	3	5	6	14
2	4	8	6	18
3	1	2	3	6

According to the MWO execution sequential constraints and work center occupation constraints, we only need to decide the maintenance schedule at work center 1, then the schedule for the following work center will be auto-formulated by following the two constraints. For example, if we decide that the execution sequence at work center 1 is MWO ID 3-1-2, the entire maintenance schedule is shown below (Table 2):

Table 2. MWO execution timespan at each work center

<table>
<tr><th>Work Center</th><th colspan="23">MWO execution at each work center (MWO ID is the number in the block)</th></tr>
<tr><td>1</td><td>3</td><td colspan="3">1</td><td colspan="4">2</td><td colspan="15"></td></tr>
<tr><td>2</td><td></td><td colspan="2">3</td><td></td><td colspan="5">1</td><td colspan="8">2</td><td colspan="6"></td></tr>
<tr><td>3</td><td colspan="3"></td><td colspan="3">3</td><td colspan="3"></td><td colspan="6">1</td><td colspan="2"></td><td colspan="6">2</td></tr>
<tr><td>Man-hr</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td></tr>
</table>

The total timespan for the execution of all MWOs is 23 h: MWO 3 has no delay or suspension; MWO 1 is delayed for 1 h at work center 1 but there is no suspension until it is complete; MWO 2 is delayed for 4 h at work center 1 and suspended for 1 h at work center 2, respectively work center 2 has 1 h idle time and work center 3 has 5 h idle time. If we decide that the execution sequence should be MWO ID 1, 2, 3 instead of 3, 1, 2, all the delays and suspensions will change accordingly.

Moreover, the operational efficiency at each work center is not constant due to seasonal environmental changes and geological impact on equipment maintainability, labor availability, resources' allocation and logistics, etc. For example at remote facility during wintertime, more time and resources may be needed to prepare for the repair following sudden equipment breakdown. The actual required man-hours at each work center may then be higher and may cause lower levels of operational efficiency (Faccio *et al.* 2014).

To generalize the calculation, given a maintenance schedule consisting of I MWOs to be executed on J work centers with identical order, we use the denotation shown in Table 3 and give the following calculation, according to the two constraints:

Table 3. Denotations and explanations

MWO_i	The i[th] of the I MWOs
WOC_j	The j[th] of the J work centers
MWO_i^j	MWO_i on WOC_j
$MOCR_i^j$	Maintenance operation cost rate of MWO_i^j
MOC_i^j	Maintenance operation cost of MWO_i^j
MOC_i	Maintenance operation cost of MWO_i
$MFLCR_i$	Maintenance function lost cost rate of MWO_i
$MFLC_i$	Maintenance function lost cost of MWO_i
R_i^j	Expected start time (moment) of the MWO_i^j
D_i^j	Delay time (period) of the MWO_i^j
S_i^j	Actual start time (moment) of the MWO_i^j
C_i^j	Actual complete time (moment) of the MWO_i^j
E_i^j	Operation efficiency of the MWO_i^j

(*continued*)

Table 3. (*continued*)

$MTTR_i^j$	Mean Time (period) To Repair of the MWO_i^j
AET_i^j	Actual execution time (period) of the MWO_i^j
Sus_i^j	Suspension time (period) of the MWO_i^j
T_i	Total time span (period) of MWO_i from expected to start to actual complete
$Idle_j$	Total idle time (period) of the WOC_j

Note:
On WOC_1, Sus_i^1 is the delay (or ahead) period of the expected start time R_i^j. On the subsequent work centers, when j > 1, Sus_i^j is the period of suspension of execution of MWO_i^j due to different reasons. For example, MWO_i has completed planning on WOC_1, but cannot perform logistic immediately on WOC_2 due to occupation by other MWOs. MWO delay and suspension represent different scheduling focuses. For instance, for one MWO, scheduler want it start as soon as possible, then the delay time should be minimized; in other case, scheduler want a MWO to be complete as soon as possible after it start, then the suspension time should be minimized.

According to the two MWO schedule constraints established previously and above denotations, the following functions are established:

- S_i^j is the later (larger in man-hour sense) one between C_{i-1}^j and C_i^{j-1}:

$$S_i^j = \max\left(C_i^{j-1}, C_{i-1}^j\right) \tag{1}$$

- ET_i^j is the $MTTR_i^j$ multiplied by the operation efficiency E_i^j:

$$AET_i^j = MTTR_i^j \times E_i^j \tag{2}$$

- C_i^j is the moment S_i^j add the actual execution time ET_i^j:

$$C_i^j = S_i^j + E_i^j \times MTTR_i^j \tag{3}$$

- When j = 1, Sus_i^1 on WOC_1 is the delay/ahead period based on R_i^1

$$Sus_i^1 = C_{i-1}^1 - R_i^1 \tag{4}$$

- When j > 1, Sus_i^j is equal to C_{i-1}^j subtract C_i^{j-1}

$$Sus_i^j = C_{i-1}^j - C_i^{j-1} \tag{5}$$

- T_i is the total time span of MWO_i from expected to start to actual complete, which is the sum of the actual execution and suspension time on all WOCs.

$$T_i = \sum_{j \in J} \left(AET_i^j + Sus_i^j \right) \tag{6}$$

- $Idle_j$ is the sum of $(C_i^{j-1} - C_{i-1}^j)$ when C_{i-1}^j is earlier than C_i^{j-1} for all MWOs on WOC_j. Ideally the scheduler will not let WOC_1 be idle. Then j > 1.

$$Idle_j = \sum_{i \in N} \left(C_i^{j-1} - C_{i-1}^j \right), \qquad j > 1 \tag{7}$$

The above Functions can be solved recursively given MWO permutations on the initial work center WOC_1.

2.2 Cost Analysis for MWO with Delays and Suspension

As defined by Lyonnet (2013), maintenance costs are grouped into two categories: (1) Maintenance operation costs, which include labor costs, maintenance equipment costs, spare parts' costs and total intervention costs, and (2) Loss due to a stoppage, which includes production losses, failed equipment amortization, energy consumption, etc. We will alter the second category from "production loss" to "function loss" to cover a wider range of maintenance activities. Moreover, since our goal here is to reduce delays and suspensions and the related cost, the static costs, regardless of delay and suspension, such as material and spare parts' costs, will be excluded from our calculation. For cost category (2), we only take function lost cost into consideration:

1. Maintenance operation cost (MOC): different MWOs have different operation cost rates (MOCR) at different WOCs, and the cost rate fluctuates with time. Thus, we define $MOCR(t)_i^j$ as the maintenance operation cost rate distribution function by time for the MWO_i^j. Then we have:

   - The maintenance operation cost distribution function of MWO_i^j

$$MOCR_i^j = MOCR(t)_i^j \tag{8}$$

   - The maintenance operation cost of MWO_i^j

$$MOC_i^j = \int_{S_i^j}^{C_i^j} MOC(t)_i^j dt \tag{9}$$

- The Maintenance operation cost of MWO_i is:

$$MOC_i = \sum_{j \in J} \int_{S_i^j}^{C_i^j} MOC(t)_i^j dt \tag{10}$$

Note: refer to Function 1 and 3, the integration of the maintenance cost rate function from $S_i^j$to C_i^j do not include the suspension time Sus_i^j.

2. Maintenance Function Lost Cost (MFLC) due to maintenance delay, suspension and execution: different MWO causes different level of function lost. Longer delay and suspension of a MWO at a period with higher function lost rate will cause a higher function lost cost. If we convert the function lost (whether the function is production or safety and environmental protection) to money terms, and we define $MFLCR\ (t)_i$ as MWO_i's function lost cost rate distribution function by time. and we assume that the function lost start at the moment when the MWO execution is expected to start, which is R_i; and recover to normal at the moment when the MWO is actually complete, which is $R_i + T_i$. Then we have:

- The maintenance function lost cost rate distribution function of MWO_i

$$MFLCR_i = MFLCR(t)_i \tag{11}$$

- The Maintenance function lost cost of MWO_i is:

$$MFC_i = \int_{R_i}^{R_i + T_i} MFLCR(t)_i dt \tag{12}$$

Given the MWOs' permutation on the initial WOC_1, $MOCR(t)_i^j$ and $MFLCR(t)_i$, based on the Function 1–12, we can calculate all and each MWO's delay and suspension time and related costs. Based on these calculations we can optimize MWOs permutation on WOC_1 towards the following typical MWO schedule objectives (Razali and Geraghty 2010):

- Obtain the shortest total work span to complete all MWOs in the backlog.
- Obtain the lowest function loss cost during intense production seasons.
- Obtain the shortest idle time for a particular work center if that work center is a costly outsourcing contractor.
- Obtain the shortest delay and suspension time for a MWO if it critically impacts production and safety functions.
- Obtain the lowest maintenance backlog to complete all MWOs in time.

The MWO scheduling optimization then become the optimization of the MWOs permutation on WOC_1 based on different schedule objectives. And it can be executed based on the above calculations by real-time data processing recursively and iteratively, the real-time data includes, but not limit to MWO backlog, maintenance operation cost rate distribution ($MOCR(t)_i^j$), maintenance function lost cost rate distribution ($MFLCR(t)_i$), operation efficiency rate (E_i^j), etc. The MWO scheduling optimization process is illustrated in Fig. 1 below:

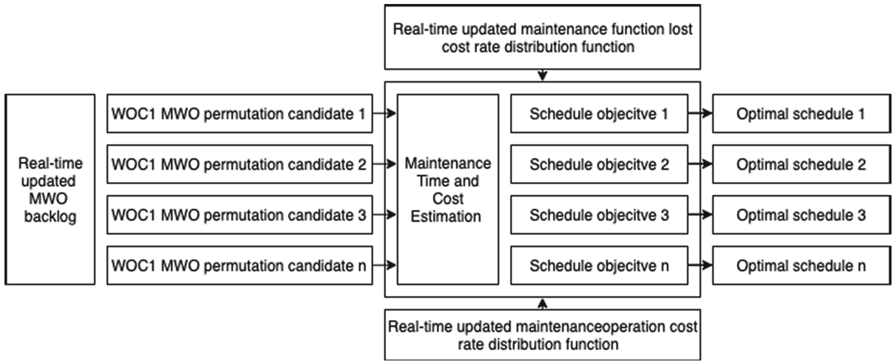


Fig. 1. MWO scheduling optimization based on real-time data

The above optimization process can be executed daily or even hourly automatically or semi-automatically, and will be used as a framework for modeling in the next section.

3 Maintenance Work Order Schedule Modeling

MWO schedule optimizations are multistage discrete optimization problems, which can be practically solved only by heuristic algorithms with industry domain knowledge. For example, Muise (2016) used a neighborhood search with a hill-climbing algorithm to find the optimal schedule for a typical flow shop schedule problem. Wang and Handschin (2000) used a genetic algorithm to find the optimal schedule for a preventive maintenance problem. In our case, we use a genetic algorithm, together with an enhanced local search, to model a MWO schedule for both preventive and corrective maintenance scenarios.

3.1 Genetic Algorithm and MWO Schedule Modeling

The genetic algorithm (GA) was originally developed by Holland (Goldberg and Holland 1988). In short, the idea is to mimic the efficient selection process of natural evolution, where the environmentally fittest chromosomes will survive from the massive chromosome populations, in which individual chromosome randomly cross over with each other and mutate to generate new chromosomes to update the population for iterative fitness selection, until convergence criteria are achieved; the process is illustrated in Fig. 2.

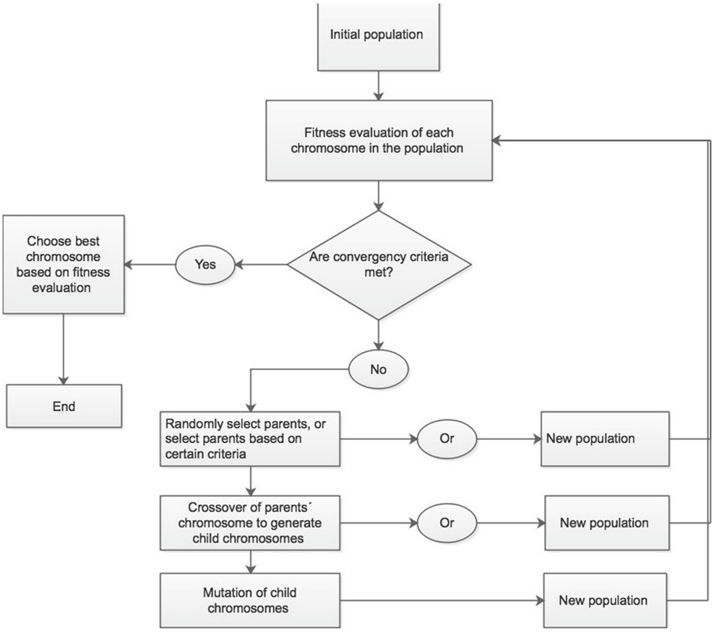


Fig. 2. Genetic algorithm procedure

By analogy, a maintenance schedule can be seen as a chromosome, and each MWO within the maintenance schedule as a gene of the chromosome. The gene (MWO) can be sequentially altered or replaced by another gene (MWO). By applying a genetic algorithm iteratively to our schedule problem, we aim to select the fittest maintenance schedule, according to our objectives, as listed in Sect. 2.2. We use MWO ID as the gene and a MWO ID permutation as a chromosome for fitness calculation against the predefined schedule objectives. Once the fittest chromosome is selected, the maintenance schedule (execution sequence at work center 1) can be decoded immediately, as shown in Table 4.

Table 4. Maintenance scheduling problem representation

Chromosome population	Genes (MWO ID)							
Chromosome 1	1	2	3	5	4	6	…	n
Chromosome 2	2	5	1	3	6	4	…	n
Chromosome 3	n	2	3	4	5	6	…	1
……	…….							
Chromosome X	1	2	3	5	6	4	…	n
MWO schedule population								

4 MWO Model Test and Demonstration

We programmed the model by Python script, based on the GA and the MWO schedule process illustrated in Fig. 1. Note that we will not elaborate on the programming and algorithm setting details, since this paper is concerned with maintenance scheduling conceptual modeling rather than computer programming. We used Microsoft Azure Machine Learning Workbench (Microsoft 2018) as the computation environment to deploy and run our model. Model input and output data were stored in the cloud. The results were demonstrated through Microsoft Power BI online visualization. The purpose of running the model and demonstrating the results on the cloud is to test the model's ability for real-time information processing and presenting. For conceptual testing, we created a test data set of 10 MWOs executed through 3 work centers. The available man-hours at each work center are 1 man-hour at a time. The data set includes the following segments (Table 5):

Table 5. Model test data set properties

Name	Property
MWO registration list	Given expected start date R_i and $MTTR_i$ for each MWO at each work center
Operation efficiency factor E_i^j	A given number range between 0 to 4 for each MWO at each work center
Maintenance operation cost rate distribution by time $MOCR(t)_i^j$	A given number range between 10 to 100 for each MWO at each work center which varied from month to month through the year
Maintenance function lost cost rate distribution by time $MFLCR(t)_i$	A given number range between 1 to 700 for each MWO which varied from month to month through the year

We ran the model according to the objectives listed in Sect. 2.2. Through 1000 iterations within eight minutes, we obtained the results as shown in Fig. 3:

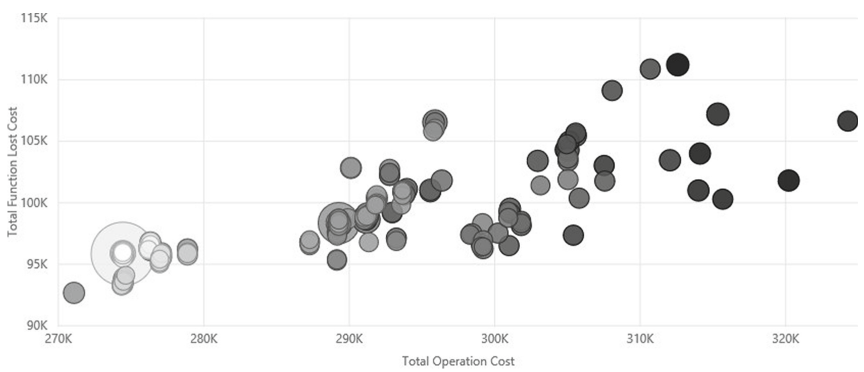


Fig. 3. MWO schedule simulation results based on genetic algorithm

- Each circle is a MWO schedule, generated by the model after one iteration; the center location of the circle represents the schedule's total maintenance operational cost and functional lost cost, as indicated on the X and Y axes respectively.
- The radius of the circle is positively correlated to the total work span of the maintenance schedule.
- The darkness of the circle is positively correlated to the total maintenance backlog of the maintenance schedule.

According to our objectives, the circle at the left most, lowest part of the chart with the smallest radius and lowest level of darkness is the favorable maintenance schedule we are looking for, since it has the lowest function lost cost and maintenance operation cost, with the smallest work span and backlog of man-hours. We can also easily choose the schedule by ranging the selection criteria from Fig. 3.

5 Conclusion

The paper presents how offshore facility maintenance performance can be improved by reducing delay and suspension times and the related cost in a dynamic environment. Two important constraints of offshore maintenance service have been introduced to build a schedule model. Based on these, we are able to generate a reasonable schedule recommendation to resolve the challenges. However, the two constraints are not universally applicable; particularly, the work center occupation constraint can be altered when the work center ongoing maintenance activities are not critical, more than one MWO can be proceed simultaneously at that work center. We will examine more complicated scenario in our further research. Finally Thank China Scholarship Council and Norwegian Research Council for supporting this research project.

References

Faccio, M., et al.: Industrial maintenance policy development: a quantitative framework. Int. J. Prod. Econ. **147**, 85–93 (2014). https://doi.org/10.1016/J.IJPE.2012.08.018

Goldberg, D.E., Holland, J.H.: Genetic algorithms and machine learning. Mach. Learn. **3**(2), 95–99 (1988). https://doi.org/10.1007/BF00113892

Lyonnet, P.: Maintenance Planning: Methods and Mathematics. Springer Science & Business Media, New York (2013)

Microsoft: Azure Machine Learning Workbench (2018). http://azureml.azureedge.net/content/apphome/index.html

Razali, N.M., Geraghty, J.: Genetic algorithms performance between different selection strategy in solving TSP (2010)

Muise, C.: A flow shop scheduler. In: 500 Lines or Less, p. 175 (2016)

Wang, Y., Handschin, E.: A new genetic algorithm for preventive unit maintenance scheduling of power systems. Int. J. Electr. Power Energy Syst. **22**(5), 343–348 (2000). https://doi.org/10.1016/S0142-0615(99)00062-9

Werner, F.: Genetic algorithms for shop scheduling problems: a survey, p. 31 (2011, preprint). http://www.math.uni-magdeburg.de/∼werner/preprints/p11-31.pdf

Integrating Maintenance into Long Term Planning

Ype Wijnia(✉)

AssetResolutions B.V., Zwolle, The Netherlands
ype.wijnia@assetresolutions.nl

Abstract. Asset management can be regarded as managing the degradation of assets by means of timing of interventions. Basically two types of interventions exist, maintaining or replacing the asset. In practice, maintenance and replacement are often optimized independently, as the timeframes and costs involved are often very different. Only at the end of asset life the disciplines connect in the choice between maintaining and replacing the asset. However, as maintenance impacts asset life, in long term planning of the asset base both should be considered in coherence. In this paper we present our simplified marginal approach for integrating maintenance in long term replacement planning considerations. The results are compared with numerical optimization by means of life cycle costing. Although the results align in general, there may be cases that a marginal approach does not provide the right result. The conditions under which this happens will require further research.

1 Introduction

Asset management can be regarded as managing the degradation of assets (Gorjian et al. 2009). Any asset will lose performance over time, and it is up to the asset manager to decide when and how to intervene. Interventions may range from maintenance fundamentals like Aligning, Balancing, Cleaning, Coating and Lubricating to end of life interventions like repairs, replacement and abandonment. Decision making on interventions is aided by inspections, monitoring, diagnostics and decision support systems. In designing the maintenance program asset managers have to make several choices, taking into account the characteristics of the assets and preferences of the asset owner:

- Maintenance regime: Preventive of corrective, and in case of preventive use based or condition based
- Maintenance tasks: which activities to perform
- Intervention timing: either time, usage or condition

Optimization of the maintenance program can be very complicated, even if only a single asset is considered. This is because the three dimensions are both internally path dependent as they are mutually interdependent. For example, if the decision is made to run an asset to its failure, it is impossible to change towards a preventive maintenance regime near the end of life because of the accumulated damages. Change of regime then requires a restoration of the asset or perhaps even a full replacement.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 628–635, 2020.
https://doi.org/10.1007/978-3-030-48021-9_70

The complexity increases further if end of life considerations are included. Basically every asset (even those properly maintained) has an end of life. Maintenance costs typically rise near this end of life. At a certain moment it may be wiser to replace the asset instead of maintaining it. But given that the asset will be replaced at a certain moment, some planned preventive maintenance may not be very valuable.

In this paper we explore the potential for integral optimization of maintenance and replacement. First we develop a simplified view on optimizing maintenance or replacement separately. This will be followed by an integration of those perspectives. These perspectives will be applied on several assets of a drinking water company. In reviewing the results we will establish whether it is feasible and valuable to integrate the perspectives, or whether organizations are better off treating maintenance and replacement separately.

2 Basic Optimization Model

Optimization is in this paper defined as finding the best possible intervention. Even though true mathematical optimization may not exist in real life, optimization can be used in a pragmatic approach to identify a range of good options (Wijnia 2016) especially for timing decisions.

In the marginal approach of such a timing decision (Frenk et al. 1997), the benefit of postponing the intervention is compared with the risk the postponement generates. The benefit roughly equals one year of depreciation given the replacement age plus interest, scaled by the probability the asset will survive the coming year. The risk of postponement is the additional cost of corrective replacement multiplied by the probability of failure. Below are the equations, with tA is the replacement age of the asset, $C_{A,I}$ as the cost of a preventive replacement, $C_{A,C}$ as the cost of a corrective replacement and r(T) the conditional survival rate for the next year.

$$Benefit(t_A) = \left(\frac{1}{t_A} + interest\right) * C_{A,I} * r(T) \tag{1}$$

$$Risk(t_A) = (C_{A,C} - C_{A,I}) * (1 - r(T)) \tag{2}$$

Because the benefit is a continuously decreasing function towards zero and risk a continuously increasing function from zero there is precisely one point where these functions are equal, this is the optimal age. Because the derivative at the optimum is zero, values around the optimum are almost as good (within 10%) as the optimum itself, indicated in Fig. 1 below by the vertical blue lines.

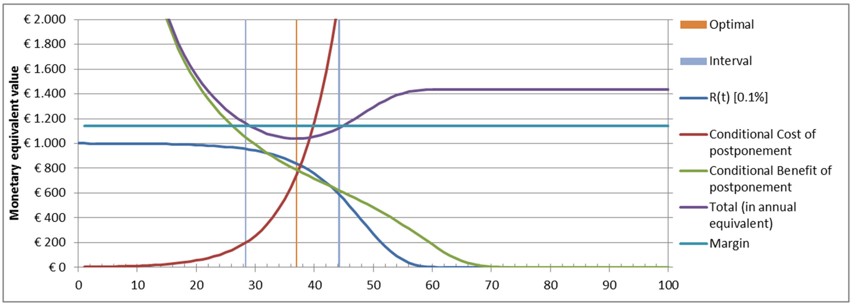


Fig. 1. Finding the optimal interval

The risk model behind this optimization is a (Weibull like) continuously growing hazard rate, estimated by two parameters: the undisturbed lifetime (T_{und}) which 95% of the assets will reach (the depreciation period typically is a good first estimate) and the maximum age for the asset (T_{max}) (Wijnia and Croon, 2017) (Fig. 2).

$$h(t) = h_0 e^{c_1 t}, R(t) = e^{-\frac{h_0}{c_1} * (e^{c_1 * t} - 1)},\ h(T_{max}) = 1,\ R(T_{und}) = 0,95 \tag{3}$$

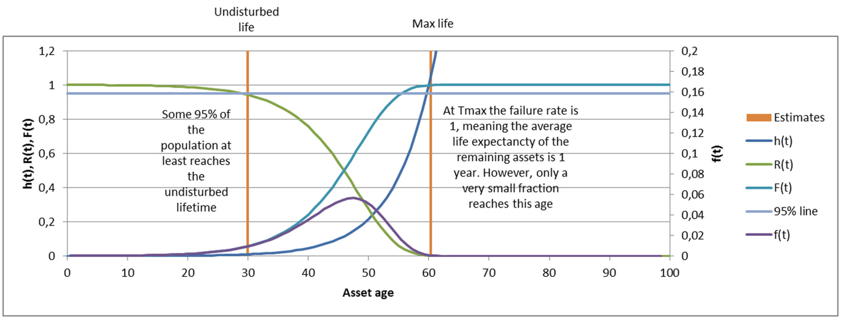


Fig. 2. Failure model based on undisturbed and maximum life

3 The Relation Between Maintenance and Ageing

The diagrams above apply to maintenance as well as they do to replacement. Both consider a deteriorating condition resulting in an increase of the risk, to be restored by an intervention (Jardine et al. 1997). The decision to intervene is an investment decision, comparing costs with benefits. However, there is a difference. Failure of (a component of) the asset will have immediate impacts, but failure of the conservation of that asset will only result in accelerated ageing of the asset. Think about a piece of steel without coating in a hostile environment, bearings without lubrication, badly aligned gearboxes, dirty and unbalanced turbines. Properly maintained life may be measured in years, unprotected it may only be days.

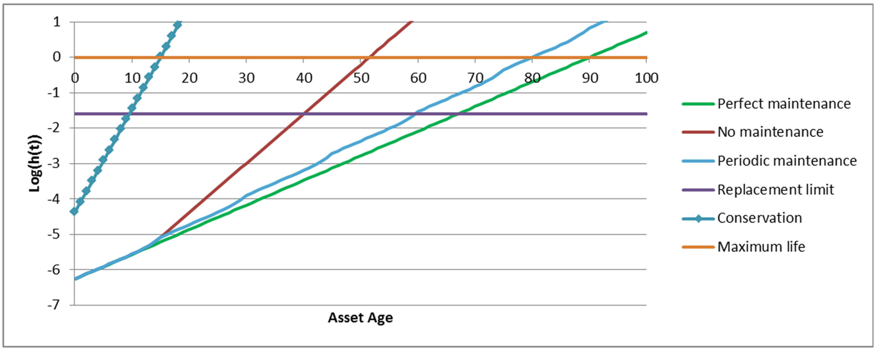


Fig. 3. Failure rate development for an asset for several conservation regimes

This can be demonstrated in the development of the conditional failure rate h(t) of an initially well conserved asset. The blue boxed line in Fig. 3 shows the failure rate over time for the conservation of the asset (= fraction of asset unprotected). The red line shows the impact on the asset failure rate. Once a significant fraction of the conservation has failed (after about 10 years), the asset ageing will increase (in the graph twice as fast). If maintenance is conducted, asset ageing will resume its normal rate, though the lifetime lost due to the accelerated ageing will not be recovered. This is shown by the stepped blue line. The green line finally shows what happens if the asset is maintained perfectly, i.e. the conservation is kept in the as new condition. Maximum asset life is defined (in a pragmatic approximation) as when the conditional failure rate reaches 1 (the orange line). Depending on the risk of failure, asset managers typically will not run the asset to failure, but will replace at a certain failure rate, indicated (for simplicity) by the purple line.

4 Combined Optimization

Now that the combined failure rate has been established it is possible to optimize on both maintenance interval and replacement interval to arrive at the lowest annual equivalent cost for the asset. First this optimization will be approached marginally (comparing the benefit of postponing the intervention with the risk this generates), similar as in optimizing the asset without considering maintenance. The second approach will be numerically, calculating the equivalent lifecycle costs for a range of combinations of maintenance and replacement intervals. The parameters for the asset and conservation are given by the Table 1 below.

Table 1. Parameters for the optimization model

Item	Asset		Conservation	
Interest = 0%	Symbol	Value	Symbol	Value
Preventive cost	$C_{A,I}$	1000	$C_{C,I}$	125
Corrective cost	$C_{A,CR}$	1500	$C_{C,CR}$	n.a.
T undisturbed	$T_{A,U}$	60	$T_{C,U}$	10
T max	$T_{A,max}$	90	$T_{C,max}$	15
Lifetime-loss per year unprotected[a]	LTL	1		

[a]This is the additional lifetime loss. If an unprotected asset ages twice as fast (lifetime 45 instead of 90 years) the lifetime-loss is 1 year per year.

4.1 Analytical Resolution by Marginal Approach

Including maintenance in this marginal approach requires a twofold optimization, both of the maintenance and the replacement interval. As the replacement interval is influenced by the maintenance regime (it is roughly when the hazard rate reaches the allowed limit for replacements) but not vice versa (except for excluding maintenance intervals longer than replacement intervals) maintenance should be optimized first.

The marginal approach to maintenance optimization would be comparing the benefit of postponing maintenance with the risk is generates. The benefit of postponing 1 year can be formulated similarly to that of postponing investments, though the survival rate of the conservation is no part of the optimization (maintenance can be applied as long as the asset exists)

$$Benefit(t_C) = \left(\frac{1}{t_C} + interest\right) * C_{C,I} \tag{4}$$

The risk of postponement however is conceptually completely different. It is not about the additional cost of corrective maintenance (for this exercise they are assumed to be zero) but in the lost life expectancy of the asset. This extra aging will mean that at the end of asset life, some book value is left to be disinvested: the lost years multiplied by the "normal" depreciation per year. As the ageing of the asset depends on the amount of conservation that has failed (F(t)), the risk can be formulated as follows:

$$Risk(t_C) = \frac{C_{A,I}}{T_{A,max}} * F(t_C)_{cons} * LTL \tag{5}$$

As the benefit decreases over time and risk increases, there is one point where they are equal. Solving for t_C gives the optimal maintenance interval. Using this interval the combined failure rate can be calculated (see Fig. 3) and the associated optimal asset replacement age. For the parameters used it results in a maintenance interval of 15 years and a replacement interval of 61 years. It is thus optimal to wait until the conservation is gone. This can be explained by the relatively high costs of conservation compared to asset value and the limited impact on asset life[1].

4.2 Numerical Resolution by Life Cycle Costing

A different approach is numerical, calculating the equivalent lifecycle costs for a limited number of combinations of maintenance and replacement intervals. Comparing the value per combination with the best possible combination reveals a relative score that can be colour coded to form an optimization map, as shown in Fig. 4.

	Repl_55	Repl_56	Repl_57	Repl_58	Repl_59	Repl_60	Repl_61	Repl_62	Repl_63	Repl_64	Repl_65	Repl_66	Repl_67	Repl_68	Repl_69	Repl_70	Repl_71	Repl_72
Main_10	118%	117%	116%	115%	114%	113%	113%	112%	111%	111%	110%	110%	110%	110%	110%	110%	110%	111%
Main_11	114%	113%	112%	111%	110%	109%	108%	108%	107%	107%	106%	106%	106%	106%	106%	106%	106%	107%
Main_12	110%	109%	108%	107%	107%	106%	105%	104%	104%	104%	103%	103%	103%	103%	103%	103%	104%	105%
Main_13	107%	106%	105%	105%	104%	103%	102%	102%	101%	101%	101%	101%	101%	101%	102%	102%	103%	104%
Main_14	105%	104%	103%	103%	102%	101%	101%	100%	100%	100%	100%	100%	101%	101%	102%	103%	104%	105%
Main_15	103%	103%	102%	101%	101%	100%	100%	**100%**	100%	100%	101%	102%	102%	103%	104%	106%	107%	108%
Main_16	102%	102%	101%	101%	100%	100%	100%	100%	101%	102%	103%	104%	106%	108%	109%	111%	113%	114%
Main_17	102%	102%	101%	101%	101%	101%	102%	102%	103%	104%	106%	107%	109%	111%	114%	115%	117%	118%
Main_18	102%	102%	102%	102%	103%	104%	105%	106%	107%	109%	110%	112%	114%	115%	117%	118%	119%	120%
Main_19	101%	101%	102%	103%	105%	106%	108%	110%	112%	113%	115%	117%	118%	119%	120%	121%	121%	121%
Main_20	101%	101%	102%	103%	105%	107%	110%	113%	115%	117%	119%	120%	121%	121%	122%	122%	122%	122%

Fig. 4. Optimization map of maintenance and replacement interval

The optimal values are boxed and in bold typeface, corresponding to a maintenance interval of 15 years and a replacement interval of 62 years. As expected the optimum (indicated by the bold typeface) is not very sharp. Replacing 2 years early or late results in the same rounded score (i.e. less than 0,5% deviation), indicated by the boxed values. With regard to maintenance the sensitivity is larger, but even then a deviation of one year still results in near-optimal scores. The diagram also shows the inverse relation between maintenance and replacements, visible in the slightly tilted green band. Shortening the maintenance interval allows for lengthening the replacement interval and vice versa. The results of this numerical optimization align very well with the results from the marginal approach, as they fall within the optimal region (boxed area).

4.3 Considerations

The two proposed methods align in their results. However, given the ratio of cost of optimal conservation (125/15 years = 8.3/yr) versus the cost difference between two extremes in asset replacement (1000/42 years and 1000/68 years respectively => maximum gain 9.1/yr) it can be discussed whether it is not better to skip maintenance altogether. The numerical approach is therefore expanded to a full scan for all combinations. This is shown in Fig. 5 (please note the axis have been reversed):

---

[1] Omitting interest results in the following relation: $t_C * F(t_C) = \frac{C_{C,I}}{C_{A,I}} * \frac{T_{A,max}}{LTL}$. A high value on the right hand side implies a relatively long maintenance interval.

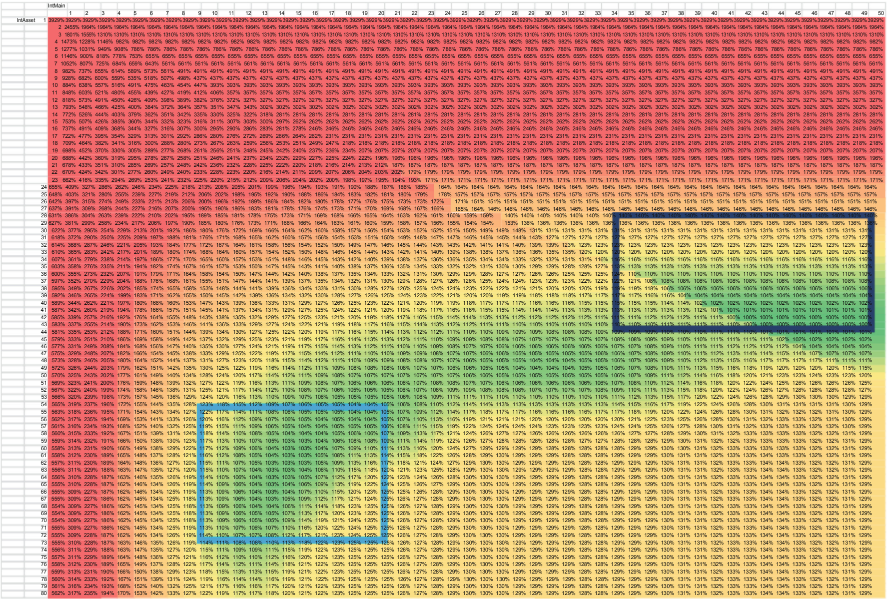


Fig. 5. Full calculation of costs

This figure shows the optimal combinations of maintenance and replacement indeed include an area without any maintenance (dark blue), which is as good as the found values in the marginal optimization and the limited numerical approach (light blue). Both are local optima, and neither can be approached from the other point by marginal improvement. Starting a marginal approach from the perspective of a perfectly conserved asset will find an optimal combination, but starting from a no-maintenance policy will not. Maintaining right before replacement only adds cost and thus the marginal approach will stop.

5 Discussion and Conclusion

In this paper it was shown that maintenance can be optimized by means of a marginal approach that takes the influence on asset life into account. The resulting degradation curve then can be used to optimize asset replacement. This approach provides basically the same results as calculating the lifecycle costs for a number of combinations of maintenance and replacement intervals. Because of its analytical nature the optimization effort is much less, which allows maintenance to be included in long term asset replacement plans. However, for certain combinations of parameters, the optimum by the marginal approach is not necessarily the best; it may be wiser not to maintain the asset and replace it in a higher rate. At this moment it is not entirely clear under what conditions the marginal approach gives an adequate answer, nor whether the region in which it does not is common in practice. Further research will be needed to clarify these points.

References

Frenk, J.B.G., Dekker, R., Kleijn, M.J.: On the marginal cost approach in maintenance. J. Optim. Theory Appl. **94**, 771–781 (1997)

Gorjian, N., Ma, L., Mittinty, M., Yarlagadda, P., Sun, Y.: A review on degradation models in reliability analysis In: Kiritsis, D.E.C., Koronios, A., Mathew, J. (eds.) 4th World Congress on Engineering Asset Management. Springer, Athens (2009)

Jardine, A.K.S., Banjevic, D., Makis, V.: Optimal replacement policy and the structure of software for condition-based maintenance. J. Qual. Maintenance Eng. **3**, 109–119 (1997)

Wijnia, Y., Croon, J.D.: Strategic asset planning: balancing cost, performance and risk in an ageing asset base. In: The 12th World Congress on Engineering Asset Management (WCEAM 2017). Springer, Brisbane (2017)

Wijnia, Y.C.: Processing risk in asset management: exploring the boundaries of risk based optimization under uncertainty for an energy infrastructure asset manager. Ph.D., Delft University of Technology (2016)

Intelligent Maintenance Practices Within Norwegian Continental Shelf Toward Industry 4.0 Vision: An Overview

Helge Nordal(✉) and Idriss El-Thalji

University of Stavanger, Stavanger, Norway
{Helge.nordal, Idriss.el-thalji}@uis.no

Abstract. Industry 4.0 is the future scenario of industrial production since it enables a new level of organising and controlling the entire value chain with the product lifecycle by creating dynamic and real-time understanding of cross-company behaviours. Therefore, most of the oil and gas (O&G) companies at the Norwegian continental shelf (NCS) are looking forward to digitalise and automatize their asset and operations in order to get the promising benefit of industry 4.0. There are recommendations from the ministry of industry toward the industrial digitalisation, together with commercial incentives taken by large-scale oil and gas operating companies. The holistic picture of the status of industry 4.0 at the Norwegian industrial sector is still missing. Industrial managers at the NCS, who are either ready for industry 4.0 or already progressing toward this industrial transition, are still lacking answers to several significant questions. Industry 4.0 originated in manufacturing sector and most of the current experiences are related to that specific sector, thus, can industry 4.0 technologies be useful for companies of operation and maintenance service providers? What would be the specific requirements to revolutionise current maintenance into more intelligent maintenance? Consequentially, the paper tries to present whether complete solutions (i.e. product availability and business models) of intelligent maintenance are available and how mature they are.

Keywords: Industry 4.0 · Cyber-Physical-Systems (CPS) · Intelligent maintenance, oil and gas industry · Norwegian industry

1 Introduction

The core concept of Industry 4.0 is the *cyber-physical-system* (CPS), which convergence the cyber space with the physical space (Lee et al. 2015). One of its main enabler is the technology concerning *Internet-of-things (IoT)* and successively the opportunities within the concepts of *big data, data mining, and Internet-of-service (IoS)*. The initiative of British Petroleum (BP) toward digitalised operation and maintenance operations was illustrated in (World_Economic_Forum 2017) where the digitalisation efforts could reduce the upstream operations' finding and development costs by 5%, maintenance costs by 20%, overtime cost by 20%, downtime by 5% (mainly due to predictive maintenance), inventory levels for spare parts by 20% and boost production by a conservative 3% (conventional land operations).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 636–645, 2020.
https://doi.org/10.1007/978-3-030-48021-9_71

In the Norwegian context, there are recommendations from the Norwegian ministry of industry (Det_Kongelige_Nærings-og_fiskeridepartementet 2017; Norsk_industri 2017) toward the industrial digitalisation, together with commercial incentives taken by large-scale oil and gas operating companies. The real applications show that design for intelligent production/operation is rapidly growing in the Norwegian industry, particularly, the oil and gas sector is pioneering the technological development e.g. unmanned platform, automate drilling rigs, remotely operated vehicle (ROV), and automated substation. However, the maintenance as a supportive function for the production and operation activities are not having the same developmental growth. This is a general observation in other industrial sectors that have progressed toward industry 4.0 vision. Bokrantz et al. (2017) highlighted that in both scientific and business literature on digitalised manufacturing, maintenance is barely even mentioned, or is perceived rather narrowly. Bokrantz et al. (2017) studied around 34 future projections of maintenance in digital industry. In Norway, some of those projections are already in the research and development phase (Marhaug and Schjølberg 2016; Rødseth et al. 2017).

However, the experts of industry 4.0 listed several challenges that is facing the implementation of industry 4.0 in general. Kagermann et al. (2013) highlighted the lack of a unified standard and reference architecture design, lack of models for processes/work organization, lack of new business models, and lack of product availability. On other hand, several other smart factory technologies, modelling and simulation techniques (Oesterreich and Teuteberg 2016) are not fully adopted with present progress. Moreover, predictive maintenance and operation optimization are the main technologies that will have the highest industrial and societal impacts (World_Economic_Forum 2017). Probably these are the opportunities and challenges that are also related to intelligent maintenance, thus, this paper provides a brief overview for each challenge and opportunity respectively.

2 Challenge 1: Standards and Reference Architecture Design for Intelligent Maintenance Systems

The standards, patents (Trappey et al. 2016, 2017), technologies and applications (Lu 2017) related to industry 4.0 are reviewed. It is obvious that most of the literatures describe industry 4.0 as an architecture of several layers. The most well-known architecture for industry 4.0 is RAMI 4.0 (VDI/VDE 2015; Zezulka et al. 2016) which consists of six layers. These associations behind RAMI 4.0 model are working hard with the International Standardisation Organization to create a standard for industry 4.0. Regards maintenance, there are some architectures that are active and leading the intelligent maintenance development in industry 4.0 context. For example, Watchdog Agent platform has been further developed to compile with industry 4.0 technologies (Bagheri et al. 2015), where 5C model (connection, conversion, cyber, cognition, configuration) was proposed (Lee et al. 2015) to illustrate the development of intelligent maintenance architecture and layers. Since the 5C's model excluded the business layer (the sixth layer in RAMI 4.0), a 8C's model has been proposed as an update to the 5C's model to cover customers, content, and coalition (Jiang 2017). However, the history of intelligent maintenance is far back.

The first 'obstacle' that in most cases is to be highlighted in both an academic and industrial context, is to agree upon a common terminology that reflects maintenance in the era of Industry 4.0, to facilitate interoperability and collaboration. Such common terminology is currently lacking and is in different papers referred to as e.g. intelligent maintenance, maintenance 4.0, smart maintenance, digital maintenance, emaintenance and deep learning maintenance etc. However, this paper is referring to maintenance executed in the era of Industry 4.0 as intelligent maintenance, reflecting a concept whereas proactivity is central. Ideally, this strategy means that the maintenance allocation is mainly based on an analytical approach, whereas corrective- and preventative maintenance is either minimized or neglected. Even though, the terminology "maintenance 4.0" is rarely used, the development toward smart or intelligent maintenance has long history. Perhaps, the most representative terminology for maintenance 4.0 is e-Maintenance (Han and Yang 2006; Muller et al. 2008; Adgar et al. 2008; Iung et al. 2009; Candell et al. 2009; Mascolo et al. 2010; Guillén et al. 2013; Holgado et al. 2016; Ferreiro et al. 2016; Jantunen et al. 2017). The e-Maintenance covers the functions from data to decision, which covers only the digital functions. However, the disruptive technologies (e.g. advanced robots, drones, and wearables) of industry 4.0 might enable maintenance to physically act in a smarter manner. Therefore, it expands the scope of maintenance intelligence to cover the physical functions. It is worthy to mention that even thought of the literature trace the development of e-Maintenance to 2006, the development was based on several incremental contributions (illustrated perfectly in Muller et al. (2008)). The process from data acquisition into decision has rapidly developed due to sensor technologies, signal-processing techniques, diagnostics and prognostic approaches, maintenance optimisation and decision making models. Several well-known techniques and algorithms have been standardised in six blocks of functionality in "condition monitoring and diagnostics of machines" standard (ISO_3374-1). In fact, MIMOSA (MIMOSA 1998–2018) created the open system architecture for condition-based maintenance (OSA-CBM) as implementation of the ISO 13374 (Thurston 2001). The six blocks of functionality are data acquisition, data manipulation, state detection, health assessment, prognostic assessment, and advisory generation. This open source standard and platform with other contributions have leaded several customised platforms for smart maintenance to be developed in the period between 2000–2008. For example, Watchdog Agent (Djurdjanovic et al. 2003; Lee et al. 2006), PROTEUS (Bangemann et al. 2006), SIMAP (Garcia et al. 2006), Siemens SIMATIC/S7 SPS (Groba et al. 2007) were among the most popular platforms for CBM. Some of these platforms are still active and leading the intelligent maintenance development in industry 4.0 context.

In summary, in the physical space, the asset layer (e.g. physical machines and human workers) will generate data related to maintenance (health and performance measurements, descriptive notifications and reports) which will be acquired (perception layer e.g. sensors, controller) and transmitted into the cyber space, as shown in Table 1.

Table 1. *Cyber-physical maintenance system*

Space	Layers	Main functions
Physical space	Business	Production, operations, maintenance, supply chair, marketing
	Asset and maintenance operations/executions	Platform, Pump, operators, maintenance staff
	Info/data perception	Data acquisition e.g. Vibration sensor, reports, notifications
Transmission	Data between physical/cyber spaces	Communications, networks
Cyber space with interface to physical space	Cyber space	Cloud solution
	Conversion (data)	Data manipulation
	Computation (information)	State detection, descriptive analytics
	Cognition (knowledge)	Diagnostic, predictive and prescriptive analytics, Health and prognostic assessment,
	Support decision making (optimised solutions)	Maintenance optimisation
	Maintenance management	Maintenance programme planning, capacity planning, spare part planning, scheduling
	Configuration	User interfaces, automatic actions

In the cyber space, several analytical functionalities will be performed: conversion (data per-processing, big data analytics), computations (state detection analytics, performance analytics), cognition (health diagnosis and prognosis assessment analytics, maintenance optimisation and decision support analytics, maintenance management analytics), and configuration (e.g. user interfaces, automatic actions, actuators). The configuration layer shall enable to transmit (via the transmission layer again) the required actions from cyberspace to physical space. The whole process looks like a loop from physical to cyber to physical space again.

However, in order to gain the expected benefits contributed by the employment of intelligent maintenance, a rigid common standardized reference architecture design, which is currently lacking, must be developed.

3 Challenge 2: Product Availability

Intelligent maintenance in industry 4.0 context requires several products at each previously described layers (illustrated in Table 1). First, the asset shall be smart to provide data (asset and perception layers) about its behaviour in terms of performance, health, context and cost effectiveness indicator. Second, the intelligent maintenance system shall be smart to process and utilise (cyber analytic layers) the collected data to support the decision making process. Third, the asset shall be smart in receiving/acting (configuration- and asset-layer) according to the recommended actions extracted based on the collected data. Respectively, there exist huge amount of industrial solutions/products providing cost effective perception which mainly include the controllers e.g. SCADA and sensors.

The O&G sector, as other sectors, comprises a large amount of systems and components, both old and new. This has urged for correspondingly many products, in order to facilitate the employment of intelligent maintenance system. For example, an offshore oil rig includes approximately 30 000 sensors (Manyika et al. 2015) and modern platforms approximately 80 000 sensors (World Economic Forum 2017), including several product suppliers providing industrial perception and control system e.g. Emerson, Siemens, and ABB. However, the vast majority of those sensors are related to detecting the operational anomalies and control (Manyika et al. 2015), and not for the health parameters. For sure, some of those operational parameters can be used for detecting faults and failures, but the utilisation is limited and indirect. For health monitoring purposes, there are several specialised product and service providers providing products related to health perception e.g. SKF, B&K, and National instruments with several techniques e.g. vibration, acoustic emission, ultrasound, thermal, and oil-debris. It is worthy to mention that the industrial experiences of those companies lead them to utilise the internet of things to move toward service-dominant logic and sell cyber-analytical service instead of only physical products e.g. measurements kits and control system. It is clear that companies are moving to cover the whole layers of industry 4.0 and intelligent maintenance. However, the coverage of all required layers of analytics might vary from one company-solution to another. Basically, taking a maintenance decision requires a managerial analytics that utilize technical recommendations/analytics (diagnostic and predictive analytics). Some companies provide the perception and connection layers for the operators where they can receive and access the collected data in the cloud. The operator might need to outsource multiple monitoring solutions since several monitoring companies provide solution specific to the asset type e.g. rotating, structural, electrical, and chemical and rarely to find one solution fits all.

However, the operators need to analyse the data by other specialised analytical products to determine the present status (diagnostic) and predict the future status (prognostics). Therefore, the diagnostic analytics shall provide good assessment of the process performance and asset health (healthy, and if faulty the fault type, severity, location are diagnosed). Such technical analytics helps the managers to take decision whether to act or not. However, if the maintenance manager wants to look if there is any opportunity to find a future optimal solution by waiting until next opportunistic production stop e.g. low production season to perform the required maintenance, then predictive analytics is needed. Predictive analytics shall estimate the useful remaining life for the faulty asset, which might provide sufficient time in hand to plan and optimize the maintenance action. Even though, the predictive analytics products and services are rapidly popping up, especially the open sources where data-driven approach are enabled e.g. data clustering, data classification and pattern recognition, the health parameters are still hard to be predicted, specifically, when the fault symptoms and loading conditions are changing over the time.

In summary, the companies might vary in their technical coverage for the state detection, diagnostic and predictive analytics, which also depends on the operator's need, desired operational scenario and the managerial analytics e.g. maintenance optimization, that are in use. The maintenance optimization are usually performed in locally-develop products e.g. software and excel sheets. Assume the maintenance action is optimized, then it should be planned, scheduled and ordered via enterprise resource planning system (ERP) or more specifically computerised maintenance management system (CMMS). This process has several available products e.g. SAP, IFS and Infor, to link the operations with the planning and management team in a closed loop.

Last type of products that intelligent maintenance requires is the products related to connection, cyber space and conversion analytics. These products might be part of the perception and cyber analytical solution (e.g. SKF IMX8 and @ptitude). However, largescale companies have several requirements related to these products: (1) they want to develop their own data-pool or -lake, (2) they want to determine their way of sharing the data and defining the data ownership and (3) they want to secure the data within the local or global clouds and related physical asset. Such requirements are clearly related to industrial internet of things (IIoT) where large scale and high-secure solution are required. However, the industrial practices tend toward centralised solutions e.g. one data lake, which might not be the cost effective solutions since it requires high investment and probably high operating cost.

Even though, all products listed are quite smart in performing their specific functions, the higher level of smartness is to use what you have smartly collected and progressed. Among the data which is collected from 30,000 sensors in an offshore oil and gas platform, one percentage was used to support decision-making process (Manyika et al. 2015; Lee et al. 2014). For example, (Rio et al. 2015) highlighted that time based maintenance strategy is not satisfying enough to avoid occurrence of the majority, 82%, of asset failures (as only 18% of asset failures are associated with age). That can define what smartness in the maintenance world could be. It is one of the great expected underlying business drivers associated with the implementation of an intelligent maintenance system.

4 Challenge 3 & 4: Processes/Work Organization and New Business Models

The above description of how intelligent maintenance and required products work, clearly illustrates the challenge of managing such complex and interconnected process. Thus, implementation of intelligent maintenance is expected to change the traditional process of work. Instead of operators executing preventative maintenance based on schedule in terms of time in operation or numbers of cycles and corrective maintenance whenever a fault has occurred, they are expected to contribute more prior to the maintenance execution phase. This includes activities in terms of analyzing health parameters and maintenance optimization, referring to the activity whereas existing maintenance schedule is optimized based on the findings from the big data monitored (Boman 2017). Moreover, the decision-making related to maintenance is thus expected to be made at operator level.

The decision-support generated by the health parameters monitored is expected to play an essential part in documentation associated with contractual and regulatory agreements. This will improve insight all down to component level and the tracking of the organizational processes and work, successively enhance the capabilities of supervision control (Kagermann et al. 2013).

Moreover, is the change in current work process/organization expected to stimulate to new business models. These aspects contribute to the need of a more horizontal integrated business model, whereas the decision making is becoming decentralized, relying on operational health parameters monitored and consecutively the results from the diagnostic and prognostic. It is clear that operators need to manage the business interactions with manufacturer of the control system, condition monitoring providers (it could be multiple providers related to rotating, structural and electrical systems), cyber-system service providers, diagnostic and prognostic service providers, maintenance optimization providers, CMMS providers, and maintenance support service providers. The interoperability is a significant issue in such complex business model, however, the cyber-interoperability is a new issue that still is on the discussion table for many companies in terms of security and data ownership.

5 Opportunistic Technologies for Intelligent Maintenance

Oesterreich and Teuteberg (2016) highlighted three groups of technologies that have higher impact on industry 4.0 vision: smart factory technology, data science and big data technologies, and modelling and simulation technologies. In more detail, relating respectively to technologies such as the application of robotics in which might lead to cut the maintenance fixed cost e.g. drone to inspect wind turbine blade, big data and cloud computing facilitating for e.g. accurate diagnosis and prognosis of equipment, and e.g. augmented reality enabling remote supervision.

The enhanced skills needed to convert the big data generated by the monitored health parameters into useful information will, most likely, increase the opportunities of current supplier's value chain in the O&G sector. Moreover, are the future suppliers expected to deliver products with value-added services instead of just the product itself – hence, continuous profit generation compared to the traditional onetime CAPEX (Lee et al. 2014). For instance, it is expected that a supplier of a gas-compressor to an offshore facility, which designs and manufactures the equipment, also monitors its future operational health parameters, and conducts diagnostics, prognostics, and maintenance optimization of equipment (Hannovermesse 2017).

6 Conclusion

In this paper, we have given a brief overview of the expected challenges and opportunities related to intelligent maintenance, respectively. We have discussed the main challenge regarding current lack of a common standardized reference architecture design for deployment of an intelligent maintenance system. Future architecture able to fulfil the requirements of Industry 4.0 is expected to become either the presented RAMI4.0-model or MIMOSA's OSA/CBM (especially after MIMOSA announced collaboration with SAP) (Johnston 2016).

Moreover, are the other challenges associated with current product availability, process/work organization and future business models discussed. Unfortunately, are current products related to intelligent maintenance at the NCS not, solely, capable of fulfilling the requirements of Industry 4.0. This has urged for final solutions comprising several suppliers and products. Additionally, are the companies operating at the NCS not applying one single solution for specific single equipment, thus having several various solutions for the same equipment, whereas only their geographical location is different. This arises additional challenges with regards to e.g. stake-holder's accessibility and ownership of data lakes, along with its required security. The practice of an intelligent maintenance system is expected to change current work process/organization and stimulate for new business models. Firstly, the maintenance-personnel is expected to contribute more prior to the execution-phase, which include activities in terms of analysing health parameters (diagnostics and prognostics) and maintenance optimization. Hence, changing the business models to become more horizontal integrated, whereas the decision-making is either made at a decentralized location or at operator-level onsite. Moreover, play an essential part in documentation and improve insight all the way down to component-level. The associated opportunistic technologies for intelligent maintenance are considered to be smart factory technology, data science and big data technologies, and modelling and simulation technologies. In more detail, relating respectively to technologies such as the application of robotics in which might lead to cut the maintenance fixed cost e.g. drone to inspect wind turbine blade, big data and cloud computing facilitating for e.g. accurate diagnosis and prognosis of equipment, and e.g. augmented reality enabling remote supervision.

Additionally, the deployment of an intelligent maintenance system is expected to stimulate to improved operational availability and safety, along with maintenance optimization and thus cost savings. However, a survey of state-of-practice of intelligent maintenance at the NCS is recommended executed in order to study the smartness of current maintenance management.

References

Adgar, A., Arnaiz, A., Jantunen, E.: Challenges in the development of an E- maintenance system. IFAC Proc. Vol. **41**, 257–262 (2008)

Bagheri, B., Yang, S., Kao, H.-A., Lee, J.: Cyber-physical systems architecture for self-aware machines in Industry 4.0 environment. IFAC-PapersOnLine **48**, 1622–1627 (2015)

Bangemann, T., Rebeuf, X., Reboul, D., Schulze, A., Szymanski, J., Thomesse, J.-P., Thron, M., Zerhouni, N.: PROTEUS—creating distributed maintenance systems through an integration platform. Comput. Ind. **57**, 539–551 (2006)

Bokrantz, J., Skoogh, A., Berlin, C., Stahre, J.: Maintenance in digitalised manufacturing: Delphi-based scenarios for 2030. Int. J. Prod. Econ. **191**, 154–169 (2017)

Boman, K.: <Filling in the gaps - OE Digital.pdf>. OE & AtComedia, LLC (2017)

Candell, O., Karim, R., Söderholm, P.: eMaintenance—information logistics for maintenance support. Robot. Comput. Integr. Manuf. **25**, 937–944 (2009)

DET_KONGELIGE_NÆRINGS-OG_FISKERIDEPARTEMENTET: Industrien – grønnere, smartere og mer nyskapende (2017)

Djurdjanovic, D., Lee, J., Ni, J.: Watchdog Agent—an infotronics-based prognostics approach for product performance degradation assessment and prediction. Adv. Eng. Inform. **17**, 109–125 (2003)

Ferreiro, S., Konde, E., Fernández, S., Prado, A.: INDUSTRY 4.0: predictive intelligent maintenance for production equipment. In: Third European Conference of the Prognostics and Health Management Society 2016, Bilbao, Spain (2016)

Garcia, M.C., Sanz-Bobi, M.A., del Pico, J.: SIMAP: intelligent system for predictive maintenance. Comput. Ind. **57**, 552–568 (2006)

Groba, C., Cech, S., Rosenthal, F., Gossling, A.: Architecture of a predictive maintenance framework. In: International Conference on Computer Information Systems and Industrial Management Applications, Elk, Poland (2007)

Guillén, A.J., Gómez, J.F., Crespo, A., Guerrerro, A., Sola, A., Barbera, L.: Advances in PHM application frameworks: processing methods, prognosis models, decision making. Chem. Eng. Trans. **33** (2013)

Han, T., Yang, B.-S.: Development of an e-maintenance system integrating advanced techniques. Comput. Ind. **57**, 569–580 (2006)

HANNOVERMESSE: Smart Factory: The Technology behind Industry 4.0. Hannovermesse.de: Hannovermesse (2017). http://www.hannovermesse.de/en/news/smart-factory-the-technology-behind-industry-4.0.xhtml. Accessed 10 Feb 2018

Holgado, M., Macchi, M., Fumagalli, L.: Value-in-use of e-maintenance in service provision: survey analysis and future research agenda. IFAC-PapersOnLine **49**, 138–143 (2016)

ISO_3374-1 Condition monitoring and diagnostics of machines – Data processing, communication and presentation – Part 1: General guidelines. International Organization for Standardization, Geneva, Switzerland

Iung, B., Levrat, E., Marquez, A.C., Erbe, H.: Conceptual framework for e- Maintenance: illustration by e-Maintenance technologies and platforms. Ann. Rev. Control **33**, 220–229 (2009)

Jantunen, E., Zurutuza, U., Albano, M., Orio, G.D.: The way cyber physical systems will revolutionise maintenance. In: 30th Conference on Condition Monitoring and Diagnostic Engineering Management, COMADEM 2017. University of Central Lancashire, UK (2017)

Jiang, J.-R.: An improved cyber-physical systems architecture for Industry 4.0 smart factories. In: Meen, P.L.E. (ed.) Proceedings of the 2017 IEEE International Conference on Applied System Innovation, IEEE-ICASI (2017)

Johnston, A.: SAP Joins MIMOSA to Support Interoperability Standards. MIMOSA.org: MIMOSA (2016). http://www.mimosa.org/articles/sap-joins-mimosa-support-interoperability-standards. Accessed 20 Feb 2018

Kagermann, H., Wahlster, W., Helbig, J.: Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Final report of the Industrie 4.0 Working Group. Acatech–National Academy of Science and Engineering (2013)

Lee, J., Bagheri, B., Kao, H.-A.: A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manuf. Lett. **3**, 18–23 (2015)
Lee, J., Kao, H.-A., Yang, S.: Service innovation and smart analytics for Industry 4.0 and Big Data environment. Procedia CIRP **16**, 3–8 (2014)
Lee, J., Ni, J., Djurdjanovic, D., Qiu, H., Liao, H.: Intelligent prognostics tools and e-maintenance. Comput. Ind. **57**, 476–489 (2006)
Lu, Y.: Industry 4.0: a survey on technologies, applications and open research issues. J. Indust. Inf. Integr. **6**, 1–10 (2017)
Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., Aharon, D.: The Internet of Things: Mapping the Value Beyond the Hype Executive Summary. McKinsey & Company
Marhaug, A., Schjølberg, P.: Industry 4.0 and smart maintenance: from manufacturing to subsea production systems. In: International Workshop of Advanced Manufacturing and Automation (IWAMA 2016), University of Manchester, Manchester, UK (2016)
Mascolo, J., Nilsson, P., Iung, B., Levrat, E., Voisin, A., Garramiola, F., Bellew, J.: Industrial demonstrations of E-maintenance solutions. In: Holmberg, K., Adgar, A., Arnaiz, A., Jantunen, E., Mascolo, J., Mekid, S. (eds.) E-maintenance. Springer, London (2010)
MIMOSA: Open System Architecture for Condition-Based Maintenance (1998–2018). http://www.mimosa.org/mimosa-osa-cbm
Muller, A., Crespo Marquez, A., Iung, B.: On the concept of e-maintenance: review and current research. Reliability Eng. Syst. Saf. **93**, 1165–1187 (2008)
NORSK_INDUSTRI: VEIKART FOR TEKNOBEDRIFTENE (2017)
Oesterreich, T.D., Teuteberg, F.: Understanding the implications of digitisation and automation in the context of Industry 4.0: a triangulation approach and elements of a research agenda for the construction industry. Comput. Ind. **83**, 121–139 (2016)
Rio, R.: Proactive Asset Management with IIoT and Analytics (2015). https://www.arcweb.com/blog/proactive-asset-management-iiot-analytics
Rødseth, H., Schjølberg, P., Marhaug, A.: Deep digital maintenance. Adv. Manuf. **5**, 299–310 (2017)
Trappey, A.J.C., Trappey, C.V., Govindarajan, U.H., Sun, J.J., Chuang, A.C.: A review of technology standards and patent portfolios for enabling cyber-physical systems in advanced manufacturing. IEEE Access, **4**, 7356–7382 (2016)
Trappey, A.J.C., Trappey, C.V., Hareesh Govindarajan, U., Chuang, A.C., Sun, J.J.: A review of essential standards and patent landscapes for the Internet of Things: a key enabler for Industry 4.0. Adv. Eng. Inform. **33**, 208–229 (2017)
VDI/VDE: Reference Architecture Model Industrie 4.0 (RAMI4.0): Status Report. VDI/VDE Society Measurement and Automatic Control (GMA) (2015)
WORLD_ECONOMIC_FORUM: Digital Transformation Initiative Oil and Gas Industry (2017)
Zezulka, F., Marcon, P., Vesely, I., Sajdl, O.: Industry 4.0 – An Introduction in the phenomenon. IFAC (International Federation of Automatic Control), pp. 8–12. Elsevier (2016)

A Predictive Maintenance Approach Toward Industry 4.0 Machines

Sangje Cho(✉), Gökan May(✉), and Dimitris Kiritsis(✉)

EPFL, ICT for Sustainable Manufacturing, EPFL SCI-STI-DK,
1015 Lausanne, Switzerland
{sangje.cho,gokan.may,dimitris.kiritsis}@epfl.ch

Abstract. Since the cycle of development of new products gets shorter and shorter, managing Product Life Cycle becomes more important and complex. The new generation products have more functionalities and sub-assemblies compared to the previous ones on applying enabling technologies. However, machine learning techniques have limitations in the initial phases of the development of new generation products as Middle-Of-Life (MOL) data is insufficient. Guided by this challenge, the study addresses predictive maintenance issues of new generation products and presents a conceptual model providing the main elements of a data sampling approach to overcome the lack of MOL data at the initial phases of predictive maintenance applications. Doing this, we aim at facilitating the information search relevant to the maintenance recommendation and providing an optimized maintenance management considering insufficient data of further functionalities of new generation products. In that regard, the implementation of the proposed approach could help manufacturing companies (i) address knowledge representation, exploitation, openness and diffusion in the maintenance domain, (ii) overcome insufficient MOL data using manufacturing intelligence, (iii) enable predictive maintenance strategies at the initial phase of application, and (iv) improve innovative maintenance plan using enabling technologies. Accordingly, the study paves the way for predictive manufacturing to early-predict equipment condition and make optimized recommendations for adjustments and maintenance to ensure normal operations.

Keywords: Predictive maintenance · Industry 4.0 · Machine learning · Big data · Ontology

1 Background

Currently, due to the development of emerging enabling technologies such as Radio Frequency Identification (RFID), Micro-Electro-Mechanical Systems (MEMS), Product Embedded Information Devices (PEID) and various types of sensors, it is becoming more convenient and also efficient to monitor the status of systems and products. On that vein, the recent trend of implementing predictive maintenance opens an efficient way to management of industrial production systems using applied sensor technologies, and such maintenance would contribute to the improvement of the

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 646–652, 2020.
https://doi.org/10.1007/978-3-030-48021-9_72

reliability of manufacturing systems and the reduction of unnecessary subsequent maintenance.

In this context, product maintenance is defined as all the technical and managerial actions taken during MOL phase to maintain or restore the required functionality of a product [1]. Toward this aim, predictive maintenance is performed based on the actual condition and trends, acting as one of the most effective proactive maintenance policies to carry out maintenance tasks before the occurrence of faults or problems, whereas the other proactive maintenance policy called preventive maintenance is time-oriented maintenance in which tasks are performed at predetermined intervals. As a result, implementing predictive maintenance enables the system to have high reliability with low cost, leading to industrial benefits such as more return on investment, reduction in maintenance costs, elimination of breakdowns, reduction in downtime and increase in production, because of minimized inventory, spare parts and overtime costs [2].

Enabling machine learning techniques support optimization of maintenance tasks in terms fault detection, diagnostics, and prognostics. Furthermore, machine learning involves algorithms based on statistical parameters to find data patterns automatically and for enhanced computational statistics. Nevertheless, it is challenging to find such patterns and thus computational statistics have limitations for new generation products as there are no available data for further functionalities or subassemblies. To overcome this problem, the approach proposed in this study suggests that not only sensor data from the previous generation products but also Beginning-of-Life (BOL) data such as design data, manufacturing data and test data should be exploited for data instance to have contextual meaning.

2 Problem Statement

Nowadays, adoption of new technologies and diversifying Voice of Customer (VOC) have encouraged competition between manufacturing companies. Therefore, to overcome competition, most of the companies reduce cycle time of the product design phase and the entire Product Life Cycle (PLC). Moreover, new generation products have more functionalities and complicated subassemblies, and therefore maintenance of these products is much more complex. For further functionalities and subassemblies, sensor data from the previous generation product can be exploited. However, even if this data has high data availability, it has low data relevance (See Fig. 1). On the contrary, the sensor data from the new generation product has high data relevance but low data availability. Therefore, this study deals with the problem as to how we could exploit the data coming from previous generation products for better maintenance of the new generation of these assets. For this purpose, we focus on generation of virtual data which are available in the initial phase of product instead of real data using Gibbs sampling. Hence, main objectives of this research are to facilitate the information search relevant to the maintenance recommendation and to provide an optimized maintenance recommendation to overcome data insufficiency.

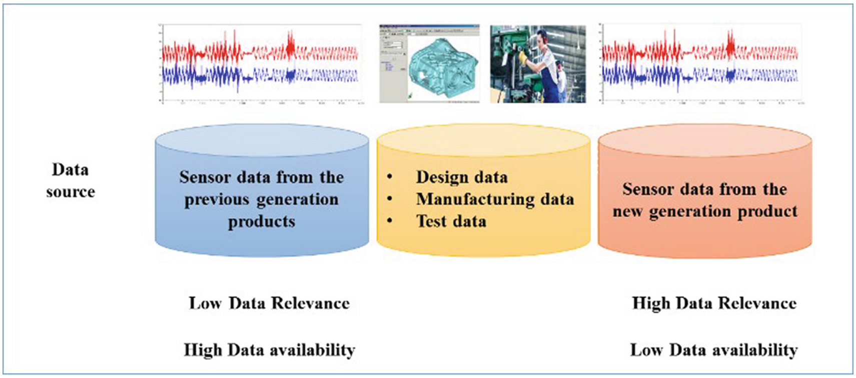


Fig. 1. Data relevance and availability of this study

3 Proposed Approach

The proposed approach is a part of a predictive maintenance framework for new generation products (See Fig. 2). To facilitate the information search, the entire structure applies an ontology based semantic technology to provide contextual meaning for each data instance. In addition, as useful method for prediction or pattern recognition through rules that machine finds itself instead of the explicit rules created by a human, machine learning algorithms is used.

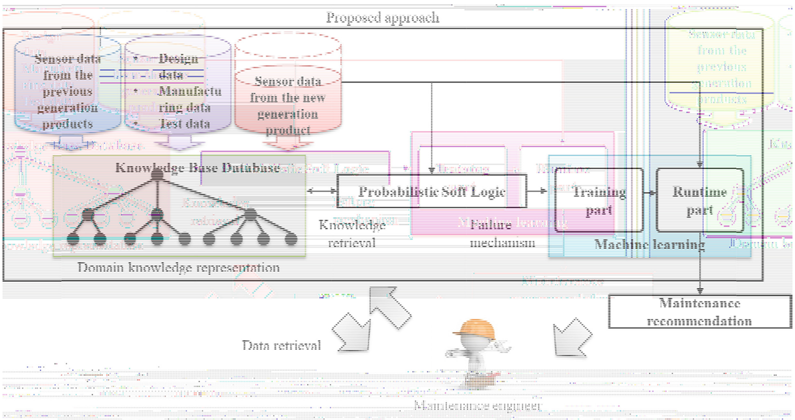


Fig. 2. The entire structure of the predictive maintenance system for new generation products

The individual modules included within the proposed predictive maintenance system is as follows:

- Data sources - i.e. BOL data sources including design data, production data and test data etc.

- Ontology - to merge and wrap data from various data sources for supporting data retrieval
- Probabilistic Soft Logic - as a set of the first order logic to define relevance between further functionalities/components and previous
- Training part - to extract significant meaning from massive data and to automatically learn the model which represents product states and probability
- Runtime part - to analyse a product state and a risk of potential failure using the model created by training part

4 Sampling Method with a Case Study

This chapter describes a sampling method with a case study as illustrated in Fig. 3.

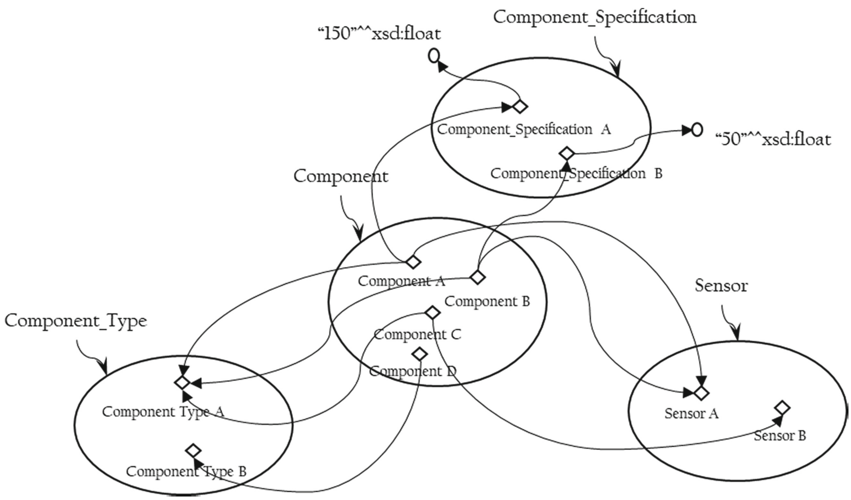


Fig. 3. Ontology for case study

Accordingly, Fig. 4 depicts the procedures of the proposed sampling method, following the entire structure presented in Chapter 3.

First, through the Probabilistic Soft Logic module which can provide a set of rules to figure out component relevance, all the possible data is collected with its relevance. For the case study with target component A, here is an example of a set of first order logics:

- If two components have similar component types, they are probably same (Component_Type(Component A, Component_Type A) & Component_Type (Component B, Component_Type B) & Similar(Component_Type A, Component_Type B) → SameComponent(A,B))

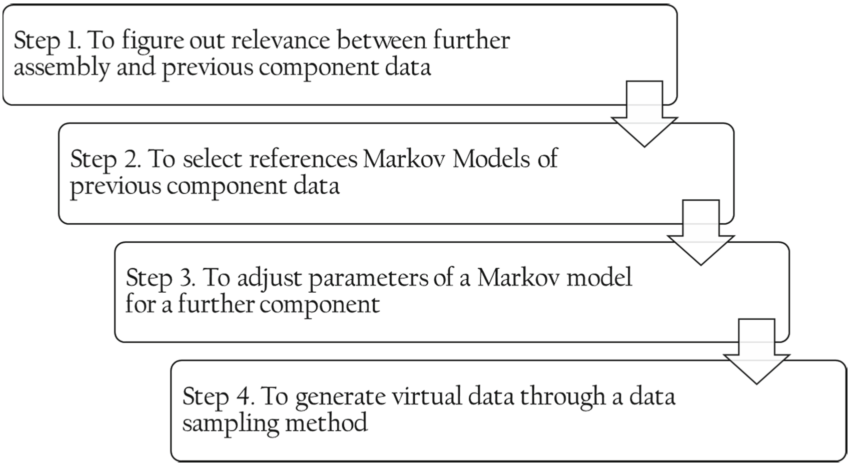


Fig. 4. Procedures of the proposed sampling method

- If two components have similar sensors, they are probably same (Sensor(Component A, Sensor A) & Sensor(Component B, Sensor B) & Similar(Sensor A, Sensor B) → SameComponent(A,B))
- Component relevance is aggregates of above logics (Average[SameComponent (A, B)]).

According to these logics, component relevance could be estimated as follows; Average[SameComponent (A,B)]) = 1, Average[SameComponent (A,C)]) = 0.5, and Average[SameComponent (A,C)]) = 0.

Second, reference data among all the possible data which can be represented in Markov model is selected based on its relevance. In this case study, only in the case that component relevance is 1, component data can be selected, so only Component B data is reference data.

Third, transaction rates of a Markov model (Fig. 5) for a further component are adjusted considering data relevance. For this case study, an assumption is that transaction rates and component specification values are in direct proportion.

Therefore, transaction rates for a further component are adjusted to 3 l, where transaction rates for Component B is l.

Last, virtual data is generated through a data sampling method. Sampling is defined as the selection of a subset of individuals within a population to estimate the characteristics of whole population, and it is widely used for gathering information about a population [3]. Gibbs sampling is a useful method to approximate in the various dimension spaces and it is a Markov Chain Monte Carlo (MCMC) algorithm to obtain a sequence of observations. Gibbs sampling is a special case of Metropolis-Hastings algorithm where Metroplois-Hastings algorithms are a Markov Chain Monte Carlo (MCMC) method, and where from a joint distribution $p(x1,\ldots, xn)$ of n-th random variables $(X1,\ldots, Xn)$, implementation of Gibbs sampling to obtain k samples of X is as follows [4]:

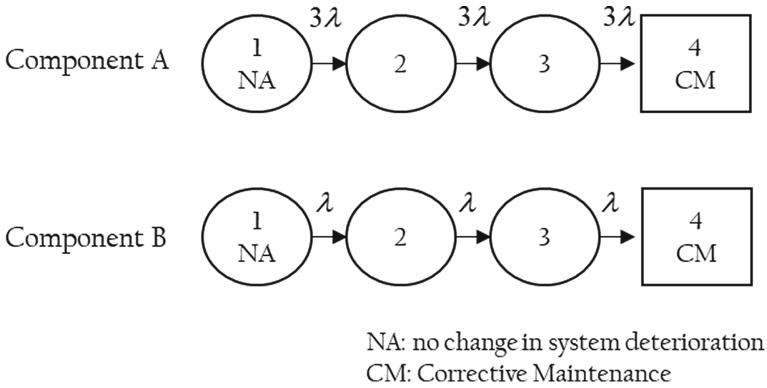


Fig. 5. Markov models for the case study

- Begin with some initial value X(0)
- To get the next sample (call it the (*i* + *1*)-th sample for generality), sample each component variable $x_k^{(i+1)}$ from the distribution of that variable conditioned on all other variables, making use of the most recent values and updating the variable with its new value as soon as it has been sampled. This requires updating each of the component variables in turn, up to the *j*-th component, it is updated according to the distribution specified by: $p(x_j^{(i+1)}|x_1^{(i+1)},\ldots,x_{j-1}^{(i+1)},x_{j+1}^{i},\ldots,x_n^{i})$. Note that the value that the (*j* + *1*)-th component had in the *i*-th sample not the (*i* + *1*)-th is used.
- Repeat the above step *k* times.

This study will apply Gibbs sampling for generating data of the new generation product instead of insufficient data.

5 Conclusion

In this research, we presented a conceptual model providing the main elements of a data sampling approach to apply machine learning techniques addressing the issue of insufficient data of new generation products. The proposed approach would foster predictive maintenance of new generation products in the context of sustainable manufacturing, and will contribute to the knowledge-based prediction modelling using machine learning techniques. Accordingly, the study paves the way for predictive manufacturing mechanism to early-predict equipment condition and make optimized recommendations for adjustments and maintenance to ensure normal operations. Further research on the subject should test and validate the proposed approach, demonstrating its applicability in several use cases of different industries, acting as a facilitator of effective implementation of advanced machine learning techniques for predictive maintenance with the expectation of massive data in the future.

Acknowledgements. This work has been carried out in the framework of Z-BRE4K Project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 768869.

References

1. Cho, S.J., Jun, H.B., Choi, S.D., Shin, J.H.: A study on the prognosis maintenance algorithm for LNG FPSO compressor. In: Proceedings of Scientific Cooperations Workshops on Engineering Branches, pp. 233–238 (2014)
2. Sullivan, G., Pugh, R., Melendez, A.P., Hunt, W.D.: Operations & Maintenance Best Practices-A Guide to Achieving Operational Efficiency (Release 3) (No. PNNL-19634). Pacific Northwest National Laboratory (PNNL), Richland, WA (US) (2010)
3. Singh, A.S., Masuku, M.B.: Fundamental of applied research and sampling techniques. Int. J. Med. Appl. Sci. **2**(4), 123–124 (2013)
4. Karyotis, V., Khouzani, M.H.R.: Malware Diffusion Models for Modern Complex Networks: Theory and Applications. Morgan Kaufmann (2016)

Intelligent Maintenance Maturity of Offshore Oil and Gas Platform: A Customized Assessment Model Complies with Industry 4.0 Vision

Santiago Echeverri Duque(✉) and Idriss El-Thalji

University of Stavanger, Stavanger, Norway
s.echeverriduque@stud.uis.no, idriss.el-thalji@uis.no

Abstract. The world economic forum in 2017 highlighted that Predictive maintenance in O&G is expected to have a total benefit impact in 160 USD billions until 2025 reducing inspection and maintenance cost. This expected impact is based on the digital transformation initiatives in the era of industry 4.0. Several companies have started the progress toward revolutionise their assets and operations to be smarter. Therefore, several assessment models for the readiness and maturity of the digital transformation are existing e.g. IMPULS. These maturity models assess the level of implementing Industry 4.0 philosophy with respect to several dimension i.e. strategy, leadership, customers, products, operations, culture, people, governance. It is worthy to note that these models are oriented to assess the manufacturing companies, as it is the origin of industry 4.0. Thus, using such models to assess the maturity of operation and maintenance systems for O&G platform requires more customised maturity model. The purpose of this paper is to present the developed maturity model, which assess the practices in eight dimensions (Asset, perception, transmission, conversion, computation, cognition, configuration and business) and generate advisory for further development e.g. roadmap, strategic plan. The eight dimensions are categorised into three main categories (Physical space, Cyber space and Business) which allow to assess specific details in the assets, IT and management configuration.

Keywords: Industry 4.0 · Intelligent maintenance · Maturity model, oil & gas platform · Norwegian continental shelf

1 Introduction

The world economic forum in 2017 highlighted that Predictive maintenance in O&G is expected to have a total benefit impact in 160 USD billions until 2025, unlocking value for the industry (World Economic Forum 2017). This expected impact is based on the digital transformation initiatives in the era of industry 4.0. The digital transformation is targeting to reduce inspection and maintenance cost. Several million dollars' worth of inspections and maintenance-related activities each year to ensure that each of these offshore platforms have not been seriously damaged by a variety of potential problems

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 653–663, 2020.
https://doi.org/10.1007/978-3-030-48021-9_73

(Perrons and Richards 2013) requires to be reduced. The key enabling technologies of the digital transformation such as autonomous operations and robots, remote operations centre, predictive maintenance, cognitive computing and wearables will be used to predict performance degradation, autonomously manage and optimize operation.

Internet of things (IoT), cloud computing, big data, Prognostics and Health Management (PHM), and cyber-physical systems (CPS) have contributed to develop the paradigm of preventive maintenance function according to Lee et al. (2014a), evolving from a reactive, traditional, problem solving approach towards a proactive, innovative, performance enhancement approach. Lee et al. (2011) has reported that machines have been given certain intelligence on maintenance to perform "predict and prevent" practice instead of "fail and fix" operation by using instruments such as sensors, meters, controllers and computational devices, and smart algorithms such as Bayesian networks, neural network, fuzzy logic, match matrix, and hidden Markov model. Intelligence or smart maintenance is part of these new revolution developed to obtain transparency of current state of the equipment, failure analysis for diagnosis purposes, and an estimate of the machine's remaining useful life achieving advantages for "worry-free uptime" such as cost reduction, operation efficiency and even product quality improvement according to Lapira et al. (2013).

Several companies have started the progress toward revolutionise their assets and operations to be smarter. Nowadays, Germany is leading a transformation toward 4th Generation Industrial Revolution (Lee et al. 2014b). Reference Architecture Model for Industry 4.0 named as RAMI 4.0 (Adolphs et al. 2015; Bitkom et al. 2016) which was developed in Germany, has helped to understand, standardize and mapping how manufacturing enterprises could implement Industry 4.0. However, according to Trappey et al. (2017) there are many numbers of layers reviewed in the literature (from five to nine layers) to structure the architecture of Industry 4.0. For instance, Lee et al. (2015) has proposed 5Cs Model (connection, conversion, cyber, cognition and configuration) to define intelligent maintenance system which supports Industry 4.0 vision.

Meanwhile, several assessment models for the readiness and maturity of the digital transformation have been defined. Models as Industry 4.0 readiness online self-check for businesses (Impuls 2015), Industry 4.0 digital operations self-assessment (PricewaterhouseCoopers 2016), and Industry 4.0 Maturity Model (Schumacher et al. 2016) have been released. It is worthy to note that these models are oriented to assess the manufacturing companies, as it is the origin of Industry 4.0. Using such models to assess the maturity of operation and maintenance systems for O&G platform requires more customised maturity models which is the goal of this research.

Therefore, the purpose of this paper is to present the development of a customised assessment model to assess the progress in implementing Industry 4.0 vision, particularly the intelligent operation and maintenance system, for oil and gas platforms. In Sect. 2, is structured the maturity model architecture. The dimensions are defined matching the layers of Industry 4.0 and existing maturity models customized to the Offshore O&G technical and business requirements. Based on literature review is defined the capabilities required to implement intelligent maintenance. Lastly, the conclusions of the present research are described in Sect. 3.

2 Development Process of Maintenance 4.0 Maturity Model

Andreas Schumacher et al. (2016) stated that the maturity of an industrial company is appreciated as "the state of growth of internal and external conditions of basic concepts of Industry 4.0 such as the vertical and horizontal integration of manufacturing systems and enterprises as well as the digital integration of engineering across the entire value chain". Therefore, the maturity model shall serve as the scale for the appraisal of the position on the evolution path (Becker et al. 2009). To reach a specific maturity level, the model provides criteria and features that need to be accomplished. Becker et al. (2009) proposed a framework to develop an assessment models which has six steps: (1) problem definition, (2) comparison of existing maturity models, (3) determination of development strategy, (4) iterative maturity model development, (5) conception of transfer and evaluation, (6) implementation of transfer media and evaluation. Therefore, in the following sub sections, a specific problem is first described to clarify what will be assessed using the developed model. Second, the existing industry 4.0 maturity models will be analysed with aim to define the required assessment dimensions, assessing style (questions, levels) and advisory generation style (overall maturity level, recommendations for enhancement). Third, the developed assessment model is described in terms of structure (dimensions, sequence), assessment style (questions and levels), assessment resolution (how detailed and level of assessment e.g. company, platform, section, machines, equipment), and advisory generation style (platform-specific roadmap, company-specific strategic plan). Finally, the way to verify the assessment model together with the concept to transfer (responses and results) and implementation are all described.

2.1 Problem Definition

Manufacturing companies are implementing Industry 4.0 concept in their process achieving great amount of benefits from digitalization, cyber physicals system, IoT, cloud-base manufacturing, etc. O&G industry has been working to translate these advantages in their operations, implementing disperse solutions. Maintenance is a core function in industry 4.0 as it is the supportive function for the production and operations. In fact, it has been under developed in term of automated and digitalised work processes compare to production operations. However, the vision of industry 4.0 will not be achievable without intelligent maintenance system that is smart enough to perceive, learn and care about its asset performance and health. To achieve this goal, a maturity model shall be developed to assess the level of maintenance smartness. It is worth to highlight that the new generation of maintenance (utilising IoT and CPS) was described by use of several terminologies (even though the processes are more or less the same) e.g. eMaintenance, intelligent maintenance, smart maintenance, digital maintenance, deep learning maintenance, maintenance 4.0. There is no standardised definition about intelligent/smart maintenance. However, the term "Maintenance 4.0" might be the best leading term to represent the development progress of maintenance systems toward industry 4.0 vision. Moreover, the standards related to industry 4.0 and IoT are coming soon, where we can foresee that the future maintenance standards e.g. dependability will follow such standards. Therefore, this paper adopt the term "Maintenance 4.0" and aims to develop a Maintenance 4.0 maturity model.

The Maintenance 4.0 maturity model should enable the assessment of maintenance systems e.g. asset and operations. It shall include the assessment of the asset by itself, assessment of the maintenance execution operations (e.g. inspections, prevention, notifications, correction, monitoring) and assessment of the maintenance management (e.g. programmes, planning, and controlling). Assessing the asset smartness is about knowing how smart your platform offshore in providing data about its performance and health. Assessing the maintenance execution operations is about knowing how smart your maintenance operations in providing data about what have been done. Lastly, assessing the maintenance management operations is about knowing how smart your management system in utilising and generating advisory for operating and maintenance execution.

2.2 Comparison of Existing Maturity Models

Based on Scopus® Database were identified the relevant documents for the research. In general, 3377 documents including "Maturity model" keywords and 146 documents including "Maturity model" and "Maintenance" were identified. Moreover, 103 documents were scanned containing "Maturity model" and "Industry 4.0" or "Digitization or Digitalization" or "Cyber-physical system" or "Internet of things" or "Cloud computing" or "cognitive computing" distributed mainly in Computer science, Engineering and Decision Science areas. Adding the keyword "Maintenance", only 3 documents were found. The most popular industry 4.0 maturity models are (1) Industry 4.0 readiness online self-check for businesses (Impuls 2015), (2) Industry 4.0 digital operations self-assessment (PricewaterhouseCoopers 2016), (3) Industry 4.0 Maturity Model (Schumacher et al. 2016), (4) The Connected Enterprise Maturity Model (Rockwell Automation 2014), (5) DREAMY (De Carolis et al. 2017) (6) Maturity Model for Industry 4.0 Strategy (Akdil et al. 2018) and (7) Industry 4.0 MM (Gökalp et al. 2017). The main capabilities assessed based on these models is matched with the main industry 4.0 dimensions i.e. that cover vertical and horizontal integration of process systems and enterprises across the entire value chain, as shown in Table 1.

Table 1. Match of capabilities and dimension of Industry 4.0

Capabilities/Dimensions	Asset	Perception	Transmission	Conversion	Computation	Cognition	Configuration	Business
Infrastructure	•	•	•				•	
Architecture				•	•	•		
Technologies	•	•	•	•	•	•	•	•
Coverage	•	•					•	
Data Management	•	•	•	•				•
Information Management				•	•	•	•	
Integration							•	•
Interoperability			•	•	•	•		
Collaboration								•
Operation	•				•	•	•	•
Implementation	•	•	•	•	•	•	•	•
Decentralization						•	•	•
Real time		•	•	•	•	•	•	•
Leadership								•
People	•	•				•	•	•
Investments	•	•	•	•	•	•	•	•
Standardization		•	•	•	•		•	•
Security				•	•	•		•
Governance								•
Culture								•

The existing industry 4.0 maturity models are useful to provide the main dimensions that any developed model in the future shall consider. Moreover, it provides exceptional way to explore the new technologies and aspects. Therefore, technologies as adaptive robotics, artificial intelligence, simulation, embedded systems, networking, cybersecurity, cloud, virtualization technologies, health sensors, performance sensors, actuators, RFID and RTLS technologies, mobile technologies, augmented reality, 3D printing, real time location system, ERP, and automation have been identified as key elements in Industry 4.0 which will be considered in the developed Maintenance 4.0 maturity model. However, it can be observed that the existing industry 4.0 maturity models have several characteristics: (1) the resolution of the assessment is up to the company level; therefore, business dimension is the most dominant, (2) the results are mainly oriented to determine a fuzzy level of maturity with almost no advisory generation. Thus, the maturity model should lead to advisory generation at both company and industrial sector levels. In addition, the assessment shall provide the summative picture of intelligent maintenance practices at the sector level e.g. O&G which shall lead the government, Petroleum Directorate, industrial associations to effectively collaborate in supporting the progress toward excellence as the case of Germany with manufacturing industry. At the same time, it shall be oriented to provide a detailed picture of intelligent maintenance practices at the factory level e.g. O&G platform which shall lead to identify what should be enhanced rather than just specifying the level of maturity.

2.3 Determination of Development Strategy and the Developed Maintenance 4.0 Maturity Model

The purpose in this research is to match the architecture of Industry 4.0 with existing maturity model dimensions of Industry 4.0 defining the backbone of self-assessment model. On one hand, RAMI 4.0 model (Adolphs et al. 2015) has defined the architecture of Industry 4.0 with six layers (Asset, Integration, Communication, Information, Functions and Business process). Moreover, Lee et al. (2015) has worked with 5C's model (Smart Connection, Data to information-Conversion, Cyber Level-Computation, Cognition and Configuration) for Cyber Physical System. Trappey et al. (2017) has proposed four layers (Perception, Transmission, Computation, Application) for Internet of Things.

However, Trappey et al. (2017) highlighted that there are many numbers of layers reviewed in the literature (from five to nine layers). In fact, the inconsistency among the definition of those layers can be easily noticed once you look into the functionalities in each layer. Therefore, to avoid such confusing definition, we present the intelligent maintenance based on the process flow and related functionalities. Asset-info/data perception-transmission-cyber (conversion, computation, cognition, support decision making, maintenance management, configuration)-asset. Cyber layer (Computation) layer acts as central information hub, from every connected machine and information is being pushed to the hub to form the machines network (Lee et al. 2015) which host several level of computations. Thus, the development strategy is illustrated in the architecture of the assessment for intelligent maintenance as shown in Fig. 1. For Physical space, specific questions have been developed about asset, perception, transmission and configuration to evaluate the capabilities in technologies, strategies and methodologies used for critical equipment in the platform. Conversion, computation, cognition, and configuration is grouped in the Cyber space dimension with general questions that cover any kind of equipment in the platform. Lastly, Business dimension is structured with subcategories as Strategy, People, Culture, Governance and Leadership.

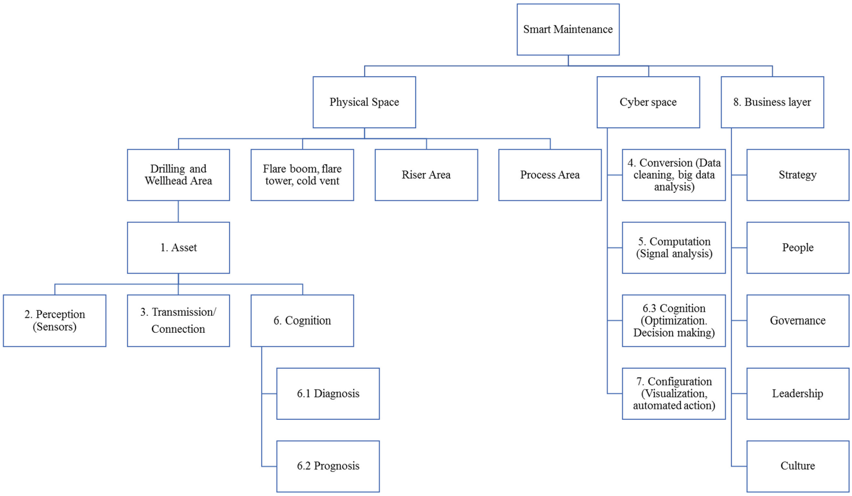
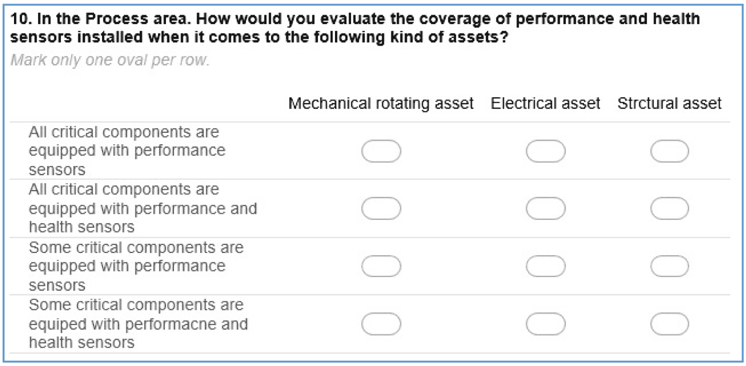


Fig. 1. Architecture of the assessment for intelligent maintenance in offshore O&G platform.

Even though the development strategic was determined to consider the eight dimensions/layers of industry 4.0 and Maintenance 4.0, a deep understanding of each layers was required to create effective assessment. For example, to assess the asset layer, we had to link that to a specific equipment to facilitate the maintenance managers in providing clear answer. The asset e.g. platform has several sections and each section has several production lines that include machineries and instrumentations. In fact, asking in general about if you have implemented intelligent operation and maintenance is fuzzy question. Therefore, we had to breakdown the asset system into several sub/systems and guide the maintenance manager to ask about each subsystem and the critical ones. Different offshore platform as Mobile offshore drilling units (Barge, Submersible, Jack up rig, Semisubmersible and Drilling-ship) and Permanent structures to production (Fixed platforms, Tension leg platform, Semi-submersible, Spars and FPSO-Floating production storage and offloading) can be assessed. In addition, the areas of the platform have been divided according to Norsok (2000): (a) Drilling and Wellhead Area, (b) Flare boom, flare tower, cold vent, (c) Riser Area, (d) Process Area, (e) Emergency Service Areas, (f) Utility Area (g) Living Quarters, (h). Helicopter Deck. After the breakdown structure was ready, we had to explore the existing and potential technologies to make the assets smart, which link the asset layer into the perception layer (how to perceive the performance and health of that asset). Therefore, the maintenance manager can assess the smartness of each level of the whole asset (even up to machine level) concerning the performance and health monitoring technologies. Consequently, we have used multiple choice questions and matrix tables to facilitate the respondents in selecting the most related possibilities and alternatives in terms of technologies, as shown in Fig. 2.


10. In the Process area. How would you evaluate the coverage of performance and health sensors installed when it comes to the following kind of assets?

Mark only one oval per row.

	Mechanical rotating asset	Electrical asset	Strctural asset
All critical components are equipped with performance sensors	◯	◯	◯
All critical components are equipped with performance and health sensors	◯	◯	◯
Some critical components are equipped with performance sensors	◯	◯	◯
Some critical components are equiped with performacne and health sensors	◯	◯	◯

Fig. 2 Example of specific question for intelligent maintenance in offshore O&G platform.

The same issue was related to maintenance operations, we had to breakdown those operations in order to facilitate the maintenance managers to answer the question without confusion. Maintenance operations includes the execution operations (repair, replace, inspect, document) and maintenance management operations (planning, organizing, ordering, performance analysis, scheduling, optimization, diagnosis, prognosis). Then for each operation, the corresponding technologies are listed for maintenance managers to assess which one is used or could be used in the future.

These two issues related to asset and operations were required to reach an effective assessment resolution. However, the developed maintenance 4.0 maturity model provides several maturity levels as follows: level 1, Low level of smartness; Level 2: Moderate level of smartness; Level 3: High level of smartness; and Level 4: Very High level of smartness.

The second challenging issue was related to consider the second requirements, which is related to utilise the assessment in order to generate clear practical advisory as roadmaps or strategic plan. First, the developed model provides a representation and visualization of the maturity level for each dimension using radar charts. It is worth to mentions that business dimension covers subcategories as strategy, people, culture, governance and leadership. Second, the results (maintenance managers' responses) are classified into four categories, two internal (strengths, weaknesses) and two externals (opportunities and threats). In order to enable such smart classification, we had to a hidden layer of analysis where four groups of questions were specified. Moreover, this requirements also encourage us to have several questioning styles, for example, we have asked for how good the used technology is (present state), how good we are in using this technology (internal strength or weaknesses), which one can be potentially used (opportunities), which one might not be used (threads).

2.4 Iterative Development, Media Transfer Concept and Implementation Plan

The maintenance 4.0 maturity model is developed in a web page using Google Form to guarantee the easy accessibility for the self-assessment in Offshore O&G Industry.

Questions are formulated to cover the capabilities identified in the literature review. Enterprises receive a questionnaire via e-mail which include the link of the self-assessment. The survey is addressed to respondents with direct connection of maintenance management. This is our media transfer concept, beside the development of the advisory generation module. The link of Maintenance 4.0 Assessment is available in the reference list developed by Echeverri and El-Thalji (2018).

However, before implementing the assessment model, couple of internal verifications should be done. Following the steps of the maturity model design, it is defined five iterations to establish the final model. In the first iteration, based on a considerable literature review is defined the first model identifying dimensions and describing general questions about Industry 4.0 which are checked inside the research group. During the second iteration, it is elaborated specific questions related to assets of offshore platforms and intelligent maintenance covering the complete value chain. This architecture is assessed in the research group. Semi-structured interviews are developed with experts in O&G Industry and Maintenance to evaluate the maturity model effectiveness in the third iteration. Consequently, the model shall be adjusted according to the feedback received during this session. Consulting maintenance managers in O&G is the last iteration, assessing the architecture in the industry to obtain suggestions. Finally, the maturity model is released for the self-assessment in the Offshore O&G Industry. These verification steps might help us to redefine better assessment scales for each aspect which will reduce the confusion while answering the survey. According to Becker et al. (2009) the web page should be provided for a census of the model acceptance, which may indicate need for further development.

3 Conclusions

Different models have been proposed to define Industry 4.0, cyber physical system, IoT etc. in the last few years, however, none are for intelligent maintenance assessment. The developed Maintenance 4.0 Maturity model is based on several Industry 4.0 maturity models; however, it is more oriented toward maintenance aspects. Three main categories (Business, Physical space, Cyber space) have been defined to assess intelligent maintenance in Offshore O&G platforms, each one contains subcategories to assess the technical and operational aspects. It can be concluded that maturity models at company level are not effective in generating advisory linked to roadmaps and strategic plans. Moreover, it is fuzzy assessment (outsider, beginner, intermediate, experienced, or absence, planned, implementing, learning, maturing). It is hard to measure the progress i.e. in percentage in the lifecycle implementation of maintenance 4.0 and industry 4.0.

There are several challenge that we have handled in this work. The main challenge to develop customised maintenance 4.0 maturity model for specific asset is to breakdown the asset into system-of-systems architecture and consider the difference in assets types' e.g. rotating, structural, electrical, equipment. Maintenance is related to the asset layer and therefore the assessment at the company level does not provide a good resolution of its maturity. The second challenge is to breakdown the maintenance operations (executions and management). The third challenge is to turn the traditional assessment method as being descriptive into being selective. It means instead of asking the responders to write which technology is used/planned to be used, we provide several options for her/him to choose. Such approach it more user/friendly and exploration/oriented as the responder might notice new technologies. However, the selective approach requires good exploratory efforts to list and classify the existing and potential technologies related to each assessed aspect. We had done that for perception layer (sensor technologies for mechanical, electrical and structural assets), cyber technologies (data transmission, data storage, cloud computing, security, etc.), maintenance cognition layer, maintenance decision-making, configuration layer. The fourth challenge is provide measurable scales for the assessment at the question level. For example, asking the managers to assess "how smart their assets are", we had to find a better scale than "low, intermediate, high" as they had no measureable scale in the mind of maintenance managers. Instead, we have asked "all system, all critical systems, some critical system, none" which is more measurable in the mind of maintenance managers.

The fifth challenge and the most challenging one is the advisory generation. In fact, the advisory generation process can represent how smart your maturity/assessment model is. You might analyse the data collected by the maturity model and present the holistic level of maturity without having traceable links to "system of actions e.g. roadmaps, strategic plans". Knowing the required links and make them traceable to the assessment questions is a challenging task and required reverse development (you need to know what the inputs of a roadmap and strategic plan are and then create data acquisition questions to collect them). For example, if the answers of the managers can be classified in smart way, then an effective roadmap can be created. The answers of managers are good in illustrating the strengths, weaknesses, opportunities and threads. Those four aspects are the core of developing an effective roadmap. However, a smart algorithm is needed to get the answers classified into such categorical aspects. The generated advice shall list and prioritise what should be done in the future, which means couple of issues: (1) what are the critical systems and how to expand/scale up the implementation, (2) what are the adopted/procured architectures, algorithms, technologies and how gradually the proofed concept(s) will be implemented (e.g. local cloud into global cloud), (3) what are the required business strategies in terms of human resource development and organization changes and how gradually the changes or resource allocations will be implemented.

The data collected shall lead us to find improvement strategies and Maintenance 4.0 roadmaps to achieve higher maturity levels at both individual and sector level in the Offshore O&G industry and open collaborative platform for Petroleum Directorate, industrial associations and industries to effectively and jointly support the progress toward higher excellence. Meanwhile, this maturity model may be transferred and customized for other industrial sectors e.g. renewables, food processing.

References

Adolphs, P., Bedenbender, H., Dirzus, D., Ehlich, M., Epple, U., Hankel, M., Heidel, R., Hoffmeister, M., Huhle, H., Kärcher, B., Koziolek, H., Pichler, R., Pollmeier, S., Schewe, F., Walter, A., Waser, B. Wollschlaeger, M.: Reference Architecture Model Industrie 4.0 (RAMI4.0) (2015)

Akdil, K.Y., Ustundag, A., Cevikcan, E.: Maturity and readiness model for Industry 4.0 strategy. In: Industry 4.0: Managing the Digital Transformation. Springer International Publishing, Cham

Schumacher, A., Erolb, S., Sihna, W.: A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises. Procedia CIRP **52**, 161–166 (2016)

Becker, J., Knackstedt, R., Pöppelbuß, J.: Developing maturity models for IT Management – a procedure model and its application. Bus. Inf. Syst. Eng. **3**, 213–222 (2009)

BITKOM, VDMA & ZVEI: Implementation Strategy Industrie 4.0. Germany (2016)

De Carolis, A., Macchi, M., Negri, E., Terzi, S.: A maturity model for assessing the digital readiness of manufacturing companies, pp. 13–20. Springer International Publishing, Cham (2017)

Echeverri, S., El-Thalji, I.: Intelligent maintenance maturity of offshore oil and gas platform (2018). https://docs.google.com/forms/d/e/1FAIpQLSfPXBbjtJqbbBtZk_g7enklog6U6iCom-voGiufA2GjgLtgjA/viewform?c=0&w=1. Accessed 25 Apr 2018

Gökalp, E., Şener, U., Eren, P.E.: Development of an assessment model for Industry 4.0: Industry 4.0-MM 2017, pp. 128–142. Springer International Publishing, Cham

IMPULS: Industrie 4.0 Readiness. German Engineering Federation (VDMA) (2015). https://www.industrie40-readiness.de/?lang=en. Accessed 02 Feb 2018

Lapira, E.R., Bagheri, B., Zhao, W., Lee, J., Henriques, R.V., Pereira, C.E., Piccoli, L., Guimarães, C.: A systematic approach to intelligent maintenance of production systems with a framework for embedded implementation. IFAC Proc. Vol. **46**, 23–28 (2013)

Lee, J., Bagheri, B., Kao, H.-A.: A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manuf. Lett. **3**, 18–23 (2015)

Lee, J., Ghaffari, M., Elmeligy, S.: Self-maintenance and engineering immune systems: towards smarter machines and manufacturing systems. Ann. Rev. Control **35**, 111–122 (2011)

Lee, J., Holgado, M., Kao, H.-A., Macchi, M.: New thinking paradigm for maintenance innovation design. IFAC Proc. Vol. **47**, 7104–7109 (2014a)

Lee, J., Kao, H.-A., Yang, S.: Service innovation and smart analytics for Industry 4.0 and Big Data environment. Procedia CIRP **16**, 3–8 (2014b)

NORSOK: Technical Safety S-001. Safety requirements to layout and arrangement, Oslo, Norway (2000)

Perrons, R.K., Richards, M.G.: Applying maintenance strategies from the space and satellite sector to the upstream oil and gas industry: a research agenda. Energy Policy **61**, 60–64 (2013)

Pricewaterhousecoopers: Industry 4.0 Self Assessment (2016). https://i40-self-assessment.pwc.de/i40/landing/. Accessed 02 Feb 2018

Rockwell Automation: The Connected Enterprise Maturity Model (2014)

Trappey, A.J.C., Trappey, C.V., Hareesh Govindarajan, U., Chuang, A.C., Sun, J.J.: A review of essential standards and patent landscapes for the Internet of Things: a key enabler for Industry 4.0. Adv. Eng. Inform. **33**, 208–229 (2017)

World Economic Forum: Digital Transformation Initiative Oil and Gas Industry. Switzerland (2017)

Part VIII: Equipment and Structural Health

Bearing Failure Prediction Technique Using Exponential Moving Average Crossover Threshold with Support Vector Regression and Kernel Regression

Aaron Wee Siong Phan(✉) and Andy Chit Tan

University Tunku Abdul Rahman, Sg Long, Malaysia
aaronphanws@hotmail.com, tanacc@utar.edu.my

An early accurate prediction of remaining useful life (RUL) is essential for improving the machine reliability and prevents system failure. This study proposes an efficient technique to evaluate the health state of bearings and estimate the RUL. It employs the Exponential Moving Average (EMA) crossover technique to actively anticipates an upcoming failure trend of the bearing and the support vector regression (SVR) to constantly predict the RUL of the bearing while its health state is still within the EMA crossover threshold. Once the health state of the bearing exceeds the EMA crossover threshold, Kernel Regression (KR) technique along with SVR will be utilized to instantly predict the failure point and estimate the RUL of the bearing. The effectiveness of the model is validated by experimental data collected from the Center for Intelligent Maintenance Systems (IMS). The proposed prognostic technique shows an effective early failure prediction with great accuracy in comparison to the common model.

1 Introduction

Over the years, maintenance action has evolved from run-to-failure maintenance to preventative maintenance and condition-based maintenance. The current approach is aim towards an intelligent proactive maintenance that is able to forecast failure and predicts the machinery RUL. This allows maintenance to be scheduled optimally to improve machine uptime and saving maintenance cost by staying ahead of any unexpected failure [1, 2]. In general, RUL prediction can be segregated into three categories, namely physics, data-driven and hybrid based model [3, 4].

Physics based prognosis employs mathematical models that describes the physics of the component. The performance of physics based methods greatly depends on the competency of the models to accurately characterize the failure and degradation phenomena [5, 6]. They require an understanding of the specific mechanistic knowledge and theories relating to the systems and therefore can be difficult to build for practical applications [4, 7]. Authors in the past have utilized the Paris law as a degradation model to represent crack growth for RUL estimation of systems subjected to fatigue [8, 9]. Recently, a modified Lundberg-Palmgren formula with modal analysis is used to accurately predict the fatigue life of bearings [10].

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 667–675, 2020.
https://doi.org/10.1007/978-3-030-48021-9_74

Data driven models are based on statistical, pattern recognition and machine-learning techniques and do not rely on the physics theory that governs the system. It is easier to build compared to physics models and has been applied in many practical cases [7]. In [11], the authors utilize the weighted average of both historical and online data-based SVR model aided by least mean square algorithm to adaptively adjust the prediction error. While [12] develops a multivariate state estimation technique and sequential probability ratio test that effectively detects the incipient stages of failure and then followed by the RUL prediction with KR technique. Researchers [13] use Gaussian mixture model to cluster features into groups of health states and then employs distance evaluation technique to predict the RUL of the bearings.

Hybrid models are the integration of both physics and data driven models that takes both theoretical and historical data into consideration while predicting the RUL. One of the recent works on hybrid model is by [14] which combines the least squares support vector regression with the hidden Markov model to attain early bearing faults detection and obtained good RUL prediction. Authors [15] propose another hybrid prognosis method which integrates the advantages of Relevance Vector Machine and exponential model to achieve better prediction performance than other exponential model based methods.

2 Proposed Technique

This paper proposes an effective prognosis technique for bearing RUL prediction. First, the health indicator signals are fed into the EMA crossover algorithm. Upon reaching the threshold of the EMA crossover, bearing failure prediction is carried out with trained SVR algorithm. Once the health indicator has surpassed the EMA threshold, it then indicates an uptrend movement of bearing failure development. From there, KR will be executed along with SVR to output the final RUL of the bearing. The flow chart of the proposed technique is illustrated in Fig. 1.

2.1 Exponential Moving Average Crossover Threshold

Moving average technique has been one of the oldest and most popular tools for trend detection. EMA has been proven to be a robust yet effective forecasting model for time series [16]. One of the reasons that makes the EMA effective as a forecasting model is that it accounts more weightage towards the most recent data into its computation, while consistently decreasing the weightage of earlier data exponentially. This outputs a curve that has more relevant information towards the time series future trend compared to the ordinary moving average process. Hence, making it very useful when it comes to dealing with very aggressive time series [17, 18]. The EMA process of the time series can be defined by Eq. (1):

$$v_t = \frac{\sum_{j=0}^{k-1}(1-a)^{k-j-1}x_{t-k+1+j}}{\sum_{j=0}^{k-1}(1-a)^j} \tag{1}$$

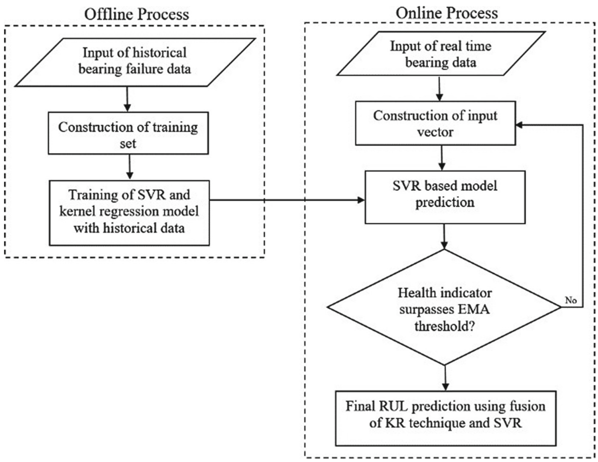


Fig. 1. Flow chart of the proposed technique.

Where, $t = k, k+1, \ldots, n$, and the smoothing factor a is defined as $a = \frac{2}{k+1}$. As k increases, the number of observations in the series of v_t decreases, where it will eventually converge to a single observation when $k = n$. When $k \rightarrow n$ the EMA series is defined by Eq. (2):

$$v_t = \frac{\sum_{j=0}^{n-1} (1-a)^{n-j-1} x_{j+1}}{\sum_{j=0}^{n-1} (1-a)^{j}} \tag{2}$$

Where, $t = k, k+1, \ldots, n$, from Eq. 2 it shows that the EMA process assigns more weightage on the latest observation and decreases exponentially as time decreases. In this paper, the two EMA processes with different number of observation are used simultaneously to plot the EMA crossover curves. The interception of the EMA crossover point will be represented as the final failure threshold of the bearing.

2.2 Support Vector Regression

The Support Vector Machines (SVM), proposed in [19], consists of supervised learning model known for analysing data and recognizing patterns. The adaptation of SVM that handles regression problems is known as Support Vector Regression (SVR). It has a main function of performing linear regression in high dimensional space for function approximation with the aid of various kernel equation. In which it uses nonlinear map defined by dot products function, $\phi : R^a \rightarrow R^a(\beta \geq \alpha)$ to map its data points into high dimensional feature space. Given the training data set is $\{(x_1, y_1), \ldots, (x_p, y_p)\}$, the SVR model applies Eq. (3) to fit the training data.

$$y = f(x) = \omega\,\phi(x) + b \quad (3)$$

where β is the dimensional vectors, $\langle\cdot,\cdot\rangle$ denotes dot product, x and b are the coefficients with target value y, and the (x) is a nonlinear mapping function which maps the input vector x into high dimension space for linear regression so that the non-linear SVR can be achieved. The estimation of these coefficients means solving criteria of Eq. (4):

$$Minimize\ \frac{1}{2}\|\omega\|_2 \quad (4)$$

$$Subject\ to\ y_i - \langle\omega, x_i\rangle - b \le \varepsilon,\ \langle\omega, x_i\rangle + b - y_i \le \varepsilon$$

where ε is a free parameter that functions as a threshold. Note that all predictions have to be within an ε range of the true predictions. Besides that, slack variables can be added into the above to give way for errors and to allow approximation in the case the above problem is infeasible.

2.3 Kernel Regression

KR has been widely used in many research areas of statistics and image processing. Another prominent application of KR is a prediction model. KR is known for its non-parametric property that is highly effective when it comes to estimating regression function *f (x, y)* that best fit non-linear data set *(Xj, Yj)* [20]. Compared to the ordinary linear regression or polynomial regression, KR does not possess any assumptions of underlying distribution to estimate its regression function [21]. The KR based prediction strategy for this paper can be separated into 4 steps. First step is the selection of 3 data points X_j in order to carry out the next-step ahead prediction. These data points are selected based on certain evaluation criteria, which is, i) data points X_j must fall after the EMA threshold, ii) data points X_j must be higher than previous peaks before the EMA threshold, iii) the adjacent data point X_{j+1} must be higher than the ones before it. These criteria are selected to acquire a gradual ascending step points for the training of the KR model. Second step is the application of Gaussian kernel to each of the data point X_j using Eq. (5).

$$K_j(x, X_j) = exp\left(-\frac{(x - X_j)^2}{2\alpha^2}\right) \quad (5)$$

Where x is a data point step vector, X_j is the selected data points and α is the kernel width. The third step is the computation of weight vector $w = (w_1, w_2, \ldots, w_p)$ using least square method to minimize the sum square error between predicted $\hat{Y}_j$ and monitored Y_j as seen on Eq. (6).

$$\varepsilon = \sum\nolimits_{j=1}^{p} \left(\hat{Y}_j - Y_j \right) \tag{6}$$

Last step is to predict the future value Y_j at point X_j by utilizing the KR formula of Nadaraya-Watson kernel weighted average [22], shown in Eq. (7):

$$\hat{Y}_j = \frac{\sum\nolimits_{j=1}^{p} w_j K_j}{\sum\nolimits_{j=1}^{p} K_j} \tag{7}$$

3 Experiment Set up

To validate the proposed prognostic technique, bearing data from platform of NSF I/UCR Center for Intelligent Maintenance Systems (IMS) are used in the analysis. A total of four Rexnord ZA-2115 double row bearings are installed on the shaft of the test rig. The shaft was kept at a constant rotating speed of 2000 r/min by an AC motor and 6000 lb of radial load was applied to the shaft and bearings via a spring loaded mechanism. The sampling rate is 20 kHz and the data length consists of 20480 points. Vibration signals are gathered at intervals of every 10 min via National Instrument DAQ Card 6062E until the test bearing has failed. The setup of the experimental test rig is shown in Fig. 2.

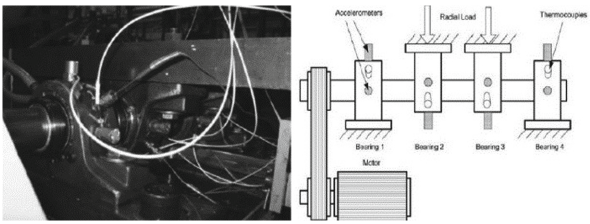


Fig. 2. Illustration of the IMS test rig [23].

4 Results and Discussion

The raw vibration bearing data from the IMS platform are extracted to perform signal analysis. In this paper, 6 features are extracted from the raw bearing signal namely mean, root mean square, kurtosis, skewness, crest factor and peak to peak. Feature reduction were performed using neighbourhood component analysis (NCA) in order to transform the 6 extracted feature into a singular health indicator feature which makes the training and prediction process much faster and efficient. The NCA algorithm function to maximizes the stochastic variant of leave-one-out KNN score on the training set and has been proven to be an effective dimensionality reduction method [24]. Figure 3a shows the normalized health indicator of the test bearing.

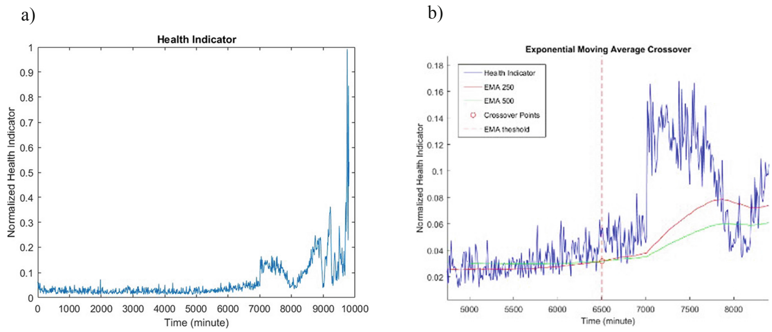


Fig. 3. Diagram for a) normalized health indicator of the test bearing and b) EMA crossover.

After obtaining the bearing health indicator, the EMA crossover technique is implemented to detect the crossover point of the bearing health state. For the time being, the sample sizes of the two EMA curve were manually set based on the best fits of historical data to achieve optimum incipient failure detection. Hence, the sample sizes of 250 and 500 were selected for the EMA curve parameter. When there's a crossover taking place between the two EMA curves as shown in Fig. 3b, it indicates a change in trend movement of the health indicator. For this case, the EMA 250 curve falls above the EMA 500 curve right after the crossover event which signifies the uptrend movement on the health indicator future state. This can also serve as an early warning that the bearing is going to fail soon. The result shows that the crossover point falls at about 6500 min, where the actual failure time of the bearing is 9735 min. This shows that the EMA crossover threshold technique is able to detect the incipient failure of the bearing when it has reached 66.8% of its life span.

Before the incipient failure threshold point is reached, RUL prediction is carried out with the trained SVR model with radial basis function being selected as the kernel, shown in Fig. 4a. When the health indicator has surpassed the EMA threshold, KR prediction will take place to predict the bearing RUL. The KR technique selects 3 health indicator points based on evaluation criteria mentioned earlier to predict the next health indicator, point 4, as shown in Fig. 4b. The KR is a robust predictor and has ability to assimilate non-uniform data points for its prediction of the RUL. The term 'non-uniform' means that the interval between one point to another is not equal, as the health indicator in real world situation often does not change uniformly. The KR will then carry out iterations multi-step ahead until the predicted health indicator has exceeded the failure threshold. Hence, in order to predict point 5, points 2, 3, and 4 are being fed into the algorithm and so forth. As seen on the Fig. 5, it takes up to 7 iterations until the predictor has surpassed the failure threshold at point 10.

The final output from the KR technique predicts the failure point at 8143 min. While the SVR prediction for this technique outputs the final RUL at 9820 min and the actual failure point of the bearing data lies at 9735 min. In this case, the SVR seems to overestimated the final failure point, while the KR underestimated it. Hence to minimize the error of the model, the mean of the predicted results is taken which is at 8982 min. To compare the percentage error in prediction between the proposed

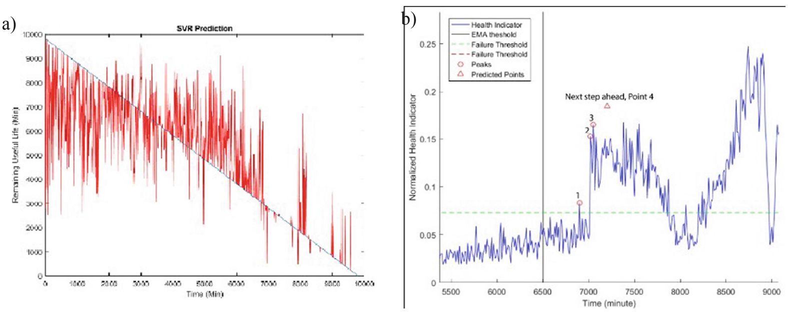


Fig. 4. *RUL Prediction with a) SVR and b) KR next step ahead.*

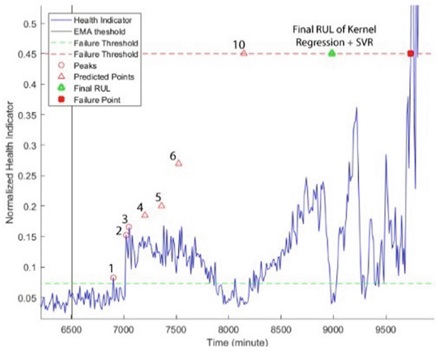


Fig. 5. RUL prediction of the proposed method.

technique and the traditional SVR, the error scoring system used in the IEEE PHM 2012 prognostic challenge is used, Eq. (8):

$$Error\ Score = \begin{cases} exp^{-\ln(0.5).\left(\frac{Er}{5}\right)} if\ Er \leq 0 \\ exp^{-\ln(0.5).\left(\frac{Er}{5}\right)} if\ Er > 0 \end{cases} \quad (8)$$

From the comparison in Table 1, it shows that the proposed technique has a slight better accuracy in RUL prediction as compared to the traditional SVR technique.

Table 1. Comparison of prediction accuracy between the traditional SVR and proposed technique.

	Traditional SVR	Proposed technique
Predicted RUL (Min)	9935	8982
Percentage error (%)	−2.1	7.7
Prognostic challenge error score	1.3	0.4

5 Conclusion

The proposed technique is applicable for various real world problem as i) the model doesn't require large amount of training set to be build, ii) the final RUL prediction does not rely entirely on the historical data being fed into the model, making it robust for machine that operates under various condition and iii) being able to provide an incipient failure warning to the machine operator due to the trend movement detection property of the EMA crossover threshold. Through experimental verification, the results show that the proposed technique can make effective incipient failure detection and a better RUL prediction performance compared to the traditional SVR technique. Future work on this technique can be done to improve the final outcome of the prediction by algorithmically assigning weight to the different predictors output in order to further improve the robustness of the model.

Acknowledgement. The authors would like to acknowledge the financial and laboratory supports of Universiti Tunku Abdul Rahman to accomplish this project. Grant: IPRS/RMC/UTARRF/2015-C1/A03.

References

1. Zhang, P., Li, B., Mi, S., Zhang, Y., Liu, D.: Bearing fault detection using multi-scale fractal dimensions based on morphological covers. Shock Vibr. **19**(6), 1373–1383 (2012)
2. Dong, S., Luo, T.: Bearing degradation process prediction based on the PCA and optimized LS-SVM model. Measurement **46**(9), 3143–3152 (2013)
3. Tian, Z., Jin, T., Wu, B., Ding, F.: Condition based maintenance optimization for wind power generation systems under continuous monitoring. Renew. Energy **36**(5), 1502–1509 (2011). https://doi.org/10.1016/j.renene.2010.10.028
4. Cubillo, A., Perinpanayagam, S., Esperon-Miguez, M.: A review of physics-based models in prognostics: application to gears and bearings of rotating machinery. Adv. Mech. Eng. **8**(8), 168781401666466 (2016). https://doi.org/10.1177/1687814016664660
5. Sikorska, J., Hodkiewicz, M., Ma, L.: Prognostic modelling options for remaining useful life estimation by industry. Mech. Syst. Sig. Process. **25**(5), 1803–1836 (2011). https://doi.org/10.1016/j.ymssp.2010.11.018
6. Tobon-Mejia, D., Medjaher, K., Zerhouni, N., Tripot, G.: A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models. IEEE Trans. Reliab. **61**(2), 491–503 (2012)
7. Heng, A., Zhang, S., Tan, A., Mathew, J.: Rotating machinery prognostics: state of the art, challenges and opportunities. Mech. Syst. Sig. Process. **23**(3), 724–739 (2009)
8. Zio, E., Di Maio, F.: Fatigue crack growth estimation by relevance vector machine. Expert Syst. Appl. **39**(12), 10681–10692 (2012). https://doi.org/10.1016/j.eswa.2012.02.199
9. Coppe, A., Pais, M., Haftka, R., Kim, N.: Using a simple crack growth model in predicting remaining useful life. J. Aircraft **49**(6), 1965–1973 (2012). https://doi.org/10.2514/1.c031808
10. Yakout, M., Elkhatib, A., Nassef, M.: Rolling element bearings absolute life prediction using modal analysis. J. Mech. Sci. Technol. **32**(1), 91–99 (2018). https://doi.org/10.1007/s12206-017-1210-1

11. Wang, X., Gu, H., Xu, L., Hu, C., Guo, H.: A SVR-based remaining life prediction for rolling element bearings. J. Failure Anal. Prev. **15**(4), 548–554 (2015)
12. Caesarendra, W., Tjahjowidodo, T., Kosasih, B., Tieu, A.: Integrated condition monitoring and prognosis method for incipient defect detection and remaining life prediction of low speed slew bearings. Machines **5**(2), 11 (2017)
13. Peng, Y., Cheng, J., Liu, Y., Li, X., Peng, Z.: An adaptive data-driven method for accurate prediction of remaining useful life of rolling bearings. Front. Mech. Eng. (2017). https://doi.org/10.1007/s11465-017-0449-7
14. Liu, Z., Li, Q., Liu, X., Mu, C.: A hybrid LSSVR/HMM-based prognostic approach. sensors **13**(5), 5542–5560 (2013). https://doi.org/10.3390/s130505542
15. Wang, B., Lei, Y., Li, N.: An improved fusion prognostics method for remaining useful life prediction of bearings. In: 2017 IEEE International Conference Prognostics and Health Management (ICPHM) (2017). http://dx.doi.org/10.1109/ICPHM.2017.7998300
16. Zakamulin, V.: Market timing with moving averages: anatomy and performance of trading rules. SSRN Electron. J. (2015)
17. Klinker, F.: Exponential moving average versus moving exponential average. Mathematische Semesterberichte **58**(1), 97–107 (2010)
18. Safi, S.: Comparative study on forecasting accuracy among moving average models with simulation and PALTEL stock market data in Palestine. Am. J. Theoret. Appl. Stat. **2**(6), 202 (2013)
19. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995). https://doi.org/10.1023/a:1022627411411
20. Atkeson, C., Moore, A., Schaal, S.: Locally weighted learning. Artif. Intell. Rev. **11**(1/5), 75–113 (1997). https://doi.org/10.1023/a:1006511328852
21. Mora, C., Schimleck, L.: Kernel regression methods for the prediction of wood properties of Pinus taeda using near infrared spectroscopy. Wood Sci. Technol. **44**(4), 561–578 (2009). https://doi.org/10.1007/s00226-009-0299-5
22. Cai, Z.: Weighted Nadaraya-Watson regression estimation. Stat. Probab. Lett. **51**(3), 307–318 (2001). https://doi.org/10.1016/s0167-7152(00)00172-3
23. Qiu, H., Lee, J., Lin, J., Yu, G.: Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. J. Sound Vibr. **289**(4–5), 1066–1090 (2006). https://doi.org/10.1016/j.jsv.2005.03.007
24. Goldberger, J., Roweis, S., Hinton, G., Salakhutdinov, R.: Neighbourhood components analysis. In: Advances in Neural Information Processing Systems 17, pp. 513–520. MIT Press, Cambridge (2005)

Bearing Fault Diagnosis Based on the Variational Mode Decomposition Technique

J. P. Xing and T. R. Lin(✉)

School of Mechanical and Automotive Engineering,
Qingdao University of Technology, 777 Jialingjiang Rd,
Qingdao Economic and Technology Development Zone,
Qingdao 266520, People's Republic of China
trlin@qut.edu.cn

Abstract. This paper presents a fault diagnosis technique for roller element bearings based on a combined de-trended fluctuation analysis (DFA) and variational mode decomposition (VMD). DFA can reveal the long-range correlation existed within a data series and is used to filter out the uncorrelated trends in a non-stationary bearing condition monitoring (CM) signal and to determine the number of decomposed modes K for VMD. VMD then decomposes the CM signal into K intrinsic mode functions (IMFs) and the fault related IMFs from VMD are selected based on DFA for the re-construction of the filtered signal. The result shows that the de-noised signal after VMD can detect an incipient bearing defect from a low signal-to-noise ratio (SNR) CM data.

1 Introduction

Vibration is the most frequently employed technique for condition monitoring (CM) of roller element bearings which can be obtained by mounting vibration sensors directly on the bearing housing of a machine. The vibration signal generated by a bearing defect in practical situations is most likely to be contaminated by ambient noise generated by other moving parts of a machine during operation. This can lead to a low signal to noise ratio (SNR) CM data, particularly for incipient bearing defects. Moreover, because of speed variation and load change of a machine, bearing CM data acquired from industrial sources often exhibits nonlinear and non-stationary characteristics (Ho and Randall 2000; Estocq et al. 2006).

Empirical mode decomposition (EMD) proposed by Huang in 1998 (Huang et al. 1998) is a powerful tool in dealing with nonlinear and non-stationary signals which has been widely employed for machine fault diagnosis nowadays. For instance, Yu (Yu et al. 2005) utilized EMD and Hilbert transform technique to obtain the local Hilbert marginal spectrum for fault identification of roller element bearings. The fault signal of a low-speed bearing was processed using EMD combining with the kurtosis to filter the trend and noise components in the signal (Xiong et al. 2017). Sharan (Sharan and Zhao 2016) presented a data driven technique which combined EMD with pseudo-fault signal (PFS) to monitor the condition of a rotating machine. Inspired by the success of

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 676–684, 2020.
https://doi.org/10.1007/978-3-030-48021-9_75

EMD method, Smith (2005) proposed another adaptive signal decomposition method termed as Local Mean Decomposition (LMD) to analyse a set of scalp electroencephalogram (EEG) visual perception data. Han (Han and Pan 2015) combined the LMD method with sample entropy and energy ratio for fault diagnosis of roller element bearings. Liu (Liu et al. 2016a, 2016b) proposed a hybrid method based on the second generation of wavelet de-noising (SGWD) and LMD for machine fault diagnosis. However, the inherent recursive sifting operation of the EMD and LMD methods can lead to mode mixing and end effect problems. Variational mode decomposition (VMD) proposed by Dragomiretskiy (Dragomiretskiy and Zosso 2014) is another adaptive signal decomposition method which uses a non-recursive processing strategy to extract signal components in a variational framework. Comparing with EMD and LMD, VMD has a clear mathematical foundation and a higher computational efficiency, though the presence of a strong background noise in a signal can severely affect the setting of mode number which can lead to information loss or over decomposition (Li et al. 2017).

Mert (Mert and Akan 2014) used DFA to define a robust threshold in the development of an EMD based de-noising technique named as EMD-DFA technique. An LMD-DFA technique was employed by Tadivaka (Tadivaka et al. 2017) to mitigate ionospheric scintillation effects of global navigation satellite system (GNSS) receivers, and the performance of the technique was found to be better than that of EMD-DFA. However, the methods cited above still have problems similar to EMD. This prompts the need to develop a more effective method in dealing with low SNR, non-linear, non-stationary signals which motivates the work presented in this study. A combined VMD and DFA technique is proposed in this study to extract useful signal components from a noise contaminated vibration CM signal for an early fault detection of incipient bearing defects. In this approach, the noise contaminated signal is first decomposed by VMD into K IMFs, and then the scaling exponent of each mode is estimated by DFA method. The rest of the paper is arranged as follows: Sect. 2 briefly introduce the VMD and the DFA algorithm. The proposed method of VMD-DFA is elaborated in Sect. 3. Section 4 presents the analysis results of bearing fault signal by VMD-DFA. Conclusions are drawn in Sect. 5.

2 An Introduction of VMD and DFA

2.1 Variational Mode Decomposition

The variational mode decomposition (VMD) utilizes the concept of intrinsic mode function (IMF) in EMD but re-defines the IMFs as AM-FM components:

$$u_k(t) = A_k(t)\cos(\phi_k(t)) \tag{1}$$

where $A_k(t)$ is the amplitude, $\phi_k(t)$ is a non-decreasing function, i.e., $\phi'_k(t) \geq 0$, $A_k(t) \geq 0$.

Each mode $u_k(t)$ has specific sparsity properties, most of them compact around the centre frequency ω_k, and the bandwidth is estimated through the squared L^2-norm of the gradient. A constrained variational mode can then be constructed using:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k(t)} \right\|_2^2 \right\} \text{ subject to } \sum_k u_k(t) = x, \tag{2}$$

where δ is the Dirac delta distribution and * denotes convolution. To solve the constraint variational problem, the quadratic penalty term and the Lagrange multiplier are utilized. The augmented Lagrangian L is given by

$$L(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k(t)} \right\|_2^2 + \left\| x(t) - \sum_k u_k(t) \right\| + \left\langle \lambda(\mathrm{t}), x(\mathrm{t}) - \sum_k u_k(t) \right\rangle \tag{3}$$

where α is the quadratic penalty term, λ(t) is the Lagrange multiplier. Equation (3) can be solved using the alternate direction method of multipliers (ADMM) (Hestenes 1969), and the expression of the modal component u_k and ω_k in the frequency domain can be obtained as:

$$\hat{u}_k^{n+1}(\omega) = \frac{f - \sum_{k \neq K} u_k(\omega) + \frac{\lambda(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \tag{4}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |u_k(\omega)|^2 d\omega}{\int_0^\infty |u_k(\omega)|^2 d\omega} \tag{5}$$

2.2 De-trended Fluctuation Analysis

Assuming a time series $x(n)$, $n = 1, 2, 3 \cdots N$, the steps of the de-trended fluctuation analysis of the time series are as follows:

(1) Calculate the integrated time series $y(k)$ from $x(n)$:

$$y(k) = \sum_{n=1}^{k} [x(n) - \bar{X}], k = 1, 2, 3, \cdots N \tag{6}$$

where $\bar{X}$ denotes the mean value of $x(n)$.

(2) The time series $y(k)$ is equally divided into N_s non-overlapping data segments termed as specifically box, each segment has s sampling points. The trend equation in each box can be expressed as follows:

$$y_s(k) = \sum_{j=0}^{p} \beta_j t^j \tag{7}$$

$\beta_j (j = 0, 1, 2 \cdots, p)$ can be obtained by the least square curve fitting the data of the box to obtain a fitting curve $y_s(k)$.

(3) Eliminate the trend item $y_s(k)$ of each box time series $y(k)$

$$\Delta y_s(k) = y(k) - y_s(k) \tag{8}$$

(4) Calculate the root mean-square of the time series $y(k)$ as a fluctuation function

$$F(s) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (\Delta y_s(k))^2} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \sum [y(k) - y_s(k)]^2} \tag{9}$$

(5) Repeat the above procedure for the segment size s according to a certain step increment. When the signal follows a scaling law, a power-law behaviour for the fluctuation function $F(s)$ can be observed:

$$F(s) \sim s^{\alpha} \Rightarrow F(s) = As^{\alpha} \tag{10}$$

For a time series, $\alpha = 0.5$, $\alpha = 1.0$ and $\alpha = 1.5$ correspond to white noise, pink noise and Brownian noise respectively. When $0 < \alpha < 0.5$, it indicates that the time series has only short-range correlation; when $0.5 < \alpha \leq 1$, it indicates that the time series is not independent of each value with long-range relevance. A persistently enhanced time series whose trend-increasing behaviour depends on the degree of $\alpha > 0.5$.

3 Signal De-noising Using VMD-DFA

The signal is first decomposed into a finite number of IMFs by VMD, the IMF components could be either noise free or contaminated by noise. Therefore, the key question is to establish a reliability metric to determine the relevant IMF components for signal reconstruction. Considering that the IMFs decomposed by VMD are not pure mono-components due to the mode-mixing issue, the threshold metric should contain a buffer interval on top of α. The threshold metric is set at $\theta = \alpha \pm 0.25$ (Mert and Akan 2014). After calculating the scaling exponent α of each IMF, the noisy components having values lower than the threshold θ will be removed. This study uses the method proposed by Liu (Liu et al. 2016a, 2016b) to determine the number of modes K in VMD decomposition:

$$K = \arg\max_{\mathrm{k}}[\alpha_{1:K} \geq \theta = J, K = 1, 2, 3, \ldots],\ J = \begin{cases} 1 & \alpha_0 \leq 0.8 \\ 2 & 0.8 < \alpha_0 \leq 1.0 \\ 3 & 1.0 < \alpha_0 \leq 1.5 \\ 4 & 1.5 \leq \alpha_0 \end{cases} \tag{11}$$

where α_0 denotes the scaling exponent of the input signal.

The processing steps for signal de-noising proposed in this study are summarized below:

Step 1: An observed signal $x(n)$ is considered as a combination of noise $\eta(n)$ and a noise-free signal component $\tilde{x}(n)$: $x(n) = \tilde{x}(n) + \eta(n)$.
Step 2: The signal $x(n)$ is decomposed into K IMFs by VMD based on Eq. (11).
Step 3: The scaling exponent α of each IMF is calculated by DFA.
Step 4: The relevant IMF components are determined based on the predetermined threshold value θ and used for signal reconstruction.
Step 5: An envelope analysis is performed on the reconstructed denoised signal for bearing fault diagnosis

4 Results and Discussion

To verify the effectiveness of the proposed technique for bearing fault detection, a bearing inner race defect signal is simulated using the following equation:

$$Q(t) = \sum\nolimits_{i=1}^{M} B(t)S_b(t - iT_b - \delta T). \tag{12}$$

The signal in Eq. (12) composes of two components, 1) an amplitude modulation component $B(t) = A * [1 - cos(2\pi f_0 t)]/2$ where A = 2 is the amplitude of the load intensity and f_0 = 10 is the shaft rotating frequency, and 2) an impulse response function (IRF) $S_b(t) = e^{-bt} sin(2f_r t)$, where b = 600 is the energy decay constant and f_r = 7 kHz is the bearing resonant frequency. T_b represents the nominal time interval between two adjacent impulses. δT denotes a random slippage effect of the rolling elements, which is a uniform distributed random variable with a zero mean and a standard deviation of 1–2% T_b. The bearing defect frequency in the simulated signal is a ball-pass frequency at inner race (BPFI) at 67.13 Hz. The simulated bearing defect signal (with and without added noise) is shown in Fig. 1.

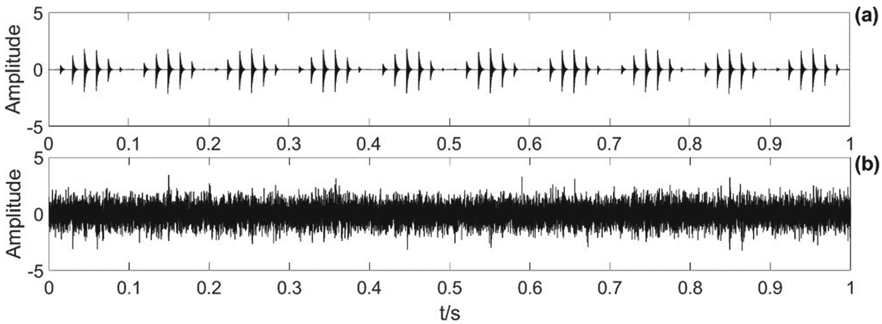


Fig. 1. The simulated bearing defect signal, (a) the noise free defect signal, (b) the noise added defect signal

The envelope spectrum of the simulated noise added defect signal is shown in Fig. 2. It is shown that the simulated defect frequency component is completed submerged by the background noise.

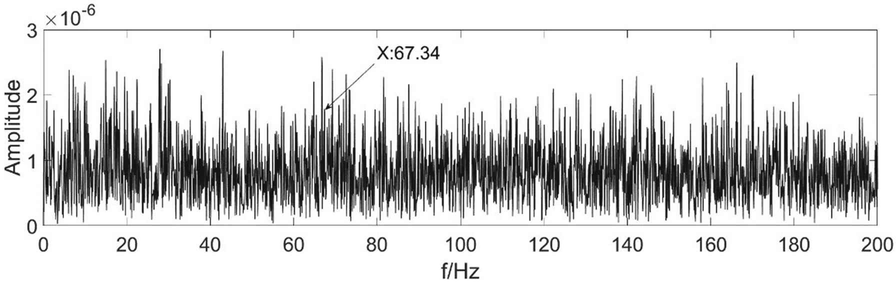


Fig. 2. The envelope spectrum of the simulated noise added defect signal

To demonstrate the effectiveness of the present method in the analysis of the noise contaminated non-linear non-stationary signal, the simulated noise added defect signal is first processed using EMD-DFA method, the reconstructed de-noise signal is shown in Fig. 3 and the envelope spectrum of the de-noise signal is presented in Fig. 4. It is shown that although the process of EMD-DFA de-noise can reduce some noise in the signal, the bearing defect frequency component is still buried under the background noise.

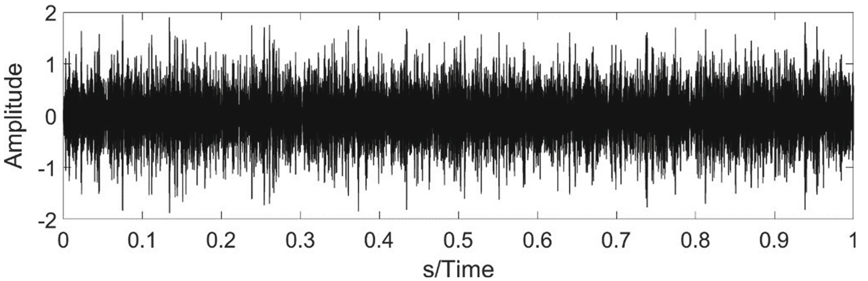


Fig. 3. The reconstructed de-noise signal using EMD-DFA technique

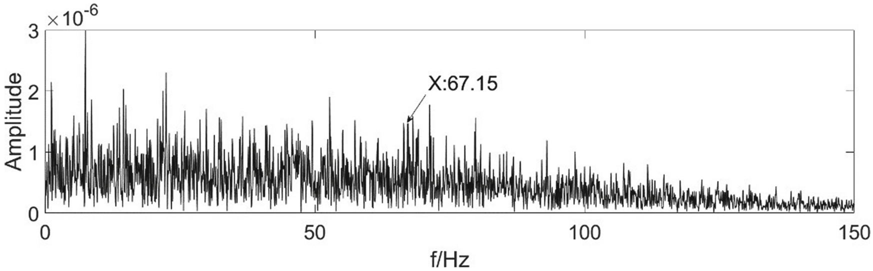


Fig. 4. The envelope spectrum of the de-noise signal using EMD-DFA technique

In the second approach, the simulated noise added defect signal shown in Fig. 1(b) is processed using VMD-DFA technique presented in this paper. According to Eq. (11), the pre-determined number of IMFs in VDM de-noise process is $K = 6$. After performing the de-trended fluctuation analysis (DFA) for each IMF and calculating the scale exponent to be in the threshold region $\theta = 0.5 \pm 0.25$. The IMF components having the scale exponent value lower than this is discarded and the relevant IMFs are used in the signal reconstruction. The reconstructed de-noise signal using the current technique is shown in Fig. 5 and the corresponding envelope spectrum is shown in Fig. 6.

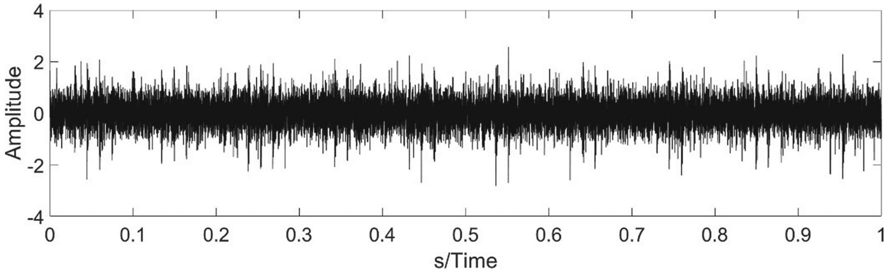


Fig. 5. The reconstructed de-noise signal using VMD-DFA technique

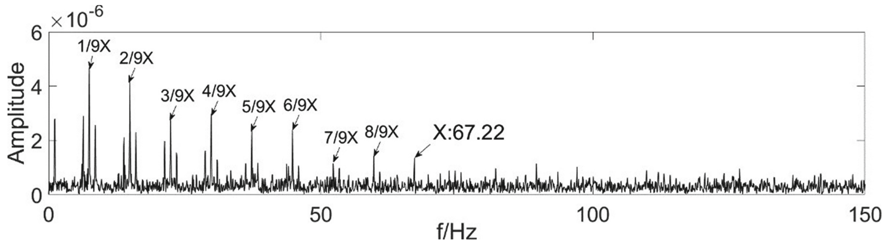


Fig. 6. The envelope spectrum of the de-noise signal using VMD-DFA technique

It is shown in Fig. 6 that the bearing defect signal component and its sub-harmonics can be clearly observed from the envelope spectrum. This then verifies the effectiveness of the current technique in the analysis of the bearing defect signal highly contaminated by noise. Compared with the results shown in Figs. 4 and 6, it is clear that the current method performs better than that of the EMD based de-noise techniques.

5 Conclusions

In this paper, a variational mode decomposition (VMD) combined with the de-trended fluctuation analysis (DFA) were utilized to analyze a noise contaminated bearing fault signal for bearing fault diagnosis. VMD was first employed to decompose the bearing signal into a given number of IMFs. DFA was then employed to calculate the scale exponent of each IMF which is used as a threshold metric to determine whether an IMF is relevant to the defect signal or is noise contaminated. The relevant IMFs were used to reconstruct the de-noise signal for the subsequent signal analysis. It was shown in the paper that comparing with EMD-DFA technique, VMD-DFA can be more efficient to remove the noise in a highly noise contaminated signal.

The proposed technique was successfully deployed to analysis a set of non-stationary bearing signal in this study though its effectiveness in bearing fault diagnosis needs to be tested further on signals acquired from practical industrial environment.

References

Ho, D., Randall, R.B.: Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. Mech. Syst. Sig. Process. **14**(5), 763–788 (2000). https://doi.org/10.1006/mssp.2000.1304

Estocq, P., Bolaers, F., et al.: Method of de-noising by spectral subtraction applied to the detection of rolling bearings defects. J. Vib. Control **12**(2), 197–211 (2006). https://doi.org/10.1177/1077546306041151

Huang, N.E., Shen, Z., Long, S.R., et al.: The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. A **454**, 903–995 (1998). https://doi.org/10.1098/rspa.1998.0193

Yu, D.J., Cheng, J.S., Yang, Y.: Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings. Mech. Syst. Sig. Process. **19**(2), 259–270 (2005). https://doi.org/10.1016/S0888-3270(03)00099-2

Xiong, Q., Xu, Y., Peng, Y., Zhang, W., Li, Y., Tang, L.: Low-speed rolling bearing fault diagnosis based on EMD denoising and parameter estimate with alpha stable distribution. J. Mech. Sci. Technol. **31**(4), 1587–1601 (2017). https://doi.org/10.1007/s12206-017-0306-y

Sharan, S.D., Zhao, Q.: Pseudo-fault signal assisted EMD for fault detection and isolation in rotating machines. Mech. Syst. Sig. Process. **81**, 202–218 (2016). https://doi.org/10.1016/j.ymssp.2016.03.007

Smith, J.S.: The local mean decomposition and its application to EEG perception data. J. R. Soc. Interface **2**(5), 443–454 (2005). https://doi.org/10.1098/rsif.2005.0058

Han, M.H., Pan, J.L.: A fault diagnosis method combined with LMD, sample entropy and energy ratio for roller bearings. Measurement **76**, 7–19 (2015). https://doi.org/10.1016/j.measurement.2015.08.019

Liu, Z.W., He, Z.J., et al.: A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery. ISA Trans. **61**, 211–220 (2016a). https://doi.org/10.1016/j.isatra.2015.12.009

Dragomiretskiy, K., Zosso, D.: Variational mode decomposition. IEEE Trans. Sig. Process. **62**, 531–544 (2014). https://doi.org/10.1109/TSP.2013.2288675

Li, Z.P., Chen, J.L., et al.: Independence-oriented VMD to identify fault feature for wheel set bearing fault diagnosis of high speed locomotive. Mech. Syst. Sig. Process. **85**, 512–529 (2017). https://doi.org/10.1016/j.ymssp.2016.08.042

Mert, A., Akan, A.: Detrended fluctuation thresholding for empirical mode decomposition based denoising. Digital Sig. Process. **32**, 48–56 (2014). https://doi.org/10.1016/j.dsp.2014.06.006

Tadivaka, R.V., Paruchuri, B.P., Miriyala, S., Koppireddi, P.R., Devanaboyina, V.R.: Detection of ionospheric scintillation effects using LMD–DFA. Acta Geophys. **65**(4), 777–784 (2017). https://doi.org/10.1007/s11600-017-0058-1

Hestenes, M.R.: Multiplier and gradient methods. J. Optim. Theory Appl. **4**(5), 303–320 (1969). https://doi.org/10.1007/bf00927673

Liu, Y.Y., Yang, G.L., et al.: Variational mode decomposition denoising combined the detrended fluctuation analysis. Sig. Process. **125**, 349–364 (2016b). https://doi.org/10.1016/j.sigpro.2016.02.011

Adaptive Canonical Variate Analysis for Performance Estimation with Application to Industrial Rotating Machines

Xiaochuan Li(✉) and David Mba

Faculty of Technology, De Montfort University, Leicester, UK
{xiaochuan.li,david.mba}@dmu.ac.uk

Abstract. An adaptive multivariate process modelling approach is developed to improve the accuracy of traditional canonical variate analysis (CVA) in predicting the performance of industrial rotating machines under faulty operating conditions. An adaptive forgetting factor is adopted to update the covariance and cross-covariance matrices of past and future measurements. The forgetting factor is adjusted according to the Euclidean norm of the residual between the predicted model outputs and the actual measurements. The approach was evaluated using condition monitoring data obtained from an operational industrial gas compressor. The results show that the proposed method can be effectively used to predict the performance of industrial rotating machines under faulty operating conditions.

1 Introduction

Multivariate statistical techniques such as principal component analysis (PCA) (Harrou et al. 2013), partial least squares (PLS) (Yacoub and Macgregor 2004) and canonical variate analysis (CVA) (Ruiz-Cárcel et al. 2016) have been widely applied for the detection of abnormalities in large industrial systems. Multivariate subspace identification models based on PCA, PLS or CVA have attracted attention over the past decades because they can be utilized for process monitoring, modelling and system identification. The authors of (Juricek et al. 2005) demonstrated that system-identification models based on CVA outperform modelling models based on regression methods such as PLS. The authors of (Ruiz-Cárcel et al. 2015) demonstrated that monitoring methods based on CVA are more suitable for systems working under changing operating conditions compared to models based on PCA and PLS. The literature provides examples of extensive application of CVA for industrial process modelling and health monitoring. Li et al. (2018) developed a prediction method based on CVA and support vector machine for modelling of industrial reciprocating compressors. The authors of (Larimore et al. 1993) proposed a state-space method using canonical variable states for modelling linear and nonlinear time series. Negiz and Cinar (1997) used a CVA-based subspace-identification approach to describe a high-temperature short-time milk-pasteurization process. CVA was utilized in (Li et al. 2018) to predict performance deterioration and estimate the behaviour of a system under faulty operating conditions. The authors illustrated the performance of the

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 685–696, 2020.
https://doi.org/10.1007/978-3-030-48021-9_76

proposed method in a large-scale 3-phase flow facility. Conventional multivariate subspace-identification approaches based on PCA or PLS are based on the assumption that the process variables are linearly correlated and are independent and identically distributed (IID) (Choi et al. 2006). The requirement that process variables be IID and linear tends to limit the scope of many subspace-identification methods to linear processes operating under steady state conditions. Occasionally, problems associated with the effectiveness of the modelling tools can arise when the underlying assumptions are violated; for instance, the presence of nonlinear distortions, time-dependency, system dynamics and varying operating conditions. Therefore, it is necessary to develop adaptive subspace-identification approaches for systems in which variations in the mode of operation and changes in the system dynamics are common. A number of recursive monitoring methods have been proposed to address these limitations. An extension to the modelling approaches based on the conventional PCA method was proposed by Lane et al. in (2003). The authors illustrated the performance of the proposed recursive PCA model in a polymer film-manufacturing process. Choi et al. (Choi et al. 2006) developed an adaptive multivariate statistical process monitoring (MSPC) for the monitoring of dynamic processes where variations in operating conditions are incurred. The authors of (Lee and Lee 2008) proposed a recursive state-space model based on CVA. In that study, the norm of the difference between consecutive measurements was used to adjust forgetting factors, and the calculation of the optimal values of the minimum and maximum forgetting factors was not detailed.

In this paper, we develop an adaptive monitoring tool based on CVA for the modelling of time-varying processes. We explore the ability of adaptive CVA to predict the behaviour of industrial rotating machines under slowly evolving faulty conditions. To obtain an accurate estimate of system outputs, forgetting factors calculated based on the residual between the model outputs and actual measurements are adopted to update the covariance and cross-covariance matrices of the system. The proposed method is validated on industrial data captured from an operational gas compressor.

2 Methodology

Given system input time-series u_t and output time-series y_t, a linear state-space model can be built as follows (Qin 2006):

$$x_{t+1} = Bx_t + Cu_t + Ke_t \tag{1}$$

$$y_t = Dx_t + Eu_t + e_t \tag{2}$$

where u_t, y_t and x_t are system inputs, system outputs and state vectors; B, C, D, E and K are model coefficient matrices; and e_t is zero-mean and normally distributed independent white noise.

The objective of CVA is to maximize the correlation of two sets of variables (Russell et al. 2000). To generate two data matrices from the measurements, the measurement vector is expanded at each sampling time by including a, the number of previous samples, and b, the number of future samples, to construct the past and future sample vectors $z_{a,t} \in \mathcal{R}^{(n_y+n_u)\cdot a}$ and $u_{b,t} \in \mathcal{R}^{n_u \cdot b}$($n_y$ and n_u are the number of output variables and input variables).

$$z_{a,t} = \begin{bmatrix} y_{t-1} \\ u_{t-1} \\ y_{t-2} \\ u_{t-2} \\ \vdots \\ y_{t-a} \\ u_{t-a} \end{bmatrix} \in \mathcal{R}^{(n_y+n_u)\cdot a} \tag{3}$$

$$u_{b,t} = \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+b-1} \end{bmatrix} \in \mathcal{R}^{n_u \cdot b} \tag{4}$$

The observations can be expanded at each sampling time t by including $a+1$ observations to form the extended past vectors $z_{a+1,t} \in \mathcal{R}^{(n_y+n_u)\cdot(a+1)}$:

$$z_{a+1,t} = \begin{bmatrix} y_{t-1} \\ u_{t-1} \\ y_{t-2} \\ u_{t-2} \\ \vdots \\ y_{t-a-1} \\ u_{t-a-1} \end{bmatrix} \in \mathcal{R}^{(n_y+n_u)\cdot(a+1)} \tag{5}$$

To avoid the domination of variables with large absolute values, the past and future sample vectors are normalized to the zero-mean vectors $\hat{z}_{a,t}$ and $\hat{u}_{b,t}$. Then, the vectors $\hat{z}_{a,t}$ and $\hat{u}_{b,t}$ at different sampling times are rearranged to produce the reshaped matrices $\hat{Z}_a$ and $\widehat{U}_b$:

$$\hat{Z}_a = \left[\hat{z}_{a,t+1}, \hat{z}_{a,t+2}, \ldots, \hat{z}_{a,t+N}\right] \in \mathcal{R}^{(n_y+n_u)\text{a}\times N} \tag{6}$$

$$\widehat{U}_b = \left[\hat{z}_{b,t+1}, \hat{z}_{b,t+2}, \ldots, \hat{z}_{b,t+N}\right] \in \mathcal{R}^{n_u\text{a}\times N} \tag{7}$$

where $N = l - a - b + 1$, and l represents the total number of samples for measurements y_t. Cholesky decomposition is then applied to the past and future matrices $\hat{Z}_a$ and $\widehat{U}_b$ to configure a Hankel matrix $\mathcal{H}$. To find the linear combination that maximizes the correlation of the two sets of variables, the truncated Hankel matrix $\mathcal{H}$ is decomposed using singular value decomposition (SVD):

$$\mathcal{H} = \Sigma_{b,b}^{-1/2}\Sigma_{b,a}\Sigma_{a,a}^{-1/2} = U\Sigma V^T \tag{8}$$

where $\Sigma_{a,a}$ and $\Sigma_{b,b}$ are the sample covariance matrices and $\Sigma_{b,a}$ denotes the cross-covariance matrix of $\hat{Z}_a$ and $\widehat{U}_b$. $\Sigma_{a,a}$, $\Sigma_{b,b}$ and $\Sigma_{b,a}$ are calculated as follows (Odiowei and Yi 2010):

$$\Sigma_{a,a} = \hat{Z}_a\hat{Z}_a^T/(N-1) \tag{9}$$

$$\Sigma_{b,b} = \widehat{U}_b\widehat{U}_b^T/(N-1) \tag{10}$$

$$\Sigma_{b,a} = \widehat{U}_b\hat{Z}_a^T/(N-1) \tag{11}$$

U, V and $\sum$ have the following form:

$$U = [u_1, u_2, \ldots, u_r] \in \mathcal{R}^{n_u\mathrm{a}\times n_u\mathrm{a}}$$

$$V = [v_1, v_2, \ldots, v_r] \in \mathcal{R}^{(n_y+n_u)\mathrm{a}\times(n_y+n_u)\mathrm{a}}$$

$$\Sigma = \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_r \end{bmatrix} \in \mathcal{R}^{n_u a\times(n_y+n_u)\mathrm{a}}$$

The columns of $U = [u_1, u_2, \ldots, u_r]$ and the columns of $V = [v_1, v_2, \ldots, v_r]$ are called the left-singular and right-singular vectors of $\mathcal{H}$. Σ is a diagonal matrix, and its diagonal elements are called singular values and depict the degree of correlation between the corresponding left-singular and right-singular vectors. The right-singular vectors in V corresponding to the largest q singular values are retained in the truncated matrix $V_q = [v_1, v_2, \ldots, v_q] \in \mathcal{R}^{(n_y+n_u)\mathrm{a}\times q}$. This matrix is used later to perform dimension reduction on the measured data.

With the truncated matrix V_q, the $(n_y + n_u)$ dimensional past vector $\hat{Z}_a \in \mathcal{R}^{(n_y+n_u)\mathrm{a}\times N}$ is further converted into a reduced q-dimensional matrix $\Phi \in \mathcal{R}^{q\times N}$ (the columns of Φ are z_t, which are called canonical state variates) by the following:

$$\Phi = [z_{t=1}, z_{t=2}, \ldots, z_{t=N}] = K\cdot\hat{Z}_a = V_q^T\Sigma_{a,a}^{-1/2}\cdot\hat{Z}_a \tag{12}$$

where $K = V_q^T\Sigma_{a,a}^{-1/2} \in \mathcal{R}^{q\times(n_y+n_u)\mathrm{a}}$ is the projection matrix that maps the past observations into the canonical variate space. In this investigation, the number of q is determined in the same way as the traditional CVA model. According to the literature (Odiowei and Yi 2010), if the number of retained states q is no less than the actual order of the system, we can substitute the state variates x_t with the canonical state variates z_t. Therefore, the state variables are defined as a linear combination of the past measurement vector $\hat{Z}_a$ (Lee and Lee 2008):

$$\hat{x}_{t+1} = [K \quad 0] z_{a+1,t} \tag{13}$$

$$\hat{x}_t = [0 \quad K] z_{a+1,t} \tag{14}$$

where $0 \in \mathcal{R}^{q \times (n_y + n_u)}$ with all zero entries. According to the literature (Shang et al. 2015), after the estimates of the state variates are calculated, the matrices B, C, D, E and K are calculated from the measurements through linear least-squares regression as follows:

$$[D \quad E] = Y_{(:,1:N-a-1)} \begin{bmatrix} \hat{X}_{(:,1:N-a-1)} \\ U_{(:,1:N-a-1)} \end{bmatrix}^{+} \tag{15}$$

$$[B \quad C \quad K] = \hat{X}_{(:,2:N-a)} \begin{bmatrix} \hat{X}_{(:,1:N-a-1)} \\ U_{(:,1:N-a-1)} \\ \hat{E}_{(:,1:N-a-1)} \end{bmatrix}^{+} \tag{16}$$

where $\hat{E} = Y - D\hat{X} - EU$.

Due to non-stationary process behaviour, many industrial processes have time-varying characteristics that may cause rapid changes in state variates over time. The sample covariance matrices $\Sigma_{a,a}$ and $\Sigma_{b,b}$ and the cross-covariance matrix $\Sigma_{a,b}$ change according to the change in operating conditions. Constant covariance and cross-covariance matrices may not be able be fully capture the system dynamics. Therefore, the exponential weighted moving-average method is employed in this investigation to update the matrices $\Sigma_{a,a}$, $\Sigma_{b,b}$ and $\Sigma_{b,a}$:

$$\Sigma_{a,a(t)} = (1-\beta) z_{a,t} z_{a,t}^T + \beta \Sigma_{a,a(t-1)} \tag{17}$$

$$\Sigma_{b,b(t)} = (1-\beta) u_{b,t} u_{b,t}^T + \beta \Sigma_{b,b(t-1)} \tag{18}$$

$$\Sigma_{b,a(t)} = (1-\beta) u_{b,t} z_{a,t}^T + \beta \Sigma_{b,a(t-1)} \tag{19}$$

where β is the forgetting factor, which is calculated according to the Euclidean norm of the residual between the predicted model outputs and actual measurements. The initial values of $\Sigma_{a,a(t)}$, $\Sigma_{b,b(t)}$ and $\Sigma_{b,a(t)}$ are determined by the traditional CVA model. Tracking time-varying parameters is an important problem in subspace modelling. A constant forgetting factor is not suitable for tracking time-varying parameters and therefore cannot fully reflect the dynamics of a process under nonstationary conditions (Leung and So 2005). Therefore, the forgetting factor must be changed according to the rate of process change to yield satisfactory predictive results in time-varying environments. In this investigation, the forgetting factor is adjusted based on the Euclidean norm of the residual between the predicted model outputs and the actual measurements as adopted in (Shang et al. 2015). When the forgetting factor is small, it gives more weight to present observations to reduce the impact of past observations on the current model. As the value of the forgetting factor approaches unity, it gives more weight to past measurements, thereby permitting long-term memory of the model. The forgetting factor used in this study is calculated as follows:

$$\beta_t = Ae^{-(\|e_{t-1}\|)/\sigma_1} \tag{20}$$

where A is a constant. The empirical parameter selection procedure proposed by (Choi et al. 2006; Shang et al. 2015) is adopted in this study to determine the value of A. Typically, a value between 0.999 and 0.9 is selected. $\| e_{t-1} \|$ denotes the Euclidean norm of the residual of the actual measurements and model outputs. The tuning parameter σ_1 controls the sensitivity of the model to prediction errors. The larger σ_1 is, the less sensitive the model is to prediction error. The value of the forgetting factor β_t is adjusted at every time instance when new measurements are available.

After the forgetting factor is determined, weighted recursive least squares (WRLS) with adaptive forgetting factor (Leung and So 2005; Turksoy et al. 2014) can be used to update model coefficient matrices B, C, D, E and K. The system described by Eqs. 1 and 2 is rewritten as follows:

$$y_t = \Theta_t^y \emptyset_t + e_t \tag{21}$$

$$\hat{x}_{t+1} = \Theta_t^x \Psi_t \tag{22}$$

where $\Theta_t^y = [\mathrm{D}_t \quad \mathrm{E}_t], \emptyset_t = \left[\hat{x}_t^T \quad \mathrm{u}_t^T\right]^T, \Theta_t^x = [\mathrm{B}_t \quad \mathrm{C}_t \quad \mathrm{K}_t], \Psi_t = \left[\hat{x}_t^T \quad \mathrm{u}_t^T \quad e_t^T\right]^T$. Θ_t^y can be calculated by using the recursive least squares (RLS) (Turksoy et al. 2013):

$$\Theta_t^y = \Theta_{t-1}^y + \left(y_t - \Theta_{t-1}^y \emptyset_t\right) \emptyset_t^T P_t \tag{23}$$

$$P_t = 1/\beta \left(P_{t-1} - \frac{P_{t-1}\emptyset_t \emptyset_t^T P_{t-1}}{\beta + \emptyset_t^T P_{t-1} \emptyset_t} \right) \tag{24}$$

The innovation noise sequence is defined as:

$$e_t = y_t - \Theta_t^y \emptyset_t \tag{25}$$

Similarly, Θ_t^x is calculated as follows:

$$\Theta_t^x = \Theta_{t-1}^x + \left(\hat{x}_{t+1} - \Theta_{t-1}^x \Psi_t\right) \Psi_t^T Q_t \tag{26}$$

$$Q_t = 1/\beta \left(Q_{t-1} - \frac{Q_{t-1}\Psi_t \Psi_t^T Q_{t-1}}{\beta + \Psi_t^T Q_{t-1} \Psi_t} \right) \tag{27}$$

The procedures for subspace identification and performance estimation using the model described above are summarized as follows:

Step 1: Calculate model coefficient matrices using the traditional CVA model.
Step 2: Calculate the forgetting factor β_t as per Eq. 20.
Step 3: Compute the updated covariance and cross-covariance matrices $\Sigma_{a,a(t)}$, $\Sigma_{b,b(t)}$ and $\Sigma_{b,a(t)}$ according to Eqs. 17–19.
Step 4: Update the Hankel matrix $\mathcal{H}$ as per Eq. 8.
Step 5: Estimate the state vectors $\hat{x}_{t+1}$ and $\hat{x}_t$ as per Eqs. 13–14.
Step 6: Update the model coefficient matrices via Eqs. 23–27.
Step 7: Estimate the model outputs y_t according to Eqs. 1–2.
Step 8: Update the forgetting factor β_t based on the residual between the estimated outputs and actual measurements.
Step 9: Repeat step 1 – step 8 iteratively.

3 Case Study

Rotating machines that operate at high speed and under high pressure are subject to performance degradation and failures. If a fault occurs and the fault evolution is slow, the machine operator may choose to keep the machine running until repair facilities and spare parts are available at the plant. In such a case, the proposed adaptive ACVA model can be used to estimate how the system will behave under faulty operating conditions given future system inputs. In this subsection, the proposed method is applied to an operational industrial centrifugal compressor to predict the performance of the machine during bearing degradation.

The measured time series from compressor A consisted of 368 observations and 13 variables. For this study, all data were captured at a sampling rate of one sample per hour. Table 1 summarizes all measured variables for this compressor. As shown in Fig. 1, the compressor is operated under healthy conditions during the first 320 samples. The readings of the four different bearing-temperature sensors start to rise at around the 321th sampling point; the machine continued to run until the 368th sampling point. At that time, site engineers shut down the compressor for inspection and maintenance. To compare the performance of the developed ACVA approach with that of the traditional CVA model, the first 240 sampling points of the monitored time series were utilized to build an offline CVA model. Then the constructed CVA model was fed with the speed set points used throughout the degradation process to estimate how the system was affected by the fault. On the other hand, the developed adaptive CVA approach was employed to update the constructed model iteratively according to steps 2–8 described in Sect. 2. The predicted outputs obtained from the adaptive CVA model were compared with those obtained from the traditional CVA model to evaluate the performance of the proposed adaptive monitoring method.

Table 1. Measured variables of compressor A

ID	Variable name	ID	Variable name	ID	Variable name
1	Speed	6	Radial vibration overall X 1	11	Radial bearing temperature 2
2	Suction pressure	7	Radial vibration overall Y 1	12	Active thrust bearing temperature
3	Discharge pressure	8	Radial bearing temperature 1	13	Inactive thrust bearing temperature
4	Discharge temperature	9	Radial vibration overall X 2		
5	Actual flow	10	Radial vibration overall Y 2		

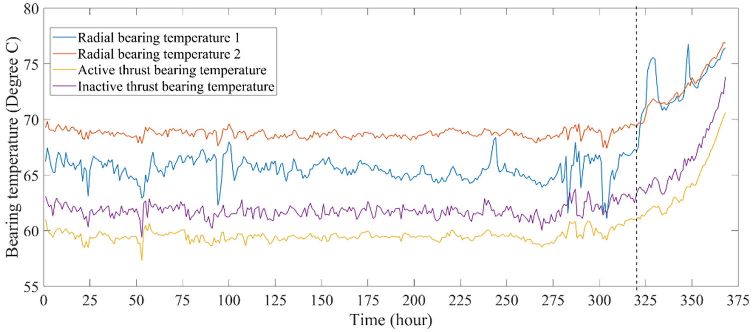


Fig. 1. Trend of four different bearing temperature sensor measurements of compressor A.

In order to determine the optimal number of retained states q, the trained offline CVA model was first utilized to predict system outputs for the data captured during the early stages of degradation. The predicted outputs are compared with the actual measurements and the mean absolute error (MAE) of all output variables are plotted against different values of retained state in Fig. 2. It can be observed from the figure that $q = 1$ gives the lowest prediction error; therefore, q was finally set to 1 to obtain the optimal model that gives the highest predictive accuracy. The value of A was set to 0.97 according to the empirical parameter-selection procedure proposed by (Choi et al., 2006; Shang et al. 2015). The value of the forgetting factor β_t can be updated at each time instance based on the difference between the predicted system outputs and the actual measurements. The value of β_t decreases to achieve faster identification with short memory when the residual of the system outputs is larger. When the residual is small, using more information about the past improves the prediction accuracy of the model. The tuning parameter σ_1 was set to 3.5 in this study, which is the minimum value that can ensure the convergence of the model while maximizing the sensitivity of the model to prediction errors.

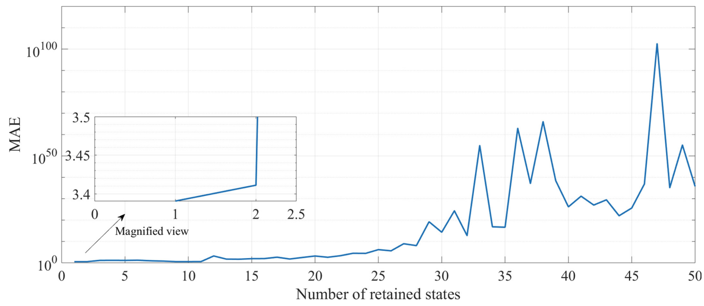


Fig. 2. MAE for all output variables for different values of retained state q

Figures 3, 4, 5 and 6 show the forecasted outputs of adaptive CVA and CVA model. The adaptive CVA model can track changes in bearing temperature measurements more accurately than the traditional CVA method. Table 2 summarises the mean absolute percentage error (MAPE) of the developed ACVA model and conventional CVA model. These results imply that the proposed method takes advantage of recursive state-space modelling to reveal the correlation between system input and output signals, thereby increasing the sensitivity of the adaptive CVA to bearing degradation compared to traditional CVA models.

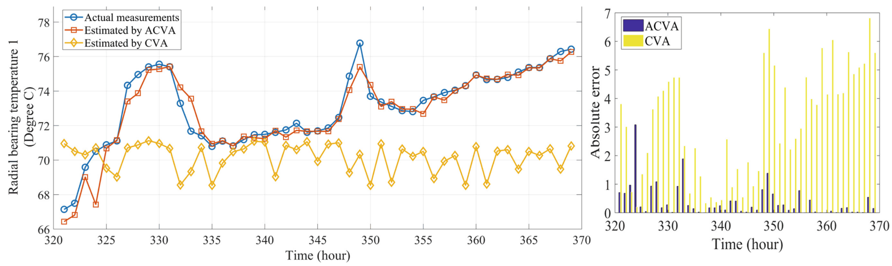


Fig. 3. Radial bearing temperature 1 under faulty operating conditions predicted by adaptive CVA and the CVA model

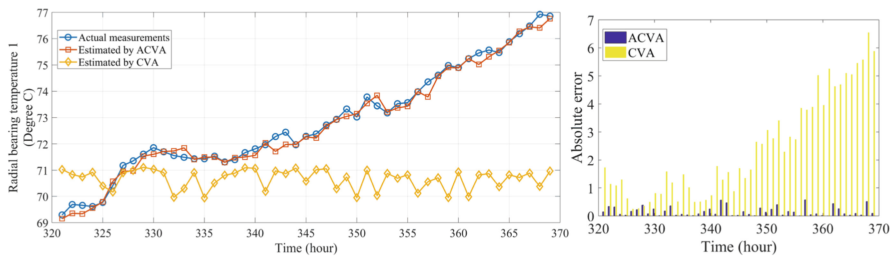


Fig. 4. Radial bearing temperature 2 under faulty operating conditions predicted by adaptive CVA and the CVA model

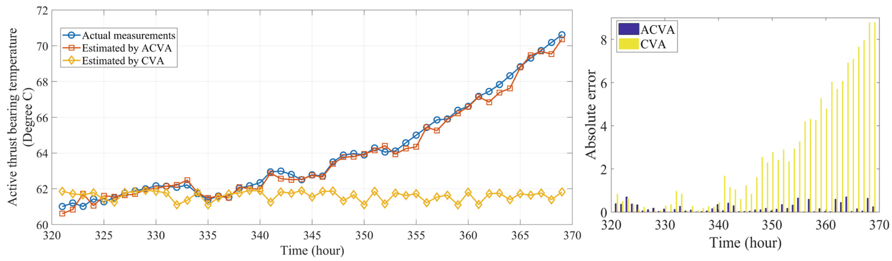


Fig. 5. Active thrust bearing temperature under faulty operating conditions predicted by adaptive CVA and the CVA model

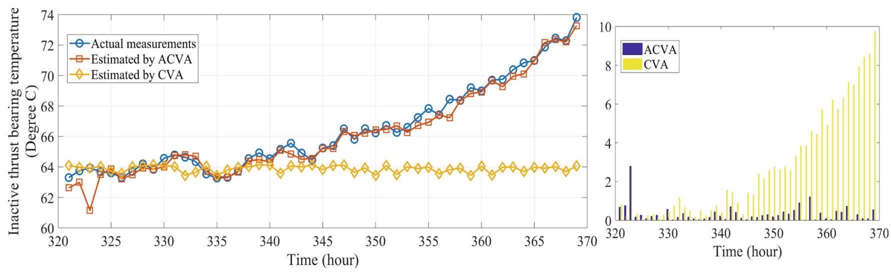


Fig. 6. Inactive thrust bearing temperature under faulty operating conditions predicted by adaptive CVA and the CVA model

Table 2. Mean absolute percentage error (100%) for different output variables

MAPE	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}
CVA	1.5	0.2	4.8	3.7	11.7	18.9	1.2	7.3	3.5	0.2	1.7
ACVA	0.1	0	1	0.1	0.03	0.9	0.7	0.1	0.2	0.1	0.2

4 Conclusion

This paper proposes an adaptive CVA modelling tool to improve the predictive accuracy of traditional CVA methods. A variable forgetting factor was adopted to update the model coefficient matrices and covariance and cross-covariance matrices according to the residuals of the model outputs. The proposed model tracks rapid changes in system outputs due to the use of the adaptive forgetting factor. Condition-monitoring data captured from an operational industrial compressor were used to test the validity of the proposed method. The predicted outputs generated by adaptive CVA highly coincide with the actual measurements. The proposed method takes advantage of recursive state-space modelling to enhance the CVA prognostic performance and increase its sensitivity to bearing deterioration. This method can be used to provide site engineers with more reliable and robust performance estimates of systems operating under varying and abnormal conditions. The information provided by the proposed method can be used to forecast the impact of a fault on the operational process and to develop appropriate production plans and optimal maintenance strategies, thereby making plant operations more safe, productive and profitable.

References

Choi, S.W., Martin, E.B., Morris, A.J., Lee, I.: Adaptive multivariate statistical process control for monitoring time-varying processes. Ind. Eng. Chem. Res. **45**(9), 3108–3118 (2006)

Harrou, F., Nounou, M.N., Nounou, H.N., Madakyaru, M.: Statistical fault detection using PCA-based GLR hypothesis testing. J. Loss Prev. Process Ind. **26** (2013). https://doi.org/10.1016/j.jlp.2012.10.003

Juricek, B., Larimore, W., Juricek, B.C., Seborg, D.E., Larimore, W.E.: Process control applications of subspace and regression-based identification and monitoring methods. In: American Control Conference, pp. 2341–2346 (2005). https://doi.org/10.1109/ACC.2005.1470316

Lane, S., Martin, E.B., Morris, A.J., Gower, P.: Application of exponentially weighted principal component analysis for the monitoring of a polymer lm manufacturing process. Trans. Inst. Measur. Control **1**, 17–35 (2003)

Larimore, W.E., Drive, F., Baillieul, J.: Identification and filtering of nonlinear systems using canonical variate analysis. In: Aerospace Control Systems, vol. 1, pp. 837–841 (1993)

Lee, C., Lee, I.B.: Adaptive monitoring statistics with state space model updating based on canonical variate analysis. Korean J. Chem. Eng. **25**(2), 203–208 (2008). https://doi.org/10.1007/s11814-008-0037-y

Leung, S.H., So, C.F.: Gradient-based variable forgetting factor RLS algorithm in time-varying environments. IEEE Trans. Sig. Process. **53**(8), 3141–3150 (2005). https://doi.org/10.1109/TSP.2005.851110

Li, X., Duan, F., Loukopoulos, P., Bennett, I., Mba, D.: Canonical variable analysis and long short-term memory for fault diagnosis and performance estimation of a centrifugal compressor. Control Eng. Practice **72**(January), 177–191 (2018). https://doi.org/10.1016/j.conengprac.2017.12.006

Li, X., Duan, F., Mba, D., Bennett, I.: Combining canonical variate analysis, probability approach and support vector regression for failure time prediction. J. Intell. Fuzzy Syst. 746–752 (2018). https://doi.org/10.1109/sdpc.2017.146

Negiz, A., Cinar, A.: Statistical monitoring of multivariable dynamic processes with state-space models. AIChE J. **43**(8), 2002–2020 (1997). https://doi.org/10.1002/aic.690430810

Odiowei, P.E.P., Yi, C.: Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. IEEE Trans. Ind. Inform. **6**(1), 36–45 (2010). https://doi.org/10.1109/TII.2009.2032654

Qin, S.J.: An overview of subspace identification. Comput. Chem. Eng. **30**(May), 1502–1513 (2006). https://doi.org/10.1016/j.compchemeng.2006.05.045

Ruiz-Cárcel, C., Jaramillo, V.H., Mba, D., Ottewill, J.R., Cao, Y.: Combination of process and vibration data for improved condition monitoring of industrial systems working under variable operating conditions. Mechanical Systems and Signal Processing, pp. 1–17 (2015)

Ruiz-Cárcel, C., Lao, L., Cao, Y., Mba, D.: Canonical variate analysis for performance degradation under faulty conditions. Control Eng. Practice **54**, 70–80 (2016). https://doi.org/10.1016/j.conengprac.2016.05.018

Russell, E.L., Chiang, L.H., Braatz, R.D.: Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. Chem. Intell. Lab. Syst. **51**(1), 81–93 (2000). https://doi.org/10.1016/S0169-7439(00)00058-7

Shang, L., Liu, J., Turksoy, K., Min Shao, Q., Cinar, A.: Stable recursive canonical variate state space modeling for time-varying processes. Control Eng. Practice **36**, 113–119 (2015). https://doi.org/10.1016/j.conengprac.2014.12.006

Turksoy, K., Bayrak, E.S., Quinn, L., Littlejohn, E., Cinar, A.: Adaptive multivariable closed-loop control of blood glucose concentration in patients with type 1 diabetes. In: 2013 American Control Conference (ACC), pp. 2905–2910. Washington, D.C. (2013)

Turksoy, K., Quinn, L., Littlejohn, E., Cinar, A.: Multivariable adaptive identification and control for artificial pancreas systems. IEEE Trans. Biomed. Eng. **61**(3), 883–891 (2014)

Yacoub, F., Macgregor, J.F.: Product optimization and control in the latent variable space of nonlinear PLS models. Chem. Intell. Lab. Syst. **70**, 63–74 (2004). https://doi.org/10.1016/j.chemolab.2003.10.004

Predictive Analytics Combining Multi-stream Data Sources

R. A. Platfoot(✉)

Covaris Pty Ltd., Sydney, Australia
r.platfoot@covaris.com.au

Abstract. Complex utilities such as power distribution, nuclear reactors or large water networks employ multiple data sources to provide key information regarding asset performance and condition. These include operational logs and fault tracking, SCADA real-time output, work and financial transactions, and a wide variety of formats of inspection and condition monitoring data. To integrate this data requires a complex view of the assets, linking network or functional connectivity with regional locations to build up an understanding of how the extensive fleet of assets must work with each other to achieve the operational objectives.

1 Introduction

Smart operation is a new data-intensive approach where very precise understanding of the entire asset portfolio's condition, capability and current level of resilience will inform a competitive and customer-engaged stance to substantially improve the future positioning of the organisation (Jeble et al. 2016). If this approach is progressed with external and internal stakeholders effectively, it will substantially change not only asset management but the organisation's operational strategy of how its services are delivered to its end customers. The actual hardware of the asset portfolio may not change, but with predictive analytics the portfolio will evolve to a smart network of integrated systems and assets, where the benefits in performance, cost and risk will be quantified (Castellanos et al. 2006).

Predictive analytics involves the use of all data sources in the organisation including the emergent opportunity offered by Smart Meters and the IIoT (Arridha et al. 2017). It includes well-established data sources such as SCADA (Kilpatrick et al. 2008) and improving transaction data in the enterprise asset management system (EAMS, covering financial, work and condition history). The intent is to integrate these sources into a seamless interoperable data lake from which effective forecasts of both asset performance and customer services delivery can be utilised to refine operations as well as long-term asset investment.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 697–703, 2020.
https://doi.org/10.1007/978-3-030-48021-9_77

2 Nature of a Networked Asset Portfolio

A representation of a power network is shown below with a mix of signals or measurements which may be obtained from the diverse asset classes as well as smart metering. This example is a useful case to explain a strategy for asset surveillance (which is a mix of monitoring and proactive testing) which has relevance for both operations support and long-term investment strategy (Nozick et al. 2005).

The web of measurements provides a detailed diagnostic of the component parts of the network and enables decision making to optimise work and investment which will balance cost, risk and performance down to a very detailed level of assets and network connections (Gaynor and Shankaranarayanan 2008). Customer support is obviously from management of the LV network and is affected by both customer behaviour (e.g. the devices they power from the system) and the network delivering power where issues on the HV circuits will impact multiple customers at a time. From an investment perspective, the aggregate of many customer experiences from the LV can inform where significant investment is needed on the expensive but sparser HV networks.

There are six drivers of work on the network split between capital works and operational support. Capital works involves network development to improve the capacity and performance of the network, asset renewal to renew/replace assets in a timely fashion before loss of function, and customer-induced work delivering additional connections to the network. Operational support involves fault response to return the network delivery back to an intended state, routine maintenance which is a mix of preventive maintenance, checks/patrols/inspections and repairs, and operations such as switching, network outage management and SCADA control (Fig. 1).

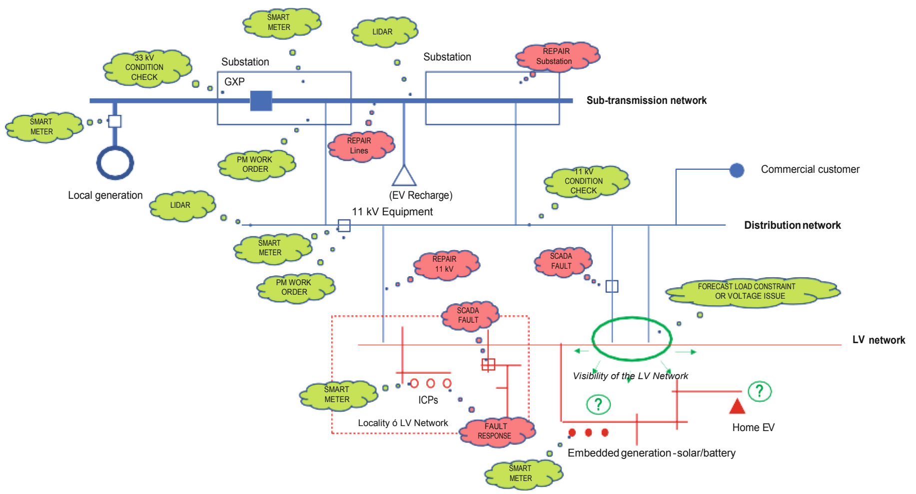


Fig. 1. Surveillance of the power network

All forms of work benefit from improved decision making based on a detailed understanding of the network across all network layers from the sub-transmission HV to the LV.

3 The Data Environment

The data environment such as for the measurements identified in Fig. 2 is shown below.

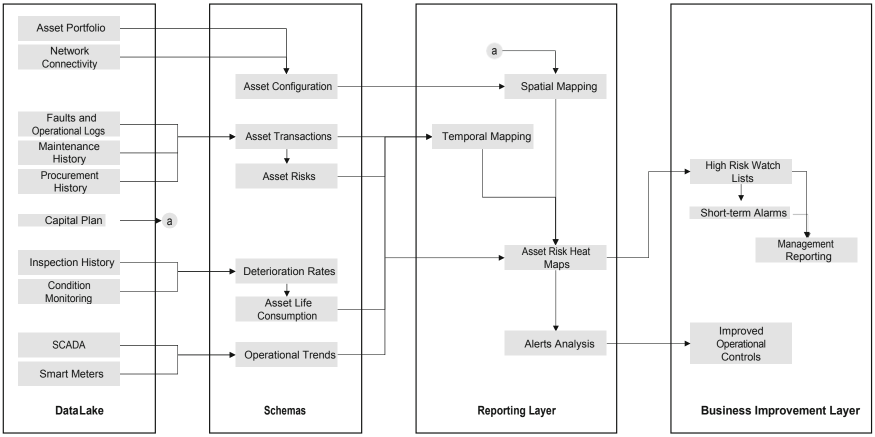


Fig. 2. Data environment – predictive analytics in asset management

There are four layers in this environment, each of which has unique challenges because the data is disparate from each of the data streams entering it from different systems across the organisation (Ferreira and Furtado 2013). The data lake is formed by of a mix of asset identifiers, transaction data and real time data, Saxena et al. (2012). Transaction data consists of discrete transactions such as individual faults, work orders, purchase orders or condition measurements.

The multiple data streams are processed by schema to support reporting of both current and future risks across the assets, which will improve decision making for both current operations and the investment portfolio.

4 Integrating the Data Streams

It is possible to seek out anomalies across the asset portfolio combining this data by correlating the timing of transaction data with changes in the real time data (Ranjan 2008). Snapshots of such changes may be stored as data signatures for future diagnostic work. Analytics by itself cannot determine if the correlation of transactions and the nature of individual signatures picked up in the data represents a problem (Saxena et al. 2012). This requires expert judgement since there is no attribute information to logically guide where to determine a systemic issue. But once such expert judgement has been applied and verified in one case, repeat behaviour can be detected as utilised as a prognostic for future problems (Fig. 3).

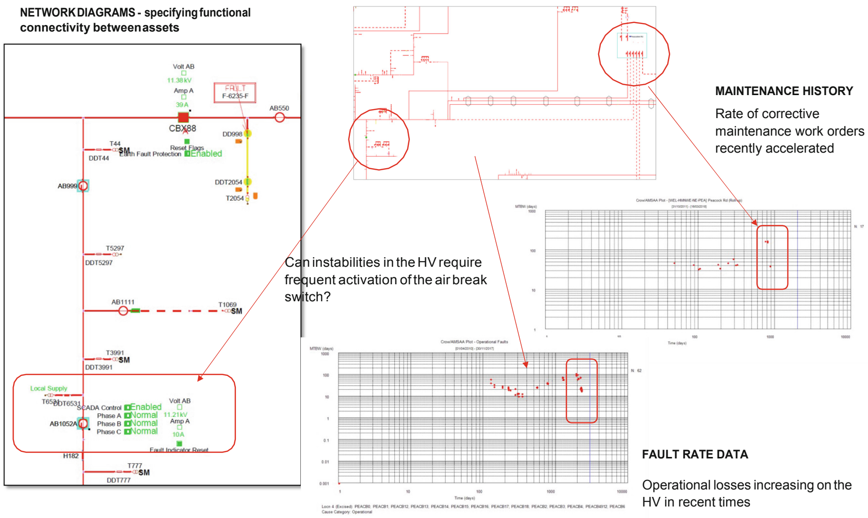


Fig. 3. Electrical connectivity of assets

Faults reported on the network report the realisation of a risk to network operations. Anticipating such faults from a history of maintenance on the assets (e.g. increasing corrective maintenance of circuit breakers) would be advantageous to trigger intervention before the loss in power distribution was realised.

The assets which were tested in the above exhibit were the main circuit breakers on the key circuits coming out of a substation and the downstream Air Break Switch on the network. In this work risk to network health can be associated with spikes in current flows or slow change in voltage levels measured within assets.

The plot below reports current through an air break switch highlighted in the lower left-hand corner of the network diagram above. The horizontal axis is time. The triggering of the switch is an indicator of network health as compared to asset deterioration since the fact that the switch operated does not automatically correlate with significant degradation of the assets. Considering the correlation shown below it could well be coincidence that the period of intensive circuit breaker faults ended around the time of three events on the network: this is a matter of current investigation. What is being tested is how the poor reliability of the circuit breakers is related to network operational concerns and hence current operational risk of the network (Fig. 4).

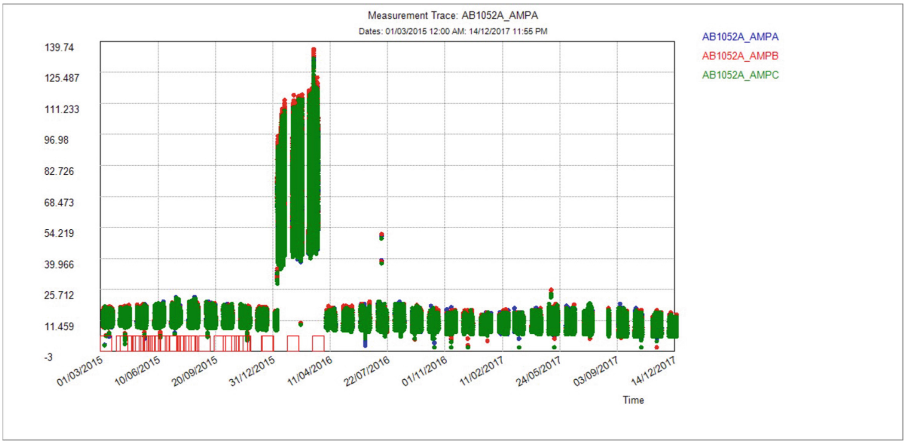


Fig. 4. Correlation of failures

The list on the right-hand side of the figure selected important circuit breakers known from the configuration to be associated with the substation. This list is a subset of the full asset listing since it only contains assets for which faults have been recorded. The real time plot has been updated with marks related to the timing of faults on these circuit breakers. The cessation of faults on specific breakers is obviously of interest and apparently coincides with substantial operational interruption marked by the high amps across all three phases of the air break switch which is being monitored.

This is an example of how multiple streams of data – configuration, faults and real time – need to be brought together to then support expert investigations (Varde and Pecht 2012). It is possible that the coincidence of the data trends is not associated with common causation: but at this stage that remains to be determined and at least can be explored.

5 Reporting

With respect to the reporting layer, the data can be presented over spatial maps (e.g. network maps overlaid on geographic surveys or satellite photographs of the land) as well as time trends in temporal mapping where multiple data streams may be combined in accordance with the time stamps on individual data packets. Both approaches assist experts to understand problems with assets with which they are familiar and for which they need to diagnose events and recommend improvements.

The common metric for interpreting the data is risk which is a function of asset criticality, severity of the issue impacting the asset, likelihood the issue could be realised to its full significance (ranging from catastrophic loss down to operational inefficiencies) and potential for mitigation (e.g. redundancy etc.), (Wlodarczyk and Hacker 2014). In its simplest form, the risk can be inferred as time to realisation of a problem. Raw risk assessments by themselves are often not helpful, and proximity to an alert level for further action is needed to assist with decision making.

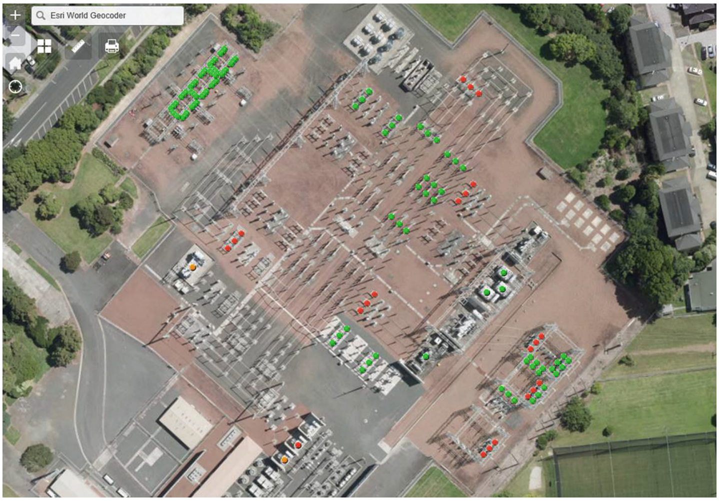


Fig. 5. Potential horizon 2 outcomes

The spatial image shown in Fig. 5 is a substation with specific transformers, circuit breakers, surge arresters or disconnectors exhibiting a level of risk based on past measurements or current performance. This is a dynamic interface which can be seen in a control room or as part of an oversight program. Attribute data associated with the overlay of red and green dots provide details of any problems.

Relating this to the example discussed in Sect. 4: we are seeking to relate network risk as a statistic which is highlighted on spatial maps above because of poor asset reliability. While recognising faults as an obvious sign of loss of network performance, this approach seeks to anticipate high network risk of future faults due to maintenance transactions on related primary assets such as circuit breakers.

6 Conclusion

Predictive analytics is the process by which data is analysed to improve the early warning of risk in a complex asset portfolio and assist with smart operations. The deterioration of assets is based on their consumption during operation and as a function of their working environment and will progress from an indistinguishable level to a state which is detectable by contemporary inspection and condition monitoring techniques. Using the analytics, it is sought to bring forward the time to detect the early stages of deterioration based on feedback from operational control systems.

In this work we seek to integrate multiple data streams which include real time but also less dynamic data such as faults, maintenance and condition measurements, to a stable representation of the health of a complex asset portfolio in terms of performance and condition at a controlled point in time. The reporting is not a real-time update for operations management, but to forecast a level of probability of future events. Hence this work is not an operational control process, but an asset management process.

This work has identified the essential requirements for integrating asset information data sets including real time operational data:

- Improvement of data management, including further refinement of the integration of the data streams, utilisation of work history and condition monitoring data, and utilisation of the Smart Meter data to determine asset risks.
- Refinement of the analysis rules which consider functional connectivity between assets, diagnostic rules to be applied real time data behaviour, and development of schemas for diagnostics which may be inserted into the organisation's data environment.
- Testing of the communication of statistics resulting from the analytics and engagement of stakeholders in case work studying specific issues such as those described in this paper.

References

Arridha, R., Sukaridhoto, S., Pramadihanto, D., Funabiki, N.: Classification extension based on IoT-big data analytic for smart environment monitoring and analytic in real-time system. Int. J. Space Based Situat. Comput. **7**, 82–93 (2017)

Castellanos, M., Salazar, N., Casati, F., Dayal, U., Shan, M.: Predictive business operations management. Int. J. Comput. Sci. Eng. **2**, 292–301 (2006)

Ferreira, N., Furtado, P.: Real-time data warehouse: a solution and evaluation. Int. J. Bus. Intell. Data Min. **8**, 244–263 (2013)

Gaynor, M., Shankaranarayanan, G.: Implications for sensors and sensor-networks for data quality management. Int. J. Inf. Qual. **2**, 75–93 (2008)

Jeble, S., Kumari, S., Patil, Y.: Role of big data and predictive analytics. Int. J. Autom. Logist. **2**, 307–331 (2016)

Kilpatrick, T., Gonzalez, J., Chandia, R., Papa, M., Shenoi, S.: Forensic analysis of SCADA systems and networks. Int. J. Secur. Netw. **3**, 95–102 (2008)

Nozick, L., Turnquist, M., Jones, D., Davis, J., Lawton, C.: Assessing the performance of interdependent infrastructures and optimising investments. Int J. Crit. Infrastruct. **1**, 144–154 (2005)

Ranjan, J.: Traditional business intelligence vis-à-vis real-time business intelligence. Int. J. Inf. Commun. Technol. **1**, 298–317 (2008)

Saxena, A., Roychoudry, I., Celaya, J., Saha, B., Saha, S.S., Goebel, K.: Requirements flowdown for prognostics and health management. In: Infotech@Aerospace 2012 Garden Grove, California, USA, AIAA 2012-2054 (2012)

Varde, P., Pecht, M.: Role of prognostics in support of integrated risk-based engineering in nuclear power plant safety. Int. J. Progn. Health Manag. 1–21 (2012). ISSN 2153-2648

Wlodarczyk, U., Hacker, T.: Current trends in predictive analytics of big data. Int. J. Big Data Intell. **1**, 172–180 (2014)

A Probabilistic Stress - Life Model for Fretting Fatigue of Aluminum Conductor Steel Reinforced Cable - Clamp Systems

Oluwafemi O. Thomas[1(✉)], Luc E. Chouinard[1], and Sebastien Langlois[2]

[1] Department of Civil Engineering and Applied Mechanics, McGill University, Macdonald Engineering Building, 817 Sherbrooke Street West, Montreal, QC H3A oC3, Canada
oluwafemi.thomas@mail.mcgill.ca, luc.chouinard@mcgill.ca
[2] Département de Génie Civil, Université de Sherbrooke, 2500 Boul de L'Université, Sherbrooke, QC J1K 2R1, Canada
sebastien.langlois@USherbrooke.ca

Abstract. This paper presents a probabilistic analysis of a compilation of test data on the fatigue endurance of Aluminum Conductor Steel Reinforced (ACSR) cable–clamp systems. A brief review of the testing and measurement methods used to perform fatigue tests on conductors are described. Theoretical arguments based on the properties of extreme value distributions and random vibrations are presented which indicate that a Weibull S – N model is the most appropriate among models previously proposed in the literature for fatigue of ACSR cable – clamp systems. Predictions from the model are presented in terms of idealized stresses using bending amplitudes. Statistical tests are performed to verify that the Weibull distribution provides a good fit to the conductor fretting fatigue data. Validation datasets independent of the training dataset are used to evaluate the predictive ability of the model. The proposed probabilistic model is shown to be a reliable means of predicting the residual life of conductors subjected to aeolian vibrations for transmission line management and conductor replacement planning.

1 Introduction

Conductors of transmission lines are subjected to a variety of cross – flow induced vibrations such as aeolian vibration, sub-conductor oscillation and galloping (Cloutier et al. 2006). Of these, the phenomenon of fatigue due to aeolian vibrations is discussed in this paper. The fatigue of materials is often treated from two approaches – *a total life approach* which includes the *stress – life (S – N) approach* and *strain – life approach; a damage tolerant approach* based on fracture mechanics (Suresh 1998). The S – N approach is often used in representing fatigue test data on conductors (Cloutier et al. 2006), where the fatigue data are generated using a test bench. The fatigue stress levels are quantified using Poffenberger and Swart (P – S) relationship (Poffenberger and

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 704–715, 2020.
https://doi.org/10.1007/978-3-030-48021-9_78

Swart 1965) or the maximum antinode amplitude (Cloutier et al. 2006) while the number of cycles is quantified as the first wire failure in Cloutier et al. (2006) or 10% of the total number of aluminum wires by CIGRE (1979). The data set and criteria set in Cloutier et al. (2006) are used herein.

There have been previous statistical analysis on conductor fatigue lifetime data (Hardy and Leblond 2001; Thi-lien 2015). These models present the S – N regression curve in terms of the conditional cumulative distribution function for the lifetime given a stress level $F\left(N_1|\sigma_{yb} = \sigma_{yb,i}\right)$, where N_1 is lifetime of the first wire of the conductor and $\sigma_{yb,i}$ is the ith stress level at 89 mm from the last point of contact between the conductor and the clamp. Confidence intervals about the 50$^{\text{th}}$ percentile curve are then produced based on an assumed lognormal distribution of the lifetime. However, $F\left(N_1|\sigma_{yb} = \sigma_{yb,i}\right)$ cannot be constructed from data because the amount of conductor fatigue data at each stress level is too small and non-constant variance of each $F\left(N_1|\sigma_{yb} = \sigma_{yb,i}\right)$. The solution usually adopted is to ignore the variation of variance with stress level or treat it as a function of the stress level (Hardy and Leblond 2001; Thi-lien 2015). Because of these, the fit of the model to data cannot be checked using simple statistical means such as probability plots or the theoretical probability density function (PDF) checked against that of the data. Rather the validity of the model is checked by minimizing the empirical risk, use of the coefficient of correlation (Hardy and Leblond 2001) or plot of the residuals. Strictly speaking minimizing the empirical risk (EMR) doesn't guarantee good prediction ability of the model (Vapnik 1998). The EMR of the model presented herein will still be compared with the one of the model previously proposed by Hardy and Leblond (2001) for the same class of conductor.

The previous works (CIGRE 1979; Hardy and Leblond 2001; Thi-lien 2015) have not assessed also the tail behaviour of their fatigue distributions. Hence there is no information on how the run – out data influence the model prediction and generalization ability. The model presented herein allows studying how the run – outs can influence the tail distribution. The region of validity of these previous models are also not known. Is it valid in the region of the mean? Is it valid at the extremes? When making predictions, an engineer would like to know the region of validity and limitations of his model.

The general objective of this paper is to present an S – N model that allows to evaluate the conductor lifetime, evaluate the accuracy of such prediction and show the effect of ignoring run – out fatigue data on conductor S – N regression models. Due to the limited amount of data in conductor fatigue testing, the model presented herein uses a normalization variable that allows data pooling from various stress levels allowing a larger amount of data to be used in generating the CDF and PDF and checking the theoretical distribution against the empirical probability density function (EPDF) and empirical cumulative distribution function (ECDF).

In the sections to follow, considerations and assumptions used to select the required distribution function are presented. This is followed by a presentation of the methodology used to determine the model parameters. Results of goodness of fit test are presented; finally, validation dataset excluded from the model training are used to check the prediction ability of the model.

2 Determination of Distribution Functions

To determine the fatigue characteristics of stranded conductors, a fatigue test is usually carried out on a test bench. The fatigue stresses are characterised by two stress indicators: P – S relationship or the maximum antinode. These relationships are presented in Cloutier et al. (2006). The stress indicators are obtained as a function of the bending amplitude at 89 mm from the last point of contact of the conductor with the clamp or as a function of the maximum antinode amplitude in the test span. A plot of these stresses against the natural logarithm of number of cycles to failure gives S – N plots which are different for both stress indicators.

The lognormal distribution is usually assumed for the CDF $F\left(N_1|\sigma_{yb}=\sigma_{yb,i}\right)$ in conductor fatigue data analysis (CIGRE 1979; Hardy and Leblond 2001; Thi-lien 2015). However, it is posited that $F\left(N_1|\sigma_{yb}=\sigma_{yb,i}\right)$ can also take an extreme value distribution. Consider a conductor with n number of wires subject to stress at $\sigma_{yb,i}$:

$$N_1|\sigma_{yb,i}=Min\left(N_1,N_2,\ldots\ldots.,N_n|\sigma_{yb}=\sigma_{yb,i}\right) \tag{1}$$

It is seen that the if the lifetime of the first wire of the conductor is used as a criterion to stop the conductor fatigue test, then the conductor lifetime can be considered from the point of view of extreme value statistics of the minimum. The condition of independence between the N_i is not necessary for admissibility of an extreme value distribution to $F\left(N_1|\sigma_{yb}=\sigma_{yb,i}\right)$ (Castillo 1988). The first wire to fail is considered as the weak wire in the whole conductor, thus the Weibull distribution for the minimum is selected to model $F\left(N_1|\sigma_{yb}=\sigma_{yb,i}\right)$.

The distribution of stress level given a number of cycle to first wire failure $F\left(\sigma_{yb}|N_1=N_{1,i}\right)$ is unknown. The normal, lognormal and Weibull distributions have been adopted in the literature to model stress distributions in fatigue (Hanaki et al. 2010). A Weibull distribution is selected to model this distribution. This selection is influenced by some statistical consideration as follows: given any other distribution of $F\left(\sigma_{yb}|N_1=N_{1,i}\right)$, such distributions could be transformed to a Weibull distribution (Castillo and Galambos 1987). A physical consideration that also influences the selection over the normal or lognormal is that under field conditions, the aeolian peak stresses have been defined to follow Rayleigh distribution et (Noiseux et al. 1986) which is a special case of the Weibull distribution (Marshall and Olkin 2007).

It is accepted that a marginal and conditional density are required to construct $F\left(\sigma_{yb,},N_1\right)$, the joint distribution of the stress level and first wire lifetime. However, it has also been accepted that joint distributions can be specified by their conditionals only (Arnold et al. 1999; Besag 1974; Bhattacharyya 1943; Castillo and Galambos

1987). In this light, it is possible to specify the form of $F(\sigma_{yb,}, N_1)$ given that both conditionals have been identified. A bivariate function when both conditionals are Weibull distributions is given in (Castillo and Fernandez-Canteli 2009; Castillo and Galambos 1987):

$$F(N_1, \sigma_a) = 1 - \exp\left[-\left(\frac{(h(N_1) - \zeta)\left(g(\sigma_{yb}) - \xi\right) - \lambda}{\delta}\right)^{\beta}\right];\ N_1, \sigma_a \in \mathbb{R}_{++} \quad (2)$$

$$where\ V = (h(N_1) - \zeta)(g(\sigma_{yb}) - \xi);\ h\,and\,g\,are\ \log functions$$

$$\zeta, \xi, \lambda, \beta\,and\,\delta\,are\,model\,parameters$$

3 Fitting the Model to Data

To fit the model to data, the parameters ζ and ξ are first estimated using a least square approach. Once the values of the parameters ζ and ξ have been determined, the location parameter λ, scale δ and shape parameter β of $F(N_1, \sigma_{yb})$ can be determined by standard methods of parameter estimation; see e.g. (Castillo and Fernandez-Canteli 2009; Kotz and Nadarajah 2000). The maximum likelihood (ML) method has been used to estimate λ, δ, β respectively as (Cousineau 2009; Crowder et al. 1991; Kotz and Nadarajah 2000):

$$maximize\left\{\prod_{i\in U} j(V_i; \beta, \delta, \lambda)\right\}\left\{\prod_{i\in C} S(V_i; \beta, \delta, \lambda)\right\} \quad (3)$$

Subject to the constraint that the Jacobian matrix is stationary. Where j is the Weibull density function and S the survival function, U is a set of failure data and C is a set of run-out data. The analysis done herein has not considered the effect of run-out on the model parameters hence the survival function in (3) was not utilized in obtaining the likelihood function. All data were treated as failure data. Because run-outs are not usually considered in the statistical treatment of conductor fatigue data, see e.g. Hardy and Leblond (2001), this approach has been followed to show the effect of not considering run – out on the S – N curve.

The stress – lifetime pair dataset available in Cloutier et al. (2006) were collected for single, double and 3-layer ACSR class of conductors in terms of bending amplitude. The fitted stress – first wire lifetime models with bending amplitude stress indicator are shown in Fig. 1. The run-out points are indicated in red on these plots.

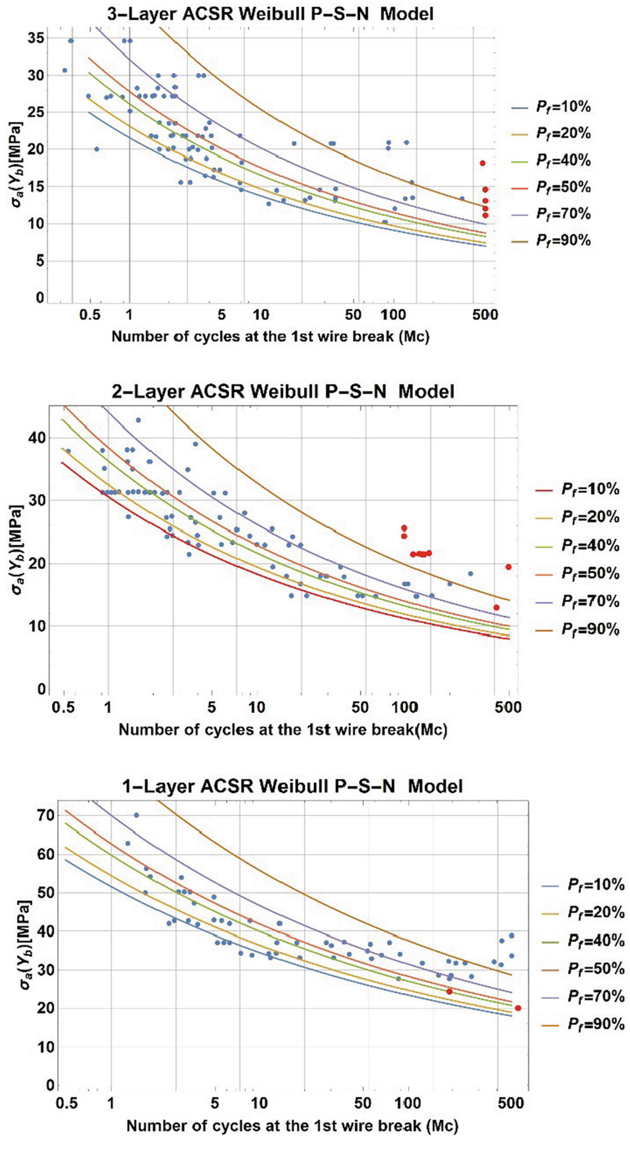


Fig. 1. S - N Curves in terms of Bending Amplitude

4 Evaluating the Goodness of Fit of the Model

There are various ways to test if a selected distribution function model is a good fit for the underlying distribution that generates the data. The simplest being the comparison of the EPDF and ECDF to that of the selected theoretical distribution. If there is a resemblance between the EPDF and the theoretical PDF, it is acceptable to conclude that the selected distribution is a probable function that generated the sample data. If the ECDF also converges well to the theoretical CDF, then it is expected that the frequency of fatigue failure events converges to their probability of occurrence.

In Fig. 2a, the theoretical CDF of the parameter V is compared against its ECDF for the single layer, two layer and three layer ACSR group of conductors. Figure 2b also presents the theoretical PDF of the parameter V against its EPDF for the same class of conductors. It is observable that the theoretical PDF and CDF are a good fit to the EPDF and ECDF respectively for all three classes of conductors. The PDF's of the two layer and single layer conductor is more right skewed than the three layer ACSR conductors confirming the already available knowledge of a decreasing fatigue lifetime with increasing number of layers of the conductor.

A problem that occurs with judging visually the closeness of the theoretical CDF to the ECDF is due to the curvature of the theoretical CDF. It is difficult to observe differences in the upper tail where the curve begins to flatten out. To remedy this, probability plots can be used to check the distribution assumption. These plots are also better in determining the appropriate distribution for small finite sample size than comparing the EPDF and theoretical PDF (Montgomery and Runger 2003). Therefore, probability paper plots are also presented to corroborate the comparison of the ECDF,

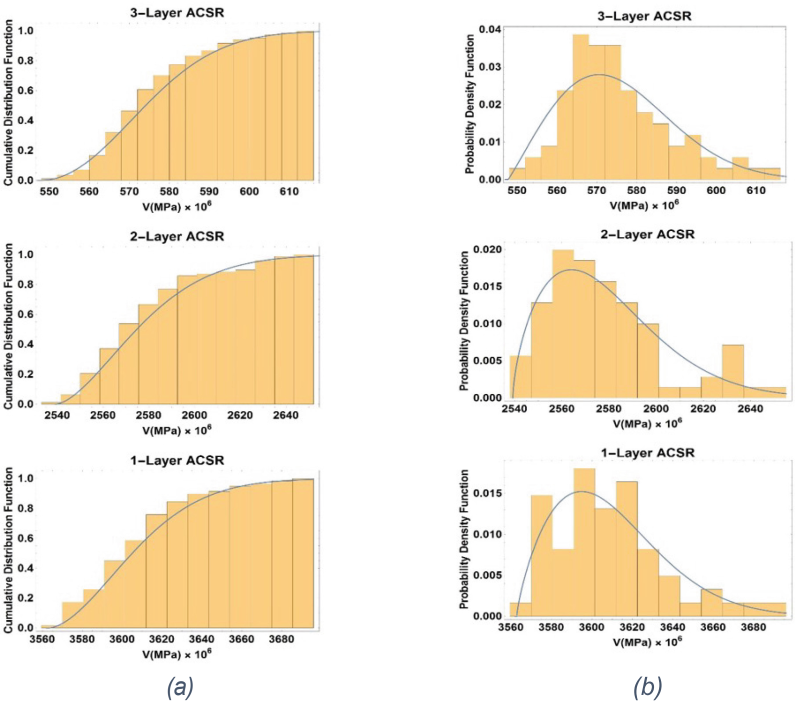


Fig. 2. (a) Comparison of CDF to ECDF (b) Comparison of PDF to EPDF

EPDF and their theoretical equivalent of the model. Approximate linearity on the probability plots allows to determine if the selected model is a plausible representation of the underlying distribution. Two types of probability paper plots namely the quantile – quantile (QQ) plot and the probability plot (PP) are used to check the model fit. The plotting coordinates of each of these probability paper plots are given respectively as follows (Crowder et al. 1991):

$$\left[\left(LogN_{1,i} - \zeta\right)\left(Log\sigma_{yb,i} - \xi\right)\right]_j, F^{-1}\left(p_j\right) \tag{4}$$

$$p_j, F\left(\left[\left(LogN_{1,i} - \zeta\right)\left(Log\sigma_{yb,i} - \xi\right)\right]_j; \hat{\alpha}\right) \tag{5}$$

where F is the selected CDF, j represents the ordering of the random sample and $\hat{\boldsymbol{\alpha}}$ the estimate of the model parameters. The point estimates of the model parameters from the ML method are used. The difference between the PP and QQ plots is in the region with highest variability (Crowder et al. 1991). In the PP plots, the points at the tails of the distribution have the lowest variability (Crowder et al. 1991). The opposite is true for the QQ plot. From an engineering point of view, the PP plot can thus be used to judge the fit of the model around the central region of the distribution while the extreme points can be judged by the QQ plot because in fatigue, the upper points are usually those with the greatest variability. The PP plots for the three classes of conductors is presented in Fig. 3a and the QQ plots in Fig. 3b. Linearity of the PP plot for all three class of conductors shows that the distribution assumption is valid. The QQ plots however show instability at the tails. Factors that could contribute to this include: the model not accounting for the effect of run - out, lesser amount of data points at the extremes, the distribution is not valid at the upper tail (D'Agostino and Stephens 1986) and the higher variability assigned to extreme points by the QQ plots (Crowder et al. 1991). Interestingly, the QQ plot for the three layer class of conductors shows a trend at the upper tail suggesting a Weibull distribution with different parameters; that is the points show a linear trend parallel to the line.

Previous models (CIGRE 1979; Hardy and Leblond 2001) have not provided any information on the region of validity of their model hence it is impossible to know where prediction can be made with the model with a high degree of accuracy. The model presented herein gives a quantitative measure of its region of validity, which can be determined from the PP and QQ plots.

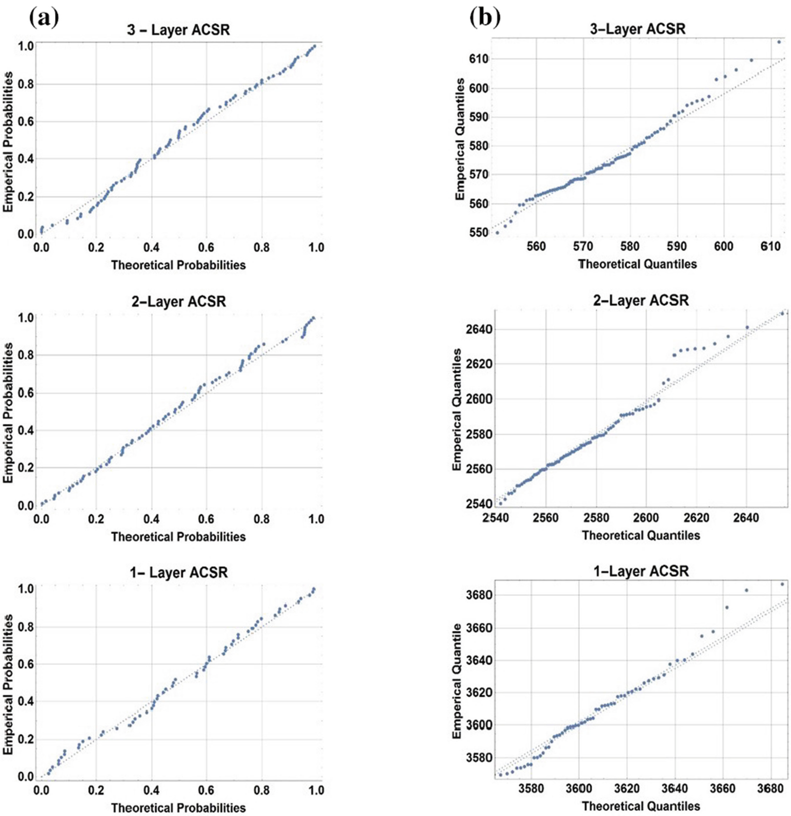


Fig. 3. (a): Probability - probability plot for all three classes of ASCR conductors; (b) Quantile - Quantile plots for three classes of ACSR conductors

5 Model Validation

To test the predictive ability of the model, validation dataset which were excluded from the training process are used. Table 1 gives the data source, the type of conductor, the stress level, data type (run out or failure), the actual number of cycles to failure recorded when the test was terminated and the probability of failure.

It is usual in the conductor fatigue literature to select a curve for failure analysis that is based on a certain probability of failure (CIGRE 1979; Hardy and Leblond 2001) and suggested as a lower bound curve. Such a selection is not made herein. However, for the sake of validation, the 50th percentile curve has been selected to predict the lifetime of the first wire failure. The model predictions are presented in Table 1. The computed probabilities of failure are computed using the actual number of cycles to failure and the stress level.

The model doesn't perform well on the run – out data. This is as expected as run – outs weren't accounted for in the model. Very good agreement is obtained between the model predicted number of cycles to failure and all data from (Fadel et al. 2012) for the Ibis 26/7 conductor. The mean number of cycles to first wire failure are approximately equivalent to that predicted by the model in Table 1 as seen in columns 5 and 6. The reason for this accuracy is that if one looks at the S-N for the class of conductors which this Ibis 26/7 belongs to (2 – layer), it is observed that those stress levels are at the extreme lower tail of the distribution. A concurrent look at the QQ and PP plots for this class of conductors shows that that the variability at the lower tail is properly captured by the model. For the Tern 45/7 conductor, the model underestimates the lifetime. To explain this, observe from the S-N for its class of conductors (3-layer) that the P – S stress level of 29.65 MPa is in the lower tail of the distribution. A concurrent look at the PP and QQ plots for the three layer class of conductors shows that the model probability of failure and quantiles are both underestimated by the model respectively. For the Bersfort conductor submitted to a P – S stress level of 11.58 MPa, the model overestimates the lifetime at the 50th percentile curve. Again, to explain this, it is observed in the S-N curve for its class of conductors (3-layer) that this stress level is in the upper tail of the distribution. A concurrent look at the QQ plots of this class of conductor shows that the model overestimates the distribution quantiles at the upper tail. This overestimation at the upper tails shows the influence of not accounting for run out on the distribution or excluding them from the analysis.

A point of caution in interpreting these results is to recall the meaning of probability from a frequentist point of view. It is expected that the frequency converges to the probability with increasing sample size (Vapnik 1998) thus it is not surprising that where the average value of a number of lifetimes is presented in Table 1, it is closer to the prediction of the model. Thus, not all the overestimation or underestimation by the model discussed above is due to tail distribution error but accounting for run-outs should decrease tail prediction error.

To further show the validity of the model, the methodology of Hardy and Leblond (2001) is compared with the model presented herein. The statistical analysis in Hardy and Leblond (2001) has been repeated with the 84 data points used for the model presented for 3 – layer class of conductors. Given the large amount of data now available, a lognormal distribution as postulated by Hardy and Leblond (2001) is used and the Strohmeyer model (Strohmeyer 1914) used to represent the 50th percentile curve as done by Hardy and Leblond (2001). To compare the performance of both models, the empirical risk of both models is computed. The model that has the minimal empirical risk is considered the better model (Vapnik 1998). The empirical risk is computed as (Vapnik 1998):

$$R_{(emp)}(\hat{\alpha}) = \frac{1}{l}\sum_{i=1}^{l}\left(LogN_{1,i} - I\left(N_{1,i}, \hat{\alpha}\right)\right)^2 \tag{6}$$

Table 1. Validation dataset for conductor vibration using bending amplitude stress indicator

Data source	*Conductor type*	$(\sigma_{a(Yb)})$ *(MPa)*	*Data type*	*Actual Number of cycles to failure* $(*10^6)$	*Predicted number of cycles to failure* $(*10^6)$	*Probability of failure*
Cloutier et al. (2006)	Rail 45/7	10.53	Run out	318.07	163.93	0.66
	Crow 54/7	16.96	Run out	24.78	11.92	0.69
Goudreau et al. (2010)	Bersfort 48/7	11.58	Failure	71.74	102	0.40
	Tern 45/7	29.65	Failure	1.11	1.00	0.61
		29.65	Failure	1.87	1.00	0.74
		29.65	Failure	2.74	1.00	0.82
Fadel et al. (2012)	Ibis 26/7	25.08	Failure	5.50	6.67	0.42
		25.08	Failure	2.98	6.67	0.18
		28.22	Failure	3.00	3.92	0.39
		28.22	Failure	1.90	3.92	0.21
		31.35	Failure	2.47 (mean)	2.45	0.5
		34.49	Failure	1.13 (mean)	1.61	0.36
		39.82	Failure	1.00 (mean)	0.86	0.56
		43.31	Failure	0.53 (mean)	0.59	0.45

Where the function I is the 50th percentile curve and l is the length of the dataset. The obtained empirical risks are presented in Table 2. On this basis, the model presented herein outperforms the previous model on the 3-layer ACSR class of conductors and is a better model in predicting the time to failure of the first wire of the 3-layer class of conductor.

Table 2. Comparison of empirical risk

Model	$R_{(emp)}(\alpha)$
Present model	1.62
Hardy and Leblond (2001)	2.29

Because the empirical risk is also a measure of the model error, the model presented herein has lower prediction error and produces tighter confidence interval about the 50th percentile curve than the type of model presented in Hardy and Leblond (2001).

6 Conclusion and Subsequent Research

A statistical analysis on the ACSR conductor fatigue data has been conducted. Non-constant variance of the conditional density of lifetime given stress level has been considered by the model. The region of validity of the model has been shown using quantile – quantile probability paper plots and probability – probability paper plots. The Normalization variable used by the model allows to construct the empirical probability density and cumulative distribution functions. It is shown that the theoretical CDF and PDF for all classes of conductors examined agree with the ECDF and EPDF respectively. The model prediction capability is demonstrated by comparing predictions of the model with a validation dataset. Good agreement is obtained in the region of the mean. It is shown that the effect of run – out on the conductor must be accounted for in conductor fatigue statistical analysis to improve the tail behaviour. Comparison of the present model and a previous model showed that the present model provide a lower prediction error and tighter confidence intervals. However, a lacuna of the present model is that one of the marginal densities doesn't exist. Nonetheless, the emphasis is on predicting as closely as possible the time to first wire failure and the model has demonstrated its capability for this. Further analysis is required to refine the model and to ensure that it provides the minimum possible risk in prediction, which will guarantee minimum error in residual life estimation of transmission line conductors subjected to aeolian vibrations.

Acknowledgements. This research project was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, Hydro – Quebec, Reseau de Transport d'Electricite and InnovEE Canada.

References

Arnold, B.C., Castillo, E., Sarabia, J.M.: Conditional Specification of Statistical Models, 1st edn. Springer, New York (1999)

Besag, J.: Spatial interaction and the statistical analysis of lattice systems. J. Roy. Stat. Soc.: Ser. B (Methodol.) **36**(2), 192–236 (1974)

Bhattacharyya, A.: On some sets of sufficient conditions leading to the normal bivariate distribution. Sankhya: Indian J. Stat. (1933–1960) **6**(4), 399–406 (1943)

Castillo, E.: Extreme value Theory in Engineering. Academic Press Inc., Carlifornia (1988)

Castillo, E., Fernandez-Canteli, A.: A Unified Statistical Methodology for Modeling Fatigue Damage. Springer, Dordrecht (2009)

Castillo, E., Galambos, J.: Lifetime regression models based on a functional equation of physical nature. J. Appl. Probab. **24**(1), 160–169 (1987)

CIGRE: Recommendation for the evaluation of the lifetime of transmission line conductors (Electra n 63) (1979)

Cloutier, L., Goudrea, S., Cardou, A.: Fatigue of overhead conductors. In: Chan, J.K., Havard, D.G., Rawlins, C.B., Weisel, J. (eds.) EPRI Transmission Line Reference Book: Wind-induced Conductor Motion, 2nd edn. EPRI, Palo Alto (2006)

Cousineau, D.: Fitting the three-parameter Weibull distribution: review and evaluation of existing and new methods. IEEE Trans. Dielectr. Electr. Insul. **16**(1), 281–288 (2009). https://doi.org/10.1109/TDEI.2009.4784578

Crowder, M.J., Kimber, A.C., Smith, R.L., Sweeting, T.J.: Statistical Analysis of Reliability Data. Springer Science+Business Media, Berlin (1991)

D'Agostino, R.B., Stephens, M.A.: Goodnss of Fit Techniques. Marcel Dekker Inc., New York (1986)

Fadel, A.A., Rosa, D., Murça, L.B., Fereira, J.L.A., Araújo, J.A.: Effect of high mean tensile stress on the fretting fatigue life of an Ibis steel reinforced aluminium conductor. Int. J. Fatigue **42**, 24–34 (2012). https://doi.org/10.1016/j.ijfatigue.2011.03.007

Goudreau, S., Lévesque, F., Cardou, A., Cloutier, L.: Strain measurements on ACSR conductors during fatigue tests II - stress fatigue indicators. IEEE Trans. Power Delivery **25**(4), 2997–3006 (2010). https://doi.org/10.1109/TPWRD.2010.2042083

Hanaki, S., Yamashita, M., Uchida, H., Zako, M.: On stochastic evaluation of S-N data based on fatigue strength distribution. Int. J. Fatigue **32**(3), 605–609 (2010). https://doi.org/10.1016/j.ijfatigue.2009.06.001

Hardy, C., Leblond, A.: Statistical analysis of stranded conductor fatigue endurance data. Paper presented at the proceeding of the fourth international symposium on cable dynamics, Montreal, Canada (2001)

Kotz, S., Nadarajah, S.: Extreme Value Distributions: Theory and Applications. Imperial Collge Press, London (2000)

Marshall, A.W., Olkin, I.: Life Distributions. Springer, New York (2007)

Montgomery, D.C., Runger, G.C.: Applied Statistics and Probability for Engineers, 3rd edn. Wiley, New York (2003)

Noiseux, D.U., Hardy, C., Houle, S.: Statistical methods applied to aeolian vibrations of overhead conductors. J. Sound Vibr. **113**(2), 245–255 (1986)

Poffenberger, J.C., Swart, R.L.: Differential displacement and dynamic conductor strain. IEEE Trans. Power Apparatus Syst. **84**(4), 281–289 (1965). https://doi.org/10.1109/TPAS.1965.4766192

Strohmeyer, C.E.: The determination of fatigue limits under alternating stress conditions. Paper presented at the proceedings of the Royal Society of London (1914)

Suresh, S.: Fatigue of Materials, 2nd edn. Cambridge University Press, Cambridge (1998)

Thi-lien, D.: Analyse des essais a amplitudes variables realises sur le conducteur Aster 570, France (2015)

Vapnik, V.N.: Statistical Learning Theory. Wiley, New York (1998)

Condition Assessment of Norwegian Bridge Elements Using Existing Damage Records

Andreu Regué Barrufet[1](✉) and Ephrem Taddesse[2]

[1] Norwegian Public Roads Administration, University of Agder, Kristiansand, Norway
andreu.barrufet@vegvesen.no
[2] University of Agder, Kristiansand, Norway
ephrem.taddesse@uia.no

Abstract. The Norwegian Public Roads Administration (NPRA) has recorded bridge element damages in a database for all the bridges it manages since the 1990s. This paper presents a comparison of three methods to establish element condition based on damage records. The methods consist in a non-parametric procedure based on the worst damage registered in the element, linear regression considering also bridge and road characteristics data and classification through an artificial neural network. The methods are assessed using a set of 159 bridges inspected in 2016. The results show that diagnostics of bridge element condition can reach high accuracy by using an artificial neural network classifier and taking into account existing damage records and bridge data.

1 Introduction

Deterioration models are a key component for the optimal management of an agency's bridge infrastructure. Statistical deterioration models need to be developed and calibrated with historical data on the condition of bridge elements. The NPRA has recorded bridge element damages for all the bridges it manages since the late 1990s. This paper presents an analysis of existing inspection records and their suitability for assessing element condition.

Bridge and damage data is recorded in a computerized database accessible to bridge inspectors through a web interface. The system is called Brutus (NPRA 2017).

1.1 A Summary of Bridge Element Damage Recording in Norway

The NPRA performs regular bridge inspections every year and every five years. Inspections are conducted according to the manual for bridge inspection, Håndbok V441 (NPRA 2014). The result of the inspection is a record of all the damages detected on each bridge element. Damages are recorded according to a predefined schema that contains six fields.

- **Skadetype**. Predefined damage type (spalling, cracking, etc.).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 716–722, 2020.
https://doi.org/10.1007/978-3-030-48021-9_79

- **Skadekonsekvens**. An assessment of damage consequences. The system allows registering consequences for load-bearing capacity (B), traffic safety (T), future maintenance costs (V), and environmental consequences (M).
- **Skadegrad**. Degree of damage (DOD) for each kind of consequence. The DOD range goes from 1 to 4, where 1 means least damage and 4 means most damage.
- **Skadeårsak**. Probable cause of damage, chosen from a predefined list. More than one can be selected.
- **Skadebeskrivelse**. Free text description. It is encouraged that bridge inspectors record the position and extension of the damage, results of tests and measurements, and any other relevant comments.
- Optional damage images (Table 1).

Table 1. Damage recording for damages of type B as described in the Inspection manual for bridges (Vegdirektoratet, 2000/2014 p62–64)

Damage	Description
1B	Small damage/deficiency not considered to represent any risk for the load-bearing capacity of the bridge. The damage shall not be repaired
2B	Medium damage/deficiency that can reduce the load-bearing capacity of the bridge if allowed to stay for more than 4-10 years. The damage shall be repaired within 4–10 years
3B	Big damage/deficiency that can reduce the load-bearing capacity of the bridge if allowed to stay for more than 1–3 years. The damage shall be repaired within 1–3 years
4B	Critical damage which has reduced or is about to reduce the load-bearing capacity of the bridge. The damage shall be repaired immediately or at most within ½ year

The bridge inspector must decide, based on prior experience and knowledge, how fast a given damage may develop in order to rate it correctly. In practice, the inspection manual provides a comprehensive catalogue of damage pictures and descriptions for comparison.

Around year 2013, the bridge management system of the NRPA was updated. The degree of damage is now divided into two components. Those components are damage degree (Skadegrad) and damage consequence (Skadekonsekvens), as in the standard. Both parameters range from 1 (best) to 4 (worst) but only a single condition degree may be assigned to a damage, while different consequence degrees may be assigned (one for impact on future maintenance costs and a different one for load-bearing capacity, for example).

The change was implemented in the bridge management system by adding a new field called "degree of damage" (skadegrad in Norwegian) while the old field, which was named "degree of damage/-consequence", and which was a combination of numerical and alphabetic values, was renamed as simply "damage consequence" (skadekonsekvens in Norwegian). Old record values were kept in the renamed "damage consequence" field.

1.2 Overview of Bridge Performance Indicators

Clear performance measures or indicators can help agencies to assess the extent to which a bridge program, project, or policy has succeeded in achieving intended goals and objectives. Yet performance measures should be realistic and operational. Data relating to the performance should be measured with available resources without excessive effort, cost or time. (Patidar et al. 2007).

In Europe, every country uses a different set of performance indicators. Cost Action TU1406: Quality specifications for roadway bridges, standardization at a European level (Bridgespec) – Performance indicators (Matos et al. 2016) aims to set a standard in both performance indicators and performance goals. The performance indicator database compiled as part of that project shows a common reliance on visual inspection for assessing most of technical indicators of element condition. Non-destructive testing, probing and structural health monitoring are used to a much lesser extent.

The taxonomy of indicators developed by Cost Action TU1406 as described by Strauss et al. (2017) provides a reference framework. Performance indicators can be set at the element level, system (bridge) level or network level. Technical indicators at the element level are those intended to assess damages or element functionality.

1.3 Bridge Element Condition Assessment for Use in Degradation Modelling

Bridge element deterioration is usually characterized with a statistical degradation model, such as a Markov process. The Markov process defines the probability that the condition of an element worsens from a given condition state to a worse one in a specified interval of time. Parameter estimation for a Markov process works best when the number of possible states is between 4 and 6. For example, AASHTO BrM, formerly known as Pontis, uses up to 5 states (Golabi and Shepard 1997).

The condition assessment of bridge elements is based on visual inspections. Bridge inspectors assign a condition state to each bridge element based on the damages observed and a pre-defined list of criteria for each element type.

Condition states describe the deterioration of bridge elements, but the relationship with load-bearing capacity or structural reliability may be weak. Yet condition states are often used for maintenance planning as they describe the extent of deterioration processes that if untreated may impair the bridge serviceability. In addition, they describe visible processes that may cause user concern if not addressed.

2 Methodology

Data was obtained from Brutus. A training data set was manually rated. After that, three models were programmed using the Pyhton programming language in a notebook environment (Pérez and Granger 2007). The first one was a non-parametric model; the second one was a multiple linear regression and the last, an artificial neural network. The models were fitted and cross-validated using the Scikit-learn library (Pedregosa et al., 2011).

2.1 Regularization of Element Condition Records

There are two different series of condition records in Brutus for most elements. One where degree of damage/consequence is summarized in one value and another one where it is split in two values, as described in Sect. 1.1.

An analysis of 73 378 structural damages registered for the first time before the new system with two numbers was implemented, shows that the value of consequence degree used in the new system is the old degree of damage in most cases.

For damages recorded for the first time after the new system was implemented it is also the case that most of the time, the consequence degree determines the consequence damage. Damage information can thus be summarized in a single value by the consequence degree and the degree of damage can be ignored.

2.2 Assessing Element Condition

The element as a whole is rated. The purpose of assigning a single condition number to every element is to characterize it in a general way, not specific to the type of damage. This allows for the development of a statistical degradation model.

It is desirable that the number of condition states be equal or less than the number possible degree of damage for each damage recorded in Brutus in such a way that a correspondence between damage degree and condition state can be easily established. Therefore, the following condition rating scale was proposed (Table 2).

Table 2. Description of proposed element condition states

Condition state	Description
0	No defects
1	Early signs of defect. Normal wear
2	Minor surface defects. Local spalling without reinforcement steel section loss. Minimal scour. Loss of surface treatment in steel elements
3	Severe defect. Comprehensive spalling with reinforcement steel section loss. Scour. Local corrosion with section loss in steel elements
2	Concern for service or limit state of element

925 bridge elements pertaining to 152 different bridges and culverts were rated according to the rating scale proposed. Element types where subsequently grouped into five main types; bearing (120 instances), culvert element (29), infrastructure (370), superstructure (391), and steel pipe culvert (15).

31% of elements were in condition 0, 49% in condition 1 and 19% in condition 2. There were only 12 elements in condition 3 and two in condition 4. For further analysis, condition 3 and 4 were grouped into a single category.

2.3 Non-parametric Method

A method loosely based on the current procedure for assessing bridge condition consists in going through all the damages registered for an element and computing the maximum value of consequence V or consequence B. This maximum value would be the elements condition. The advantage of this method is that the resulting indicator is readily understandable and can be related back to the damage classification scale.

2.4 Multiple Linear Regression

The second procedure tested consisted in building and fitting a classifier based on multiple linear regression. To validate the model 10-fold cross validation with stratified sampling was used.

Recursive feature elimination was used to assess the optimum number of features to incorporate into the model.

2.5 Artificial Neural Network

The last classifier tested was a neural network (multi-layer-perceptron with back-propagation) with two hidden layers. Data is normalized before being fed to the classifier. A grid search over the size of each layer, activation function and regularization parameter was performed. Again, the model was optimized using 10-fold cross validation with stratified sampling.

The features used for fitting the model were the same as the initial set of features used for the multiple linear regression model.

3 Results

3.1 Non-parametric Method

By applying the method described in Sect. 2.3, the percentage of elements correctly classified was 67%.

3.2 Multiple Linear Regression

The highest accuracy was obtained using 52 features out of 60 possible (Table 3).

The average classification accuracy using multiple linear regression is 75.9% with a 95% confidence interval of (69.1%, 82.7%).

3.3 Artificial Neural Network

The highest accuracy was obtained using two hidden layers of size 60 and 40, rectified linear unit function ($f(x) = \max(0, x)$) as activation function and an L2 penalty (regularisation term) parameter of 1.0.

The mean accuracy obtained with the ANN is 79.7% with a 95% confidence interval of (72.6%, 86.8%).

Table 3. Features used in the optimal set.

Feature group	Feature	Part of best model?
None	AADT	No
	Any damages	Yes
	Age	No
	Highest damage consequence	Yes
Element type	Bearing, Culvert element, Infrastructure, Pipe culvert, Superstructure	Yes
Material	Concrete, pre-stressed concrete, steel, masonry	Yes
Design standard	12 possible alternatives, from SVV 1/30 to other/unknown	Yes
Climatic zone	Coastal, inner coastal, inland	Yes
Damage type	Other, leakage, lack of part, material-independent damage, damage on concrete honeycombing	No
	24 other types of damage	Yes

4 Discussion

The learning curve of the ANN model shows a significant gap between the training dataset and the test dataset. The accuracy of the training dataset is quite constant and close to 95% as the number of training samples increases, while the accuracy on the test set keeps increasing. It is to be expected that the accuracy on the test dataset would continue to increase.

In fitting the models, element data has been treated as being independent, but it is reasonable to expect some correlation between elements pertaining to a single bridge. The effects that this may cause in the models have not been researched.

5 Conclusions

This article demonstrates the use of machine learning techniques to assess the condition of bridge elements based on damage records and other bridge data.

The ANN approach is slightly more accurate than the multiple linear regression, but more importantly, the learning curve shows room for improvement by increasing the training set. Accuracy rates near 90% should be possible.

The automatic assessment of bridge condition based on machine learning techniques has several advantages. It relies on quite objective data, such as damage records and bridge data, combined with a training dataset assessed in advance. Therefore, some of the variability due to the subjective criteria of bridge inspectors is eliminated. The training dataset can be small enough to be compiled by a single group of experts. If needed, the training dataset can be reassessed to recalibrate the model without the need to manually reassess each element individually.

Finally, the results demonstrate that this approach can be used to assess the historical condition of bridge elements based on the records kept. The assessment of the historical condition and its changes can be used as input data to calibrate degradation models. Such models can be used to assess the remaining useful life of bridge elements, determine expected time to intervention according to predefined condition levels and plan risk-based inspection intervals, instead of current time-based inspections.

The condition assessment of bridge elements is a first and necessary step to implement bridge management systems that optimize both inspection and intervention costs.

References

Golabi, K., Shepard, R.: Pontis: a system for maintenance optimization and improvement of US bridge networks. Interfaces **27**(1), 71–88 (1997)

Matos, J.C., Casas, J.R., Strauss, A., Fernandes, S.: COST ACTION TU1406: quality specifications for roadway bridges, standardization at a European level (bridgespec)–performance indicators. In: fib Symposium 2016 (2016)

NPRA: Håndbok V441: Inspeksjonshåndbok for bruer. Vegdirektoratet (2014)

NPRA: Bruforvaltning (2017). https://www.vegvesen.no/fag/teknologi/bruer/bruforvaltning. Accessed 14 May 2018

Patidar, V., Labi, S., Sinha, K., Thompson, P., Shirolé, A., Hyman, W.: Performance measures for enhanced bridge management. Transp. Res. Rec. J. Transp. Res. Board **1991**, 43–53 (2007)

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**(1), 2825–2830 (2011)

Pérez, F., Granger, B.E.: IPython: a system for interactive scientific computing. Comput. Sci. Eng. **9**(3), 21–29 (2007)

Strauss, A., Bergmeister, K., Ivanković, A.M., Matos, J.C.: Applied and research based performance indicator database for highway bridges across Europe. In: IALCCE 2016, The Fifth International Symposium on Life-Cycle Civil Engineering, pp. 1503–1510. Taylor and Francis (2017)

Gas-Liquid Ratio Imbalance Indicators of a Fighter Aircraft Shock Absorber

Arttu Heininen[1(✉)], Jussi Aaltonen[1], Kari T. Koskinen[1], and Juha Huitula[2]

[1] Tampere University of Technology, Tampere, Finland
{arttu.heininen, jussi.aaltonen, kari.t.koskinen}@tut.fi
[2] Finnish Defense Forces Logistics Command, Tampere, Finland
juha.huitula@mil.fi

Abstract. An oleo-pneumatic shock absorber of a fighter aircraft loses some of its nitrogen charge and oil fill over time, due to leakage through seals. Usually, only gas can be added to shock absorber in line maintenance. This creates an imbalance to the ratio between gas and oil, and affects both the damping and the stiffness of the shock absorber. Measuring the gas and liquid content inside a shock absorber is infeasible. Thus, other variables that can act as indicators of an imbalance, must be identified. A coupled model of a fighter aircraft landing gear and its shock absorber is created. Using the model, landings with different sink speeds and aircraft masses has been simulated. Simulations with a standard, and later imbalanced, gas-liquid ratio are conducted. Results from these simulations are discussed and presented here, and variables that can be used as indicators of this imbalance are identified. This information can be used to potentially improve maintenance planning.

Nomenclature

A	Orifice area
C	Damping coefficient
C_L	Lift coefficient
C_q	Flow coefficient
F	Reaction force
F_C	Friction force
g	Gravitational acceleration
K	Spring stiffness
L	Lift
m	Aircraft mass
p	Pressure
S	Wing area
ρ	Density
v	Horizontal velocity
$\dot{x}$	Stroke velocity
x	Stroke

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 723–730, 2020.
https://doi.org/10.1007/978-3-030-48021-9_80

1 Introduction

One of the main parts of an aircraft landing gear is its shock absorber, which dissipates the kinetic energy related to landing to heat. Almost every medium to large sized modern aircraft uses an oleo-pneumatic shock absorber in its landing gear. In the shock absorber hydraulic fluid is forced through an orifice, which dampens the motion. Also, it has a nitrogen or an air charge that acts as a spring. The fluid flow is controlled by a metering pin or a hydraulic valve. To reach a high efficiency, the flow rate is varied during stroke. This is achieved with a varying metering pin cross-section or with a control valve. Using a metering pin is simpler, more reliable and maintenance-free (Currey 1988).

Common carrier-based fighter aircraft have a shock absorber with two chambers. A primary chamber filled with fluid and gas, and a secondary chamber filled only with gas. The secondary chamber acts as a secondary spring and activates during hard landings, i.e., when the dissipated kinetic energy is high due to a high sink speed, i.e., the vertical velocity at touchdown (Niu 1988). A schema of a typical fighter aircraft shock absorber is shown in Fig. 1. High temperature and pressure fatigues seals and O-rings, which leads to leakage due to deformation (Yang et al. 2016). Even without a full seal failure, pressure is lost over time.

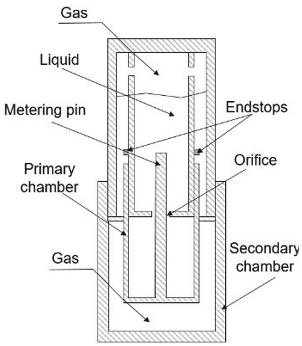


Fig. 1. A schematic view of a two-chamber oleo-pneumatic shock absorber.

As the pressure loss is compensated with added gas, there is an imbalance i.e., the proportion between gas and liquid is distorted. This causes a malfunction in the shock absorber operation. Measuring the amount of liquid or gas is infeasible. Therefore, other variables that are more easily measured and describe the condition inside the shock absorber, must be identified. Measuring everything is inefficient and costly, and a landing gear is under heavy loads and vibrations during landing, which complicates measurements. Furthermore, aerospace regulations require that the measuring system does not interfere with safety or flightworthiness (Phillips 2011). Other restrictions concerning the size, weight, and complexity of the system must also be considered.

As field measurements are expensive and difficult, mathematical simulation and modelling can be used as a tool to find the indicating variables. To have realistic results, a coupled landing gear and shock absorber model of a fighter aircraft is used to simulate real landings. In this paper, using a coupled model, the behaviour of the main variables is analysed. The main variables are stroke velocity, stroke, temperature, and pressure inside the shock absorber primary chamber. Firstly, design values for the gas-liquid ratio are used to understand how the variables behave in different landings. Secondly, the gas-liquid ratio is distorted to have an imbalance, and a similar analysis is made. The objective is to find out the variables that can be used as indicators of the gas-liquid ratio imbalance.

Construction of the shock absorber requires that the shock absorber is removed from the aircraft for full service. Thus, a usual line maintenance service procedure is to check pressure of the shock absorber and add gas to it, if required. This can be done on ground. A full service is done as shop maintenance periodically, or if there is an obvious need for it. However, the aircraft industry is moving from traditional preventive condition monitoring methods towards predictive methods (Phillips 2010). This concerns military operators also, as they are required to operate economically. To achieve this, the identified indicators should be measured during landing. The knowledge of the level of the imbalance inside the shock absorber can be used to decide if the shock absorber requires full service or not. This helps in the maintenance planning, making it potentially more cost-effective.

2 Coupled Landing Gear and Shock Absorber Model

The coupled model includes a landing gear model and a shock absorber model. The landing gear is modelled as a multibody system (MBS model) and the shock absorber as a one degree-of-freedom dynamical multi-domain system (dynamical model). The coupling is in the equation of motion which, just after touchdown, is:

$$F = mg - F_c - L - C\dot{x} - Kx. \tag{1}$$

The MBS model calculates the stroke velocity and the stroke, and these are transferred through a control interface to the dynamical model. Using the transferred variables, the dynamical model calculates the shock absorber reaction force using Eq. 1. This is repeated during every time step. The operating principle of the coupled model is shown in Fig. 2. The MBS model also handles the calculation of lift, which depends on the initial orientation and horizontal velocity.

$$L = \frac{1}{2}\rho v^2 S C_L. \tag{2}$$

The dynamical model calculates the friction force using LuGre-friction model (Lund-Grenoble) (Åström 2008). The damping properties are related to the varying cross-section orifices, whose mass flow is calculated from

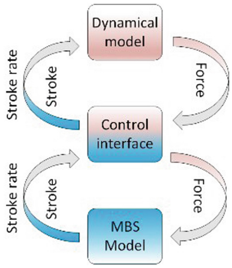


Fig. 2. A schematic view of the coupled simulation model.

$$dm = C_q A(x)\sqrt{2\rho \Delta p}. \tag{3}$$

The gas spring compression is often modelled as an adiabatic process using the ideal gas law. However, it becomes inaccurate at high pressure so Peng-Robinson model is used (Peng and Robinson 1970). The damping coefficient and the spring stiffness depend on the stroke and the shock absorber is highly non-linear to avoid considerable rebound (Wright 2014). A detailed presentation and validation of the shock absorber model can be found in (Heininen 2015).

3 The Influence of the Aircraft Mass and Sink Speed on the Main Variables

The kinetic energy that the shock absorber dissipates is proportional to the square of the sink speed and directly proportional to the mass of the aircraft. Increasing the sink speed or the aircraft mass means that more energy has to be dissipated. This should be seen as a change in the main variables between simulations. Three values for the aircraft mass were chosen: 13 608, 15 422, 17690 kg. These represent a wide but realistic range of different fighter aircraft landing weights. Also, three different sink speeds were chosen: 3.0, 4.0, and 4.9 m/s. Maximum values of each main variable during simulation is shown in Fig. 3. A non-linear polynomial curve has been fitted along the points in each dataset.

The maximum stroke, pressure, and temperature are non-linear with respect to the kinetic energy. However, the relation between the maximum stroke velocity and the aircraft mass is almost non-existent. In Fig. 4, the stroke velocity is shown during simulation. M1V1 refers to case with lightest aircraft and slowest sink speed, and M3V3 refers to heaviest and fastest. It shows the same behaviour. The differences between simulations with different mass is insignificant. The relation between sink speed and stroke velocity is non-linear. Faster sink speeds are damped relatively more than slower ones.

In conclusion, the stroke velocity can be eliminated from the variables that are studied in the distorted gas-liquid ratio simulations. The shock absorber model is adiabatic, i.e., all the dissipated heat is transferred to the fluid. This is a valid assumption, because a landing is a fast process, and therefore, heat transfer from the

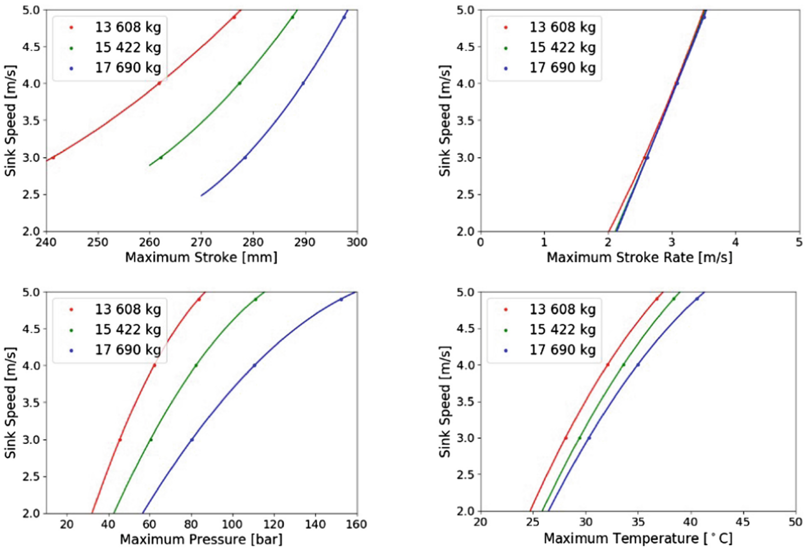


Fig. 3. Sink speed versus maximum stroke, stroke velocity, pressure and temperature inside the primary chamber with different aircraft masses and sink speeds.

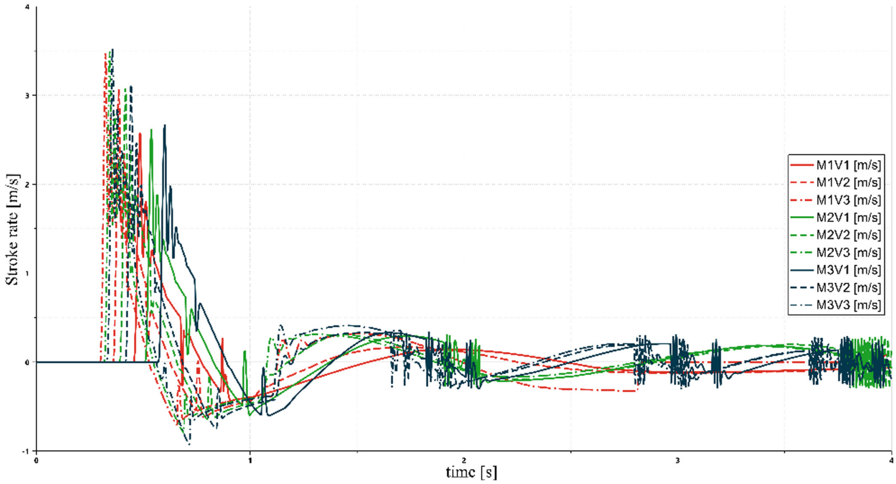


Fig. 4. Stroke velocity vs time during landing simulation.

shock absorber to the environment is negligible. Temperature change from the initial value to the maximum is distinguishable between different landing cases. Same applies for the maximum stroke and pressure. These three variables must be included in the gas-liquid imbalance analysis.

4 Imbalance in the Gas-Liquid Ratio

To find out which variables would indicate an imbalance in the gas-liquid ratio, the ratio was distorted for the simulations. The same cases as in the previous chapter were simulated. This time some of the hydraulic oil (2, 5, 10, and 20%) was replaced with gas. Again, the maximum values of the remaining variables were considered. The maximum temperature of the cases with lightest and heaviest aircraft loading is shown in Fig. 5. When the vertical kinetic energy is low, the effect of the imbalance on the maximum temperature is not clearly seen. However, as the energy increases so does the influence of the distorted gas-liquid ratio. This indicates that as a stand-alone variable temperature is not fitting as an indicator of the imbalance, but it should not be entirely disregarded.

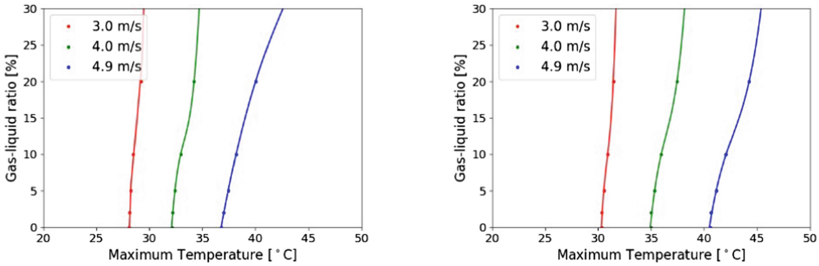


Fig. 5. Gas-liquid ratio versus maximum temperature. Left: aircraft mass 13 608 kg. Right: aircraft mass 17690 kg.

Figure 6 shows the distorted gas-liquid ratio versus the maximum stroke and pressure. The distortion affects the maximum stroke only slightly. If the distortion is small, the difference in maximum stroke can be less than one percent. This is due to increasing the gas stiffness non-linearity near the end of the stroke. The impact force related to a landing generates vibrations in multiple degrees-of-freedom and noise (Jebáček 2016). This makes the measuring of the stroke accurately during a landing difficult, as relatively small differences must be identified from a noisy signal. This means that stroke is not a suitable variable, as an imbalance indicator.

The maximum pressure is sensitive to the gas-liquid imbalance. A small change in the gas content shows a distinct difference in the maximum pressure. This is valid even for landings with low kinetic energy. As the gas amount increases, the gas spring stiffness decreases strongly. Commonly, a 4th generation fighter aircraft have a mechanical pressure gauge in the shock absorber. Replacing a pressure gauge with a measurement system that measures both temperature and pressure actively during landing is technically feasible, without risking the airworthiness. Therefore, pressure and temperature are identified as variables that together can indicate there is an imbalance in the gas-liquid ratio.

Using this kind of measuring system, the imbalance could be observed which would help to plan the maintenance accordingly. Instead of periodically servicing the shock absorber, the level of imbalance could be measured, and the decision to service

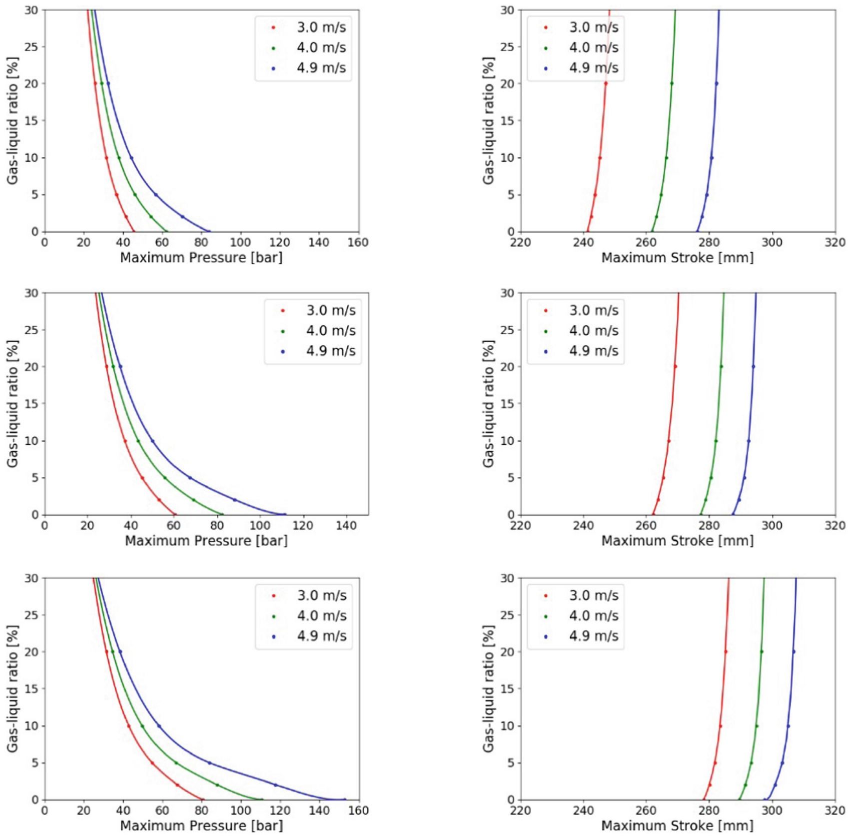


Fig. 6. Gas-liquid ratio versus maximum pressure and stroke. Top: aircraft mass 13 608 kg. Middle: aircraft mass 15 422 kg. Bottom: aircraft mass 17690 kg.

the shock absorber could be based on this information. The imbalance is generated over time as the number of landings increases. So, a certain limit that determines the need for full service could be defined.

5 Conclusions

Using a coupled landing gear and shock absorber model the behaviour of the main variables of a fighter aircraft shock absorber was studied. The main variables include stroke velocity, stroke, and temperature and pressure in the shock absorber primary chamber. Their usefulness as an imbalance indicator of the gas-liquid ratio was investigated. It was noticed that the maximum stroke velocity is mostly influenced by the sink speed making it unsuitable as an indicator. The gas spring stiffness is highly non-linear, and the differences in the stroke are small. Measuring these small differences with the required accuracy during landing is difficult, as there are heavy loads and vibration. Therefore, the stroke is not suitable as an indicator.

The influence of distortion to the temperature inside the shock absorber is noticeable in landings with high vertical kinetic energy. Also, the inside pressure shows clear difference, when the ratio is distorted. Even when the kinetic energy related to the landing is low. Technically active temperature and pressure measurements are possible. Therefore, these two variables were identified as good indicators of an imbalanced gas-liquid ratio. The level of imbalance could be monitored by measuring these two indicators. This information can be used to improve maintenance planning in the future.

References

Currey, S.: Aircraft landing gear design: principles and practices. In: AIAA, p. 373 (1988)

Heininen, A.: Modelling and simulation of an aircraft main landing gear shock absorber. Master's thesis, p. 56 (2015). https://dspace.cc.tut.fi/dpub/handle/123456789/23593

Jebáček, I., Horák, M.: Measuring of a nose landing gear load during take-off and landing. trans tech publications. Appl. Mech. Mater. **821**, 325–330 (2016)

Åström, K., Canudas-De-Wit, C.: Revisiting the LuGre friction model. IEEE control Syst. **28**(6), 101–114 (2008)

Niu, C.: Airframe Structural Design: Practical Design Information and Data on Aircraft Structures. Conmilit Press, p. 612 (1988)

Peng D.Y., Robinson D.B.: A new two-constant equation of state. In: 4th International Heat Transfer Conference, pp. 59–64 (1970)

Phillips, P., Diston, D., Starr, A., Payne, J., Pandya, S.: A review on the optimisation of aircraft maintenance with application to landing gears. In: Engineering Asset Lifecycle Management, pp. 68–76. Springer, London (2010)

Phillips, P., Diston, D.: A knowledge driven approach to aerospace condition monitoring. Knowl.-Based Syst. **24**(6), 915–927 (2011)

Yang, H., Wan, F., Cui, W.: Failure simulation and identification of shock absorber in carrier-based aircraft landing gear. In: 2016 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 1–6 (2016)

Wright, J.R., Cooper, J.E.: Introduction to Aircraft Aeroelasticity and Loads, p. 605. Wiley, Hoboken (2014)

Observation and Processing of Instantaneous Frequency Variations During Bearing Tests

Tor Inge Waag[1(✉)] and Martin Hemmer[2,3]

[1] NORCE Norwegian Research Centre AS, Grimstad, Norway
towa@norceresearch.no
[2] University of Agder, Grimstad, Norway
martin.hemmer@uia.no
[3] MHWirth, Butangen 21, 4639 Kristiansand, Norway

Abstract. Laboratory experiments have been performed on medium sized roller bearings with two levels of artificial damage. Recordings of long time series from accelerometers at a wide range of different radial loads and rotation speeds has been performed. Probably due to non-perfect performance of the control systems for the rotational speed or frequency, significant fluctuations were observed at all rotation speeds. The highest relative variation was observed at the lowest rotational speeds. These variations were recorded using a rotary encoder, which allows order tracking of the vibration signal. In real life condition monitoring, tachometers or rotary encoders are not always present, this can be due to the cost of installation or space limitations for introducing additional sensors. The present work describes a method for instantaneous frequency estimation, making it possible to perform order tracking without relying on a rotary encoder. These results are then verified by comparison to the rotary encoder data.

1 Introduction

In all condition monitoring of rotating applications, there is an advantage in using information from tachometers or rotary encoders. Any fluctuation in shaft instantaneous frequency cause frequency smearing, which can make even relatively serious damage hard to diagnose through frequency analysis. However, encoder data is not always present, and it can be both difficult and expensive to retrofit new sensors. The difficulties can be due to access problems, temperature, explosive atmosphere and the cost of cabling, scaffolding and restrictions on wireless communication. For the cases where a good tachometer or encoder signal cannot be provided, it is therefore of interest to investigate if there are ways to extract information about instantaneous speed variations of rotating machinery for the purpose of improving spectral resolution of vibration data. Several methods for order tracking exist. Bonnardot et al. (2004) presented a method for envelope demodulation under very limited fluctuations in speed. Zhao et al. (2013) and Wang et al. (2016) utilizes the generalized Fourier Transform and maxima tracking to isolate high energy harmonics used for demodulation. Schmidt et al. (2018) improves maxima tracking by incorporating acceleration information. These methods all utilize instantaneous phase estimation from the analytic signal.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 731–739, 2020.
https://doi.org/10.1007/978-3-030-48021-9_81

The authors have performed laboratory tests using vibration sensors on medium sized rolling element bearings. For these experiments a rotary encoder was present, and useful for order tracking and increasing the resolution of envelope spectra. However, the same data sets can be processed omitting the encoder data to check if similar resolution improvement can be achieved from information inherent in the vibration data. The presented method for utilizing transients in the vibration signal as a replacement for tachometer pulses to perform order tracking. The method is in the "elapsed time" category (André et al. 2013), utilizing peak detection of transients in the vibration signal to replace encoder pulse and subsequent outlier identification to stabilize RPM estimates.

It is required that the transients are detectable, but they could originate from bearing damage, gear meshing or other phenomena that are related to shaft rotation. For efficient implementation, one should have an idea of the peak spacing for the specific application, but exact knowledge is not required. For this particular case, a vibration signal from bearing fault is used to improve itself, the obvious drawback being that the fault was already present and detectable. It is emphasized that a more useful case would be to use signals independent of the fault, e.g. gear meshing and motor piston vibration. However, the test setup used in this experiment did not have a gearbox, and the presented results serves more as a "proof of concept". Note that the peaks were visible in the time-waveform without pre-processing, and a pre-processing step might be necessary, such as bandpass filtering of high kurtosis bands (Randall and Antoni 2011).

2 Experimental Setup

A series of experiments using undamaged and artificially damaged bearings exposed to different radial loads and rotational speeds were performed by the authors at the facilities of RWTH Aachen as part of the research program SFI Offshore Mechatronics at the University of Agder and Teknova AS. The actual bearing used was a 180 mm roller element bearing, NU220 ECP, seen in Fig. 1. The artificial damage was a groove in the outer race made by a hand held rotating tool with an abrasive tip.

Fig. 1. NU200 ECP bearing used in testing, artificial damage seen in the outer race

The test rig applied pure radial load only. Test were done at 4 different load cases, as seen in Table 1. In this paper, results for load case 2 are presented.

Table 1. Radial loads

Load case	1	2	3	4
Load [kN]	5	10	15	20

An induction motor was driving the shaft directly without a gearbox. A rotary encoder supplied RPM data. Table 2 shows the motor nominal test speeds in RPM and Hz.

Table 2. Motor test speeds

Nominal speed	[RPM]	100	200	300	400	500	750	1000
Shaft frequency, f_{snom}	[Hz]	1.67	3.33	5	6.67	8.33	12.5	16.67

Vibration data was collected using a triaxial accelerometer, location shown in Fig. 2. All data was sampled at 19,200 Hz.

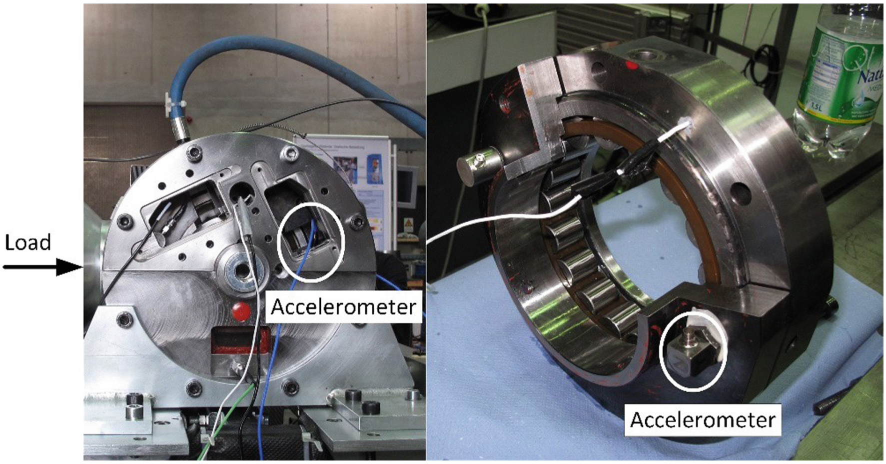


Fig. 2. Test rig (left) and bearing housing with accelerometer placement (right)

3 Method

Two channels of radial acceleration data, horizontal and vertical, were added to find the magnitude of radial acceleration. The resulting vector is taken as input to an enveloping algorithm. The artificial defect introduced to the outer race causes repetitive transients, and the local maxima of the combined envelope signal are detected. The time differences between the local maxima was used for estimating an instantaneous frequency inherent in the envelope signal. Knowing the nominal shaft frequency f_{snom} of the motor, a Peak-to-Shaft Ratio (PTSR) was calculated, describing the ratio between mean impact frequency and nominal shaft frequency. Any distinguishable, repetitive signal of higher frequency than the shaft can be used, for instance gear mesh or other bearing fault frequencies. In cases where the source of the transients is known, the observed PTSR can be used to correct mean value offsets of the nominal RPM.

The vibration signal contained transients known to originate from the outer race fault. These transients arrived an assumed constant ratio to the shaft frequency, neglecting the randomness in arrival time. To capture the speed fluctuations within one revolution, peaks of the signal envelope were detected, and the time between peaks was calculated. To locate the peaks, the Matlab function *findpeaks* is used. To minimize detection of false peaks, the parameter *MinPeakDistance* is used. Minimum peak distance limits the maximum number of peaks to be recorded per shaft revolution, effectively limiting the upper shaft frequency fluctuation that and can be captured. It also prevents one peak to be counted twice. For this case, the order normalised fault frequency is 7.174X, and a 30% fluctuation in speed was allowed. Additionally, the parameter *MinPeakProminence* was set to the RMS of the input signal. This removes small peaks, by setting a lower threshold of the peak prominence. Both these parameters can be tuned.

The times between the i^{th} and $(i-1)^{th}$ peak in the envelope signal is measured as the number n_i of time samples multiplied by the sample period T_s. We call this $T_{p,i}$, representing the time interval between the envelope maxima.

$$T_{p,i} = n_i \cdot T_s \tag{1}$$

Instantaneous peak frequency $f_{p,i}$ is then the inverse of $T_{p,i}$. The estimate is then shifted back by half an impact period to place it between the impacts it was calculated from. The mean value of all $f_{p,i}$ is added to each end of the vector.

$$f_{p,i} = \frac{1}{T_{p,i}} \tag{2}$$

Knowing the nominal shaft frequency, and mean time between impact, an estimate of the *PTSR* was obtained.

$$PTSR = \frac{1}{f_{snom}} \cdot \frac{1}{N} \sum_{i=1}^{N} f_{p,i} \tag{3}$$

If the source of the transient has a known frequency f_{source}, a correction factor k can be calculated remove a constant RPM offset.

$$k = \frac{PTSR}{f_{source}} \tag{4}$$

An estimate of f_s, the instantaneous shaft frequency estimate was then obtained by dividing the instantaneous impact frequency by the *PTSR*.

$$f_{s,i} = \frac{f_{p,i}}{PTSR} k \tag{5}$$

Expressed in revolutions per minute RPM_i

$$RPM_i = f_{s,i} \cdot 60 \tag{6}$$

Occasionally, peaks are ambiguous, not present or not distinguishable, which leads to erroneous instantaneous RPM estimates. Large deviations in time between peaks are detected and replaced using a two-stage outlier identification method. An outlier is defined as any point more than three scaled median absolute deviations (MAD), which provides more robust outlier detection in tail-heavy distributions (Leys et al., 2013). The Matlab function *filloutliers* provides this functionality. This stage removes large outliers and replace them with linear interpolations between nearest valid neighbours. Since each semi-periodic frequency fluctuation was represented by only a few data samples, a Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation function was applied to smooth the resulting plot after peak detection and conversion to instantaneous shaft frequency, so that both measured and estimated frequency can be

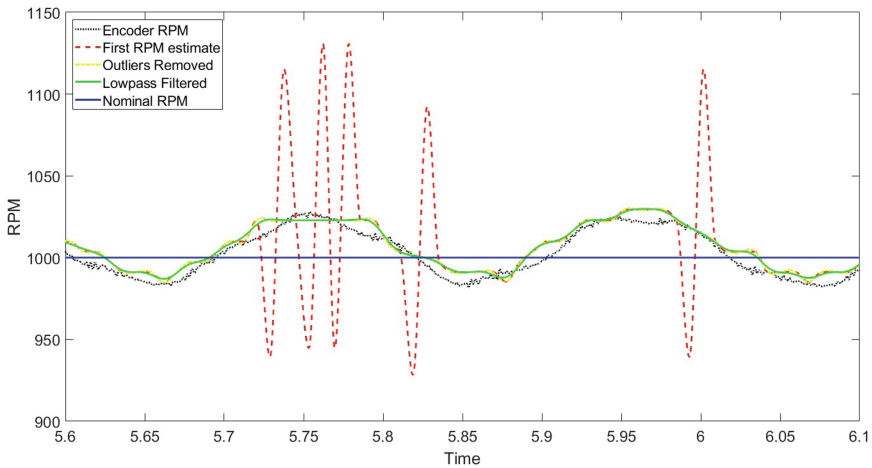


Fig. 3. Stages of RPM estimation at 1000 RPM

meaningfully plotted on the same scale for comparison. The limited number of impacts per revolution also impose a theoretical limit on the frequency of the fluctuations that can be captured. In the second stage, the estimate is lowpass-filtered using a 2^{10} order equiripple FIR filter with a cutoff frequency of $f_{snom}*PTSR/2.56$. The factor of 2.56 was used to provide some safety margin to the Nyquist frequency. Figure 3 illustrates the different stages of RPM estimation.

4 Results

The data sets analysed contained rotational speed fluctuations. The relative speed variation was highest at 100 RPM, at about ±15% at most. At higher speeds, the absolute variation increased, but not proportionally to nominal RPM. Figure 3 shows the encoder RPM at 100 and 1000 RPM in a 10 s window. A summary of the severity of the speed fluctuation is given in Table 3.

Table 3. Speed fluctuation

Nominal speed	[RPM]	100	200	300	400	500	750	1000
Actual speed	[RPM]	99.6	200.5	296.4	392.8	497.8	755.8	1005.1
RMS error	[%]	6.55	3.45	1.28	1.3	1.83	1.52	1.29

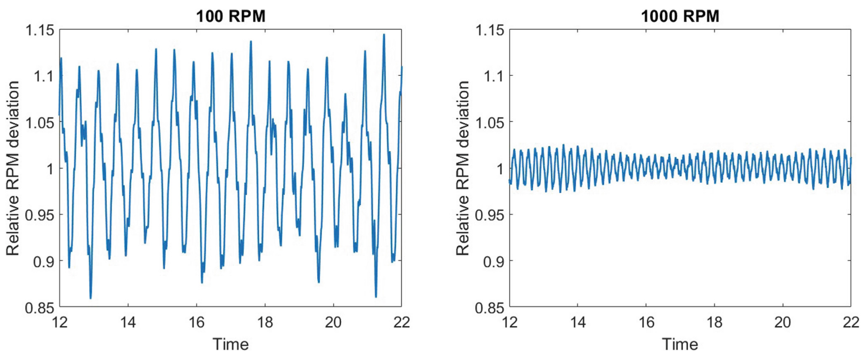


Fig. 4. Relative fluctuation in speed at 100 RPM and 1000 RPM

In Fig. 4, estimated instantaneous RPM is compared to measured RPM from the tachometer. There are both short term and long term variations present in this recording. The rotational speed fluctuations at 1000 RPM nominal show a temporal variation of approximately five periods per second, compared to two periods per second at 100 RPM nominal rotational speed. The estimated RPM is able to capture both variations. The agreement between estimated and measured RPM is good, but the

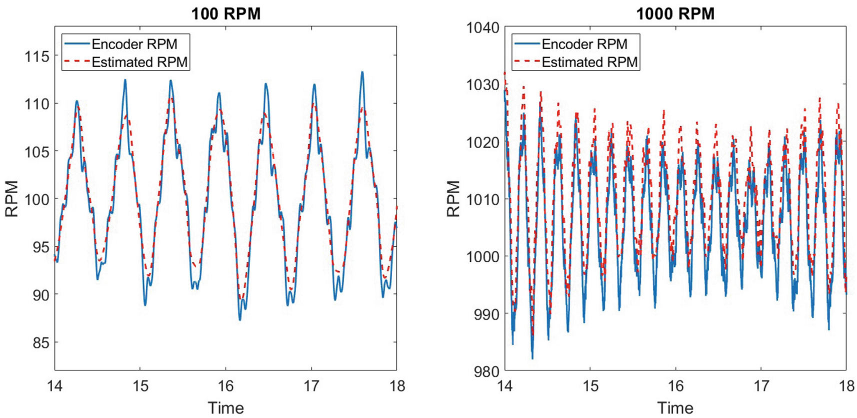


Fig. 5. Measured and estimated RPM at 100 and 1000 RPM

highest peaks are not captured, and there is still an offset in the 1000 RPM recording even after the correction factor is applied (Fig. 5).

Table 4. Performance of the order tracking method at load case 2

RPM	Relative RMS error [%]	Est. RPM FWHM reduction [%]	Enc. RPM FWHM reduction [%]	Est. RPM peak increase [%]	Enc. RPM peak increase [%]
100	1.65	24.15	23.33	55.29	54.17
200	0.64	13.44	7.03	11.57	8.23
300	0.52	39.10	40.16	45.79	45.54
400	0.58	67.35	71.70	143.76	157.14
500	0.53	40.81	42.23	87.59	59.40
750	0.53	46.86	47.86	74.81	75.04
1000	0.88	68.75	70.91	70.54	74.16

4.1 Performance

Performance is evaluated from relative RMS error of the RPM estimate, relative Full Width Half Maximum (FWHM) reduction and relative increase in peak value. Order tracking using both the encoder RPM end the estimated RPM, resulted in a visible sharpening of square envelope spectrum (SES), seen in Fig. 4. The sharpness is measured by relative reduction in FWHM at the highest peak. An overview of the results for load case 2 is given in Table 4 (Fig. 6).

In most cases, the estimated RPM provides performance comparable to the encoder data. In a few cases, it actually seems to perform equally good or better. However, as the frequency resolution is high, small differences in energy distribution between bins will affect both FWHM and peak height, and this is assumed to be the cause.

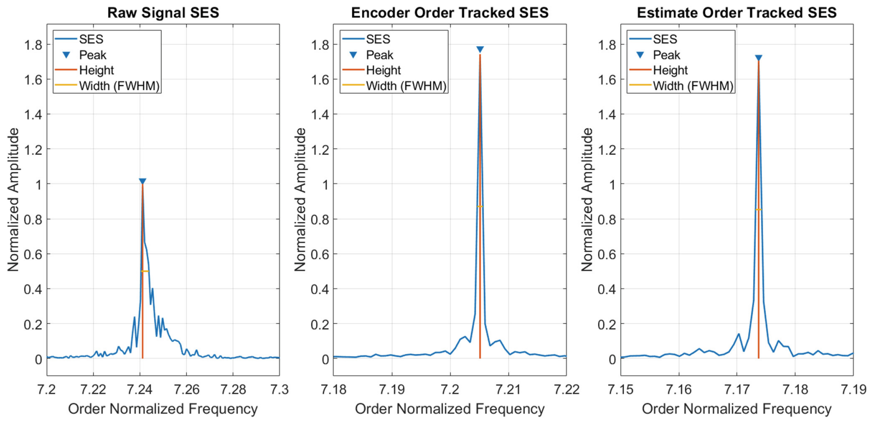


Fig. 6. Squared Envelope Spectrum of the vibration signal, centred around the largest peaks

5 Conclusion

In this paper it is demonstrated that rotational speed variations can be estimated from processing of the vibration signal itself, although with reduced resolution both in amplitude and temporally compared to an encoder. However, for situations where a tachometer is not present and is too costly or time consuming to install, the estimates obtained from accelerometer signals can prove to be useful. The rotational speed estimated can then be used for order tracking of the vibration signal, reducing spectral smearing. The method is suitable for applications where a tachometer or encoder signal is missing but a periodic transient is present.

Acknowledgement. The research presented in this paper has received funding from the Norwegian Research Council, SFI Offshore Mechatronics, project number 237896.

References

André, H., et al.: Precision of the IAS monitoring system based on the elapsed time method in the spectral domain. Mech. Syst. Sig. Process. **44**(1–2), 14–30 (2013). https://doi.org/10.1016/j.ymssp.2013.06.020

Bonnardot, F., et al.: Use of the acceleration signal of a gearbox in order to perform angular resampling (with limited speed fluctuation). Mech. Syst. Sig. Process. **19**, 766–785 (2004). https://doi.org/10.1016/j.ymssp.2004.05.001

Leys, C., et al.: Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J. Exp. Soc. Psychol. **49**, 764–766 (2013). https://doi.org/10.1016/j.jesp.2013.03.013

Randall, R.B., Antoni, J.: Rolling element bearing diagnostics—a tutorial. Mech. Syst. Sig. Process. **25**(2), 485–520 (2011). https://doi.org/10.1016/j.ymssp.2010.07.017

Schmidt, S., Heyns, P.S., De Villiers, J.P.: A tacholess order tracking methodology based on a probabilistic approach to incorporate angular acceleration information into the maxima tracking process. Mech. Syst. Sig. Process. **100**, 630–646 (2018). https://doi.org/10.1016/j.ymssp.2017.07.053

Wang, Y., et al.: An online tacholess order tracking technique based on generalized demodulation for rolling bearing fault detection. J. Sound Vib. **367**, 233–249 (2016). https://doi.org/10.1016/j.jsv.2015.12.041

Zhao, M., et al.: Tacholess envelope order analysis and its application to fault detection of rolling element bearings with varying speeds. Sensors **13**, 10856–10875 (2013). https://doi.org/10.3390/s130810856

Acoustic Response of Roller Bearings Under Critical Operating Conditions

Daniel Cornel(✉), Francisco Gutiérrez Guzmán, Georg Jacobs, and Stephan Neumann

Chair for Wind Power Drives (CWD), RWTH Aachen University, Aachen, Germany
daniel.cornel@cwd.rwth-aachen.de

Abstract. Gearbox roller bearing failures in wind turbines lead to long downtimes and high repairing costs. In this paper an acoustic emission based condition monitoring system is applied to a roller bearing test bench with the aim of identifying critical operating conditions, which could lead to early failures. The investigations show a sensitivity of the system towards lubricating conditions, pressure and slip. By applying multiple uni- and multivariate collectives it is shown that the resulting acoustical signals have characteristic attributes, allowing therefore the identification of the critical parameters.

1 Introduction

In the last decades the global wind power capacity has increased quickly, with an estimated cumulative global capacity of 539 GW (GWEC 2017). Besides the continuous growth efforts have been made to optimize the reliability and the operation and maintenance (O&M) strategy of a wind farm. Nowadays, the life time of a wind turbine (WT) is expected to reach 20 years. However, most of the failures occur within the first 6 years (Fraunhofer 2010). It has been shown that plant availability is mostly limited by the gearbox (Faulstich et al. 2010), whose reliability is mostly limited due to bearing failures (Sheng 2015). The reliable design of roller bearings for WT gearboxes has not been successful mostly because the critical operating conditions leading to premature bearing damage remain mostly unknown. Modern tests in laboratory-scale have not been able to fully correlate critical operating conditions with operating conditions in the field, which lead to premature bearing failures. A different approach can be realised by determining realistic operating states on system test benches, such as the 4 MW nacelle test bench at the Center for Wind Power Drives, RWTH Aachen University. However, system test benches are not suitable for bearing life tests, mostly due to high operating costs and long runtimes.

Based on these limitations, a new approach is followed in the project WT-Bearing Center.NRW. In this project WT roller bearings will be tested under realistic conditions. This will reduce the costs compared to tests at a system level and improve the transferability. The main objective within this project is to develop a test method, including the necessary test benches and test procedures (Kock et al. 2018), for the qualification and approval of roller bearings for WT (Neumann et al. 2017). The test

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 740–749, 2020.
https://doi.org/10.1007/978-3-030-48021-9_82

benches will be equipped with an acoustic emission system (AES) in order to detect critical operating conditions and bearing failures such as cracks, as early as possible. Prior to implementing the AES in the newly build test benches, a smaller version of the high speed shaft bearing test bench is used to determine a correlation between signal waveform, operating conditions and the sensitivity of the AES towards parameter variation. Within this work, it was shown that the operating conditions which could lead to bearing failures such as accelerated fatigue and smearing can be detected by the AES by detecting high-frequency acoustic emissions.

2 Roller Bearing Failures Modes and Their Influencing Factors

Rolling Contact Fatigue (RCF): A typical failure which can appear through subsurface originated spalling and/or surface originated pitting. Extensive experimental studies have been conducted to investigate the influences of various potential factors such as surface quality, lubricant cleanliness and material quality. Under ideal conditions such as correct mounting, alignment, and lubrication it can be expected that this failure mode occurs towards the end of life (Harris 2007). Unfavourable conditions such as shock loads can, however, accelerate this process.

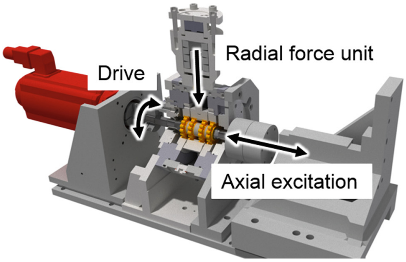
Micro-pitting: A localized surface fatigue damage caused primarily by microcracks growing against the sliding direction (Erichello 2012). The process of micro-pitting progressively alters the geometries of the contact surfaces and, therefore, the stress distribution, leading ultimately to spalling. According to (Bongardt et al. 2016) micro-pitting is caused by the interaction of the contact bodies asperities. It is now widely understood that this usually occurs when the lubricant film thickness is insufficiently to separate the raceways during sliding. However, it has been discussed that besides the contact conditions certain additives can favour the formation of micro-pitting (Brechot et al. 2000).

Smearing, scuffing and rehardening: Different types of surface damages that can occur if excessive sliding between the contacting surfaces is present. In contrast to fatigue induced failure there is no incubation phase, so that the bearing failure occurs suddenly. In the case of smearing and scuffing material is transferred from one surface to another and both are preceded by the failure of the protective lubricant films (Fowell et al. 2014). On the other hand during rehardening the material is locally heated to such temperatures that microstructural changes takes place. This produces a local change of the hardness which lead to localized stress concentrations that may cause flaking (van Lier 2015).

White Etching Cracks (WEC): A still unresolved mechanism associated with the premature failure of rolling bearings at 5–20% of the nominal life due to axial cracking or spalling of the bearing's raceway (Evans 2016). The damage pattern is characterized by sub-surface crack networks within regions of altered microstructure, which are resistant to etching and are called white etching areas (WEA). Although this mechanism had been well investigated over the course of the last years, the relevant drivers and formation mechanisms are still under discussion. Some authors propose that the crack initiation and propagations is a consequence of the formation of WEA, while other authors have suggested that the cracks are the precursor of the WEA. Besides sliding and lubricating conditions (Gutiérrez Guzmán 2010), other WEC influence factors such as a local hydrogen ingress, tensile stresses, lubricant composition and electrical effects had been proposed.

3 Test Bench and Damage Detection

Four-bearing test bench: The investigations were carried out on a four-bearing test bench (Fig. 1). This test bench consist mainly of a drive unit, a radial load unit, an oil chamber and four test bearings from type NU206-TVP2 with a bore diameter of 30 mm. The radial load unit consists of a disc spring stack, which can be statically pre-loaded. The radial load is applied through the two middle test bearings. The test bearings are mounted on a shaft, which is driven by an electric motor (max. 6,000 rpm). The loaded shaft itself is mounted in the test head by the outer test bearings, which are each arranged symmetrically to each other ensuring a symmetrical load distribution on the four test bearings.



Test bench	RLP
Shaft diameter [mm]	30
No. of test bearings [-]	4
Bearing type	NU206
Radial load [kN]	100
Axial load [kN]	8/100 (high frequency/static)
$p_{Hertz,\,max}$ [MPa]	3,600
Test temperature [°C]	40-110

Fig. 1. Four-bearing test bench (RLP), shaft diameter 30 mm

Damage detection: Beside classic acceleration or vibration sensors, several acoustic emission sensors were used. Although both methods rely on the same operating principle, the acoustic emission sensors have no seismic mass. As a result, AES can be used in higher frequency bands (Cockerill 2017), often with a resonance range of 100 kHz to 1,000 kHz. On the other hand, classical acceleration sensors operate in a band between 20 and 50 kHz. Therefore, AES operates outside the frequency range of classic machine vibrations and is less sensitive for typical back ground noises such as test bench cooling system (PAC 2015). To increase the accuracy, the sensors need to be installed as close as possible to the contact surfaces. In the case of the middle test bearings (2 & 3) the sensors can be easily installed within the test head, shown in Fig. 2. In the case of the outer test bearings (1 & 4) the sensors need to be installed outside the housing due to lack of space.

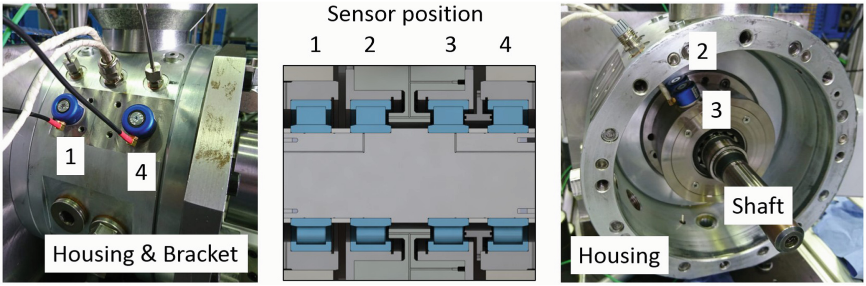


Fig. 2. RLP test bench: position of the installed AE sensors

4 Experimental Method and Results

The aim of this study is to correlate the behaviour of the AES with the operating conditions. In the first step, a four-bearing test bench was used. The critical conditions implemented in the test bench are suspected of favouring one or multiple damage mechanisms introduced in Sect. 1. The testing points were run for a defined time (t_{test} < 2 min.) so that a bearing failure can be excluded.

In this work, the specific film thickness lambda (λ) is used as a quantitative indicator of the lubricating conditions in the point of contact. The λ-value, is determined using the lubrication film thickness h_{min} – according to DOWSON and HIGGINSON (Dowson 1977) – and the root-mean-squared surface roughness of the contact partners $R_{a,1,2}$. According to (Czichos et al. 2015), values of $\lambda \geq 3$ indicate full fluid lubrication conditions and values of $\lambda \leq 1$ indicate boundary lubrication conditions. The arithmetic mean surface roughness R_a of the bearing inner rings and rolling elements was measured using a tactile roughness measuring device on three different measuring points per body. The R_a-values lie in the range from 30 to 50 nm for the inner rings and from 25 to 46 nm for the rollers. As already mentioned, the applied load collectives contain operating points which are intended to favor an early bearing failure. In this study it is assumed that in the case of subsurface RCF full fluid lubrication, i.e. a $\lambda > 3$ is required (Czichos et al. 2015), whereas micro-pittings (Bongard et al. 2016) and WEA/WEC (Evans 2016) could be promoted by boundary lubrication. These aforementioned damages are also promoted by high loads. Among other things, high accelerations (>314 rad/s^2) and high dynamic operating conditions are conducive to new hardening and smearing (FVA 2018). These boldly summarized conditions are used in the following test series.

Shaft speed & Load: In the first step two univariate collectives were applied. At first the shaft speed was varied in the range from 250 to 3,000 rpm under constant bearing′s outer ring temperature (t = 100 °C) and contact pressure at the inner ring (IR) ($p_{Hertz,IR}$ = 2.0 GPa). By varying the speed, both the over-rolling frequency and the specific lubricating film are varied. The specific lubricating film thickness and therefore the separation of the contact surfaces decreases with decreasing shaft speed. On the next step the load and consequentially the resulting pressure was varied in the range from 1 kN to 22 kN ($p_{Hertz\text{-}IR}$ = 0.9 – 3.0 GPa). It is well known that a higher contact pressure affects mostly the fatigue lifetime (Harris 2007). The results of both test series are shown in Fig. 3. While the y-axis shows the qualitative energy value of the acoustic signal (designated as amplitude), the x-axis shows respectively the shaft speed and the contact pressure. For all measuring points in Fig. 3 it can be seen that the amplitude generally rises with increasing speed and with increasing load. This acoustic response is against the expectation, that by increasing the speed and decreasing the load and, therefore, increasing the lubricant film height the amplitude should decrease. The observed tendency cannot be explained by the increased cycling frequency caused by the increased speed, due to the fact that the direct excitation lies in the range of 1 kHz and the recorded frequency lies between 100 kHz – 1000 kHz. It could also be conceivable that at certain speeds additional effects could occur, e.g. lubricant flow effects or cage excitation and affect the measurements. Furthermore, the middle sensors 2 & 3 show a much stronger sensitivity than the outer sensors 1 & 4. This effect can be traced back to the closer proximity of sensors 2 & 3 to the test bearings. Moreover, a comparison of the gradients shows that the amplitude is influenced much less by the load. This is presumably due to the fact that an increased speed provides for a stronger excitation of the system resulting in a higher amplitude compared to one caused by an increased load.

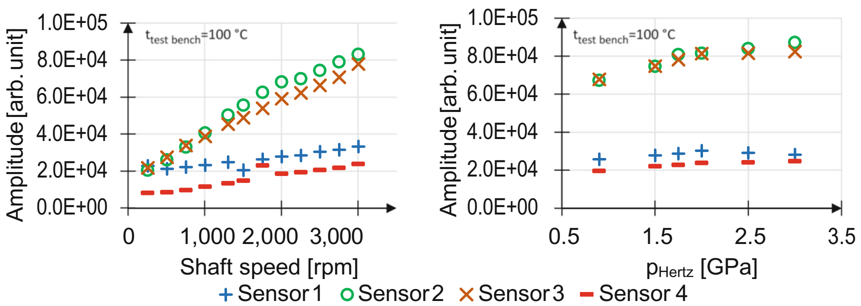


Fig. 3. Influence of the shaft speed and the Hertzian pressure on the AE Energy level

<u>Lambda:</u> The strong dependence of the specific lubricant film thickness on the oil temperature was used to adjust different lubrication conditions. More precisely, the temperature was varied from 30 to 130 °C in 10 °C steps, whit a shaft speed of 500 and 3,000 rpm. Within this test series, λ-values between 4.1–0.3 (500 rpm) and 14.5–1.2 were achieved. The result of the investigations carried out at a constant pressure of 2.0 GPa are shown in Fig. 4. It can be seen that the amplitude decreases as λ increases. This is especially noticeable in the areas around λ = 0.5–2. At 500 rpm and λ > 2 the amplitude remains almost constant. On the other hand, at 3,000 rpm this effect can only be seen by the external sensors at λ > 5. As explained before a lower λ value indicates an increased contact of the asperities which could lead to an increase of the general noise. This fact correlates with the observation from the experiment. As observed in the previous experiment, increasing the shaft speed leads to an increased amplitude.

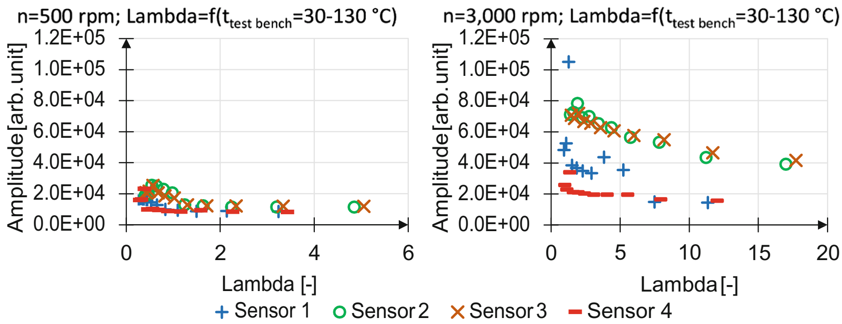


Fig. 4. Influence of the lubrication regime and the shaft speed on the AE Energy level

Angular acceleration: Within these tests speed ramps (0–2,500 rpm) were applied within 0.025–5 s. This speed ramps result in angular accelerations from 52 to 10,480 rad/s^2. The results from the tests carried out at a temperature of 100 °C are shown in Fig. 5. Due to the strongly varied value range and in contrast to the previous figures the y-axis shows the amplitude gradient and both axis are scaled logarithmic. The angular acceleration can be correlated to the amplitude gradient. Thus, occurring critical accelerations could be detected without specific bearing measurement technology.

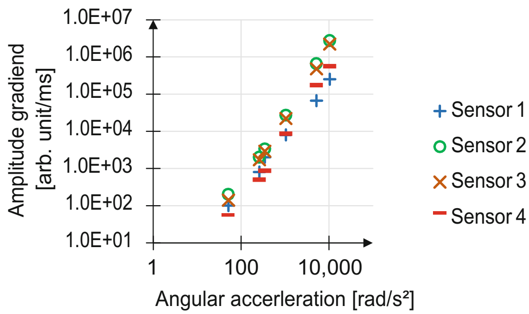


Fig. 5. Influence of rotational acceleration on the AE Energy level

Axial slip: The signal behaviour under axial excitation was investigated using the following conditions: axial amplitudes of ±0.15 and ±0.3 mm at shaft speeds of 500 and 3,000 rpm, constant pressure of 2.0 GPa and constant temperature of 80 °C. The figures below (Fig. 6 and Fig. 7) show that the excitation frequency for an amplitude of 0.15 mm plays a secondary role compared with the influence of the shaft speed. On the other hand, with an amplitude of 0.3 mm it can be seen that the excitation frequency clearly determines the amplitude of the AE sensors at both shaft speeds. It is interesting that the results of the respective sensors are at a comparable level (apart from sensor 1, which shows clear deviations). The axial excitation leads to axial slippage conditions, which cannot be determined in detail, but possibly lead to a reduction of the λ-value and thus increase the general noise.

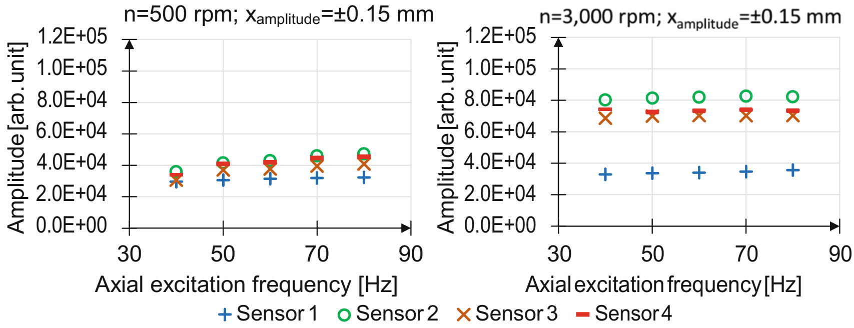


Fig. 6. Influence of the axial excitation frequency on the AE Energy level for an amplitude of ± 0.15 mm and different shaft speeds

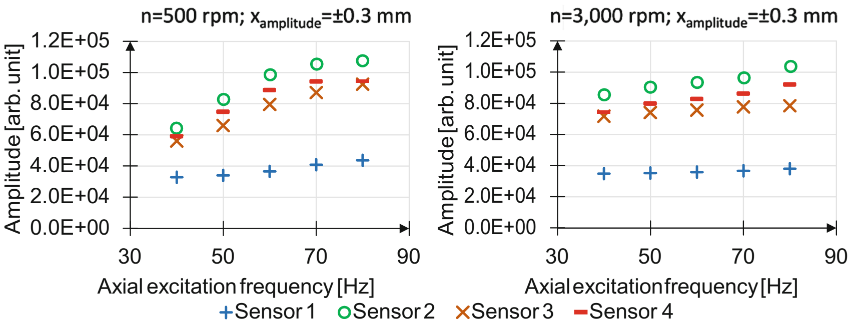


Fig. 7. Influence of the axial excitation frequency on the AE Energy level for an amplitude of ± 0.3 mm and different shaft speeds

5 Summary and Outlook

The investigations have shown that the lubricating conditions of the system have a clear influence on the signals of the AES. Figure 8 compares the results. The previously used amplitude is plotted above the specific lubricating film height lambda. This representation allows a direct assignment to the lubrication condition and thus an initial assessment of the criticality of the conditions. The figures shows the reduction of the load, yellow arrow, at a speed of 3,000 rpm and the reduction of the temperature also at 3,000 rpm, blue arrow. In addition, the temperature variation at 500 rpm is shown by the red arrow and the variation of the speed is kept in green. Figure 8 does not include the findings of axial excitation, as the authors are not aware of any possibility to include the influence of axial excitation in the lubricant film height calculation.

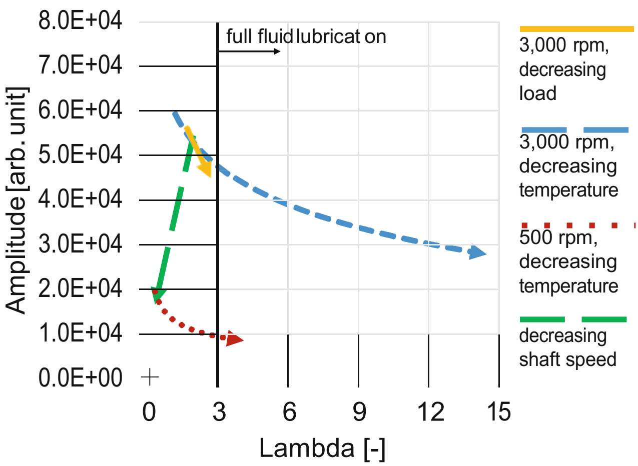


Fig. 8. Influence of load, temperature and shaft speed on the specific lubricating film height and on the amplitude of the AE system

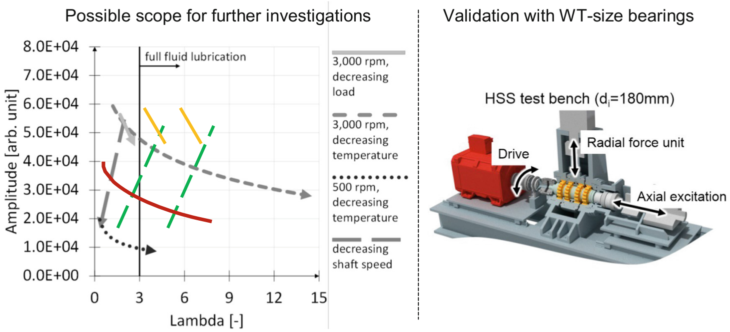


Fig. 9. Further steps: Validation of the results with real-sized WT-gearbox bearings on the HSS test bench

As expected a decrease of lambda-value decreases leads to an increase of the measured amplitude. An explanation for this tendency could be provide by λ itself due to the fact, that as λ increases, the solid state contact (λ < 3) decreases and thus the possible contact between the roughness peaks of the bearing components. A possible explanation for the continuous amplitude decrease at λ > 3 – by reducing the temperature – could be increased damping properties of a higher lubricating film. A more detailed examination of the speed influence shows that, as expected, an increase in the speed leads to a significant increase of the amplitude. This effect can probably be explained by the generally increased system excitation at a higher speed. The individual changes of the amplitude can be assigned to individual parameters (e.g. load, speed, temperature changes). Furthermore, the different test series share some identical test points, which resulted in comparable amplitudes, indicating a good reproducibility of the results. Based on these results it can be concluded that critical operating points can be detected using a AES. However, the results of the variation of lambda via the shaft speed shows that further effects such as lubricant flow effects or cage excitation should be considered.

In following studies the examined parameter field will be extendend with intermediate stages in order to detect possible non-linearities (Fig. 9). Furthermore, the results are going to be validated on WT-size bearings. For this purpose the four bearing test bench, mentioned in the beginning (shaft diameter d = 180 mm), wil be used.

Acknowledgements. The authors would like to thank the European Union and the German federal state of North-Rhine Westphalia (NRW) for the financial funding. They would also like to thank the *Leitmarktagentur.NRW* for the academic and administrative project support and all industrial partners for the good teamwork.

EUROPÄISCHE UNION
Investition in unsere Zukunft
Europäischer Fonds
für regionale Entwicklung

EFRE.NRW
Investitionen in Wachstum
und Beschäftigung

References

Bongardt, C.: Wälzlagergraufleckigkeit. Ph.D. thesis, Aachen University (2016). ISBN: 978-3-95886-089-6

Brechot, P., Cardis, A.B., Murphy, W.R., Theissen, J.: Micropitting resistant industrial gear oils with balanced performance. Ind. Lubr. Tribol. **52**(3), 125–136 (2000). https://doi.org/10.1108/00368790010371762

Cockerill, A.: Damage Detection of Rotating Machinery. Ph.D. thesis, Cardiff University (2017). orca.cf.ac.uk/id/eprint/105671

Czichos, H.; Habig, K.H.: Tribologie-Handbuch, Springer Vieweg, Wies- baden (2015). https://doi.org/10.1007/978-3-8348-2236-9

Dowson, D., Higginson, G.R.: Elasto Hydrodynamic Lubrication. Pergamon Press, Oxford (1977). https://doi.org/10.1016/C2013-0-05764-7

Evans, M.H.: An updated review: white etching cracks (WECs) and axial cracks in wind turbine gearbox bearings. Mater. Sci. Technol. **32**, 1133–1169 (2016). https://doi.org/10.1080/02670836.2015.1133022

Faulstich, S., Hahn, B., Tavner, P.: Wind Turbine Downtime and Its Importance for Offshore Deployment. Wiley, New York (2010)

Fowell, M., Ioannides, S., Kadiric, A.: An experimental investigation into the onset of smearing damage in non-conformal contacts with application to roller bearings. Tribol. Trans. **57**, 472–488 (2014). https://doi.org/10.1080/10402004.2013.875607

Fraunhofer IWES: Erhöhung der Verfügbarkeit von Windkraftanlagen (2010). https://doi.org/10.2314/gbv:634597957

FVA, Forschungsvereinigung Antriebstechnik. Schädlicher Wälzlager- schlupf II, FVA 663 II, vorläufiger Abschlussbericht (2018)

Gutierrez Guzman, F., Oezel, M., Jacobs, G.: Investigation of the formation and morphology of White Etching Cracks using a two-disc test rig. In: Conference Transcript, TAE 21st International Colloquium Tribology, Stuttgart (2018)

GWEC: Global Wind Energy Council. Global Wind Report – Annual Market Update, Brussels (2017)

Harris, T.A., Kotzalas, M.N.: ROLLING BEARING Analysis. CRC/Taylor & Francis, Boca Raton (2007). ISBN 978-0-84937-182-0

Kock, S.; Jacobs, G.; Hirt, A.: Oberdörfer, S.; Neumann, S.; Bosse, D.: Robustness test for wind turbine gearbox bearings. J. Phys. Conf. Ser. (2018). https://doi.org/10.1088/1742-6596/1037/5/052012

Li, S., et al.: Microstructural evolution in bearing steel under rolling con- tact fatigue. Wear **380–381**, 146–153 (2017). https://doi.org/10.1016/j.wear.2017.03.018

Neumann, S.; Jacobs, G.: Reduzierung von Lagerschäden im Antriebs- strang von Windenergieanlagen – WEA-Lagerzentrum.NRW, Ingenieur-Spiegel (2017). https://doi.org/10.18154/rwth-2017-09861

Physical Acoustics Corporation. Acoustic Emission Measurement training materials (2015)

Scott, K., Infield, D., Barltrop, N., Coultate, J., Shahaj, A.: Effects of extreme and transient loads on wind turbine drive trains. In: 2012 50th AIAA Aerospace Sciences Meeting, Nashville (2012). https://doi.org/10.2514/6.2012-1293

Sheng, S.: Wind Turbine Gearbox Reliability Database, Condition Monitoring, and Operation and Maintenance Research NREL, GRC Annual Meeting (2015)

van Lier, H.: Neuhärtungsgefährdung von Radial-Zylinderrollenlagern durch Lastaufschaltungen in Betriebspunkten mit Käfigschlupf. Ph.D. thesis, Aachen University (2015). ISBN 978-3-95886-052-0

Rib-Roller Wear in Tapered Rolling Element Bearings: Analysis and Development of Test Rig for Condition Monitoring

Martin Hemmer[1,2(✉)], Kjell G. Robbersmyr[1], Tor I. Waag[3], Rolf Albrigtsen[4], Torfinn Pedersen[4], Thomas J. J. Meyer[3], and Chloë Vercammen[5]

[1] University of Agder, Jon Lilletuns Vei 9, 4879 Grimstad, Norway
{martin.hemmer,kjell.g.robbersmyr}@uia.no
[2] MHWirth, Butangen 21, 4639 Kristiansand, Norway
[3] NORCE Norwegian Research Centre AS, Jon Lilletuns vei 9, 4879 Grimstad, Norway
{towa,thme}@norceresearch.no
[4] MacGregor Norway AS, Kystveien 18, 4841 Arendal, Norway
{rolf.albrigtsen,torfinn.pedersen}@macgregor.com
[5] ArianeGroup, Rue de Touban, 33185 Le Haillan, France
chloe.vercammen@hotmail.fr

Abstract. Rolling Element Bearings (REBs) are present in virtually all machines with moving or rotating parts, and are vital for proper performance and safe operation. Condition Monitoring (CM) of bearings often receive particular interest, as this component group rarely reach design lifetime and hence is responsible for unplanned machine downtime. Unplanned maintenance can represent a large cost which motivates development of improved CM methods for implementation of advanced maintenance regimes. Based on observations of a used bearing from an offshore drilling machine, wear on roller ends in the rib-roller contact area was identified as an area of interest for future research. A test rig for creating and observing accelerated roller end damage is developed, intended for use with vibration and Acoustic Emission (AE) sensors. In addition to normal continuous rotation of the bearing, the test rig is also designed with performing oscillation motion tests in mind. This mode of operation is of interest to manufacturers and end users of cranes and winches with heave compensation. Plans and challenges for future work are also discussed, in conjunction with the experimental setup.

1 Introduction

Dynamic mechanical systems of all types depend on bearings; a basic machine element that both allows and constrains movement by transferring loads to support structure while minimizing friction for the desired rotation or translation. As the diversity of applications is great, so is the number of bearing types and sizes. Given the widespread use and importance, bearing failure is also a large contributor to reduced asset availability. Under nominal operating conditions, fatigue will be the main failure mode of bearings. However, degradation and contamination of the lubricant can create wear that

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 750–757, 2020.
https://doi.org/10.1007/978-3-030-48021-9_83

changes the internal geometry of the bearing (Fitzsimmons and Clevenger 1977), leading to premature failure and unplanned maintenance.

This paper focuses on roller end wear in tapered REBs under axial load, rooted in a previous case study of an offshore drilling machine. The machine was taken out of operation for maintenance, and available for testing using both new and worn bearings. A schematic drawing of an axial spherical tapered roller bearing is seen in Fig. 1. Because of the tapered roller design, a rib is necessary to keep the rollers in place, applying a seating force to the roller end. The rib-roller contact area is separated by an oil film, and the mating surfaces are sliding relative to each other. A thorough dynamic analysis of tapered REBs is given by (Crep et al. 1995).

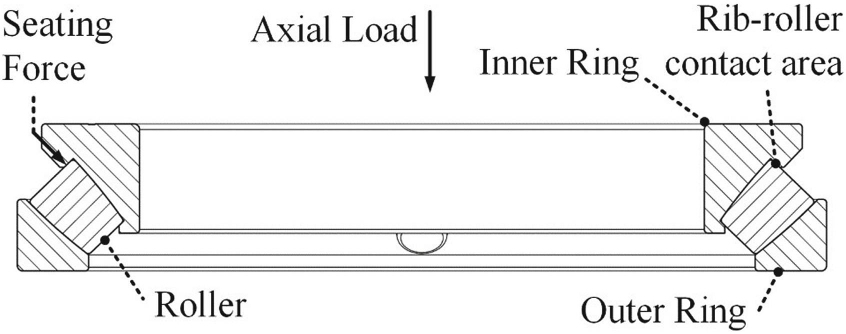


Fig. 1. Schematic drawing of an axial spherical tapered roller bearing

Tapered REBs are sensitive to proper lubrication of the rib-roller contact area, and lubrication starvation is a root cause of failure for bearing wear. In extreme cases, wear can lead to bearing seizure and catastrophic failure (Parker 1983). Ideally, online lubricant monitoring systems could detect wear particles at an early stage and prevent breakdowns. While such systems exist (Jae and Yoon 2014; Zhu et al. 2017) this approach has not replaced periodic sampling and analysis of the lubricant. Due to the inherent lag in this process, damage can occur and progress unnoticed. Alternative methods for online detection of wear will reduce the interval between detectable wear and maintenance action. Research of such methods is the main motivation for the development of this test rig.

Vibration monitoring using accelerometers is the current industry practice, and a range of well-established methods exists for fault detection (Randall and Antoni 2011) and health assessment (Wang et al. 2018). However, these methods are more suitable for detection of repetitive transients from localized damage. Roller end wear differs from localized faults in the load path, as scratch formation is not assumed to be related to the shaft frequency. This indicates that CM methods based on detection of repetitive transients at bearing defect frequencies will be ineffective.

It is known that AE has potential for monitoring of wear and sliding friction (Jiaa and Dornfeld 1990). Investigation of AE activity in sliding friction has shown a proportional relation between sliding speed and amplitude, and frictional work and energy.

Additionally, the frequency components associated with sliding was not present in the noise, which could simplify wear detection (Taura and Nakayama 2018). In the early stages of degradation, formation of scratches should therefore result in acoustic activity. AE has also been used for detection of wear in metal cutting tools (Bhuiyan et al. 2014). As wear progresses, local changes in internal geometry and clearance can lead to increased mechanical vibration detectable by accelerometers.

2 Roller End Wear

In a previous study by the authors (Hemmer and Waag 2017), one new and one previously used main bearing from an offshore drilling machine was run in a controlled environment. Scoring, a type of abrasive wear, was observed on all roller ends during visual inspection of the bearing, as seen in Fig. 2. However, any formation of new scratches could be not detected in the collected data. It is assumed that each scratch originates from a particle in the lubricant which enters the rib-roller contact area, then sliding across the roller surface, stationary with respect to the rib. If the particle size is greater than the oil film thickness, mechanical contact occurs, and a scratch is formed. The abrasion then contributes to further contamination of the oil. As this process is not expected to be periodic, any detection method based on detecting the presence of bearing fault frequencies will not be effective. Instead, an event-based approach is assumed to be more effective. If one can successfully detect the formation of new scratches, both the total number of events and changes in mean time between events are potential measures of bearing health and fault propagation. Tapered REBs are also sensitive to scuffing, a form of adhesive wear. As scuffing can be caused by excessive surface roughness, it is a possibility that the roller end scoring develops to scuffing.

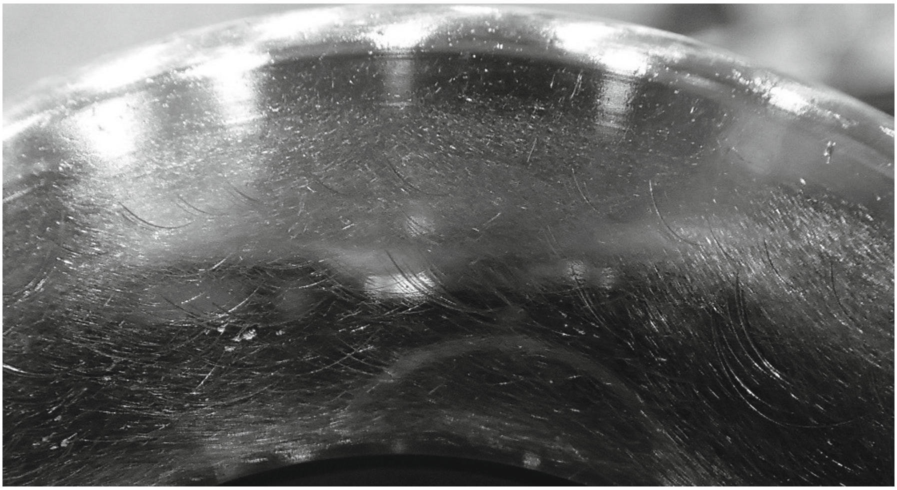


Fig. 2. Scoring on roller end of tapered roller bearing

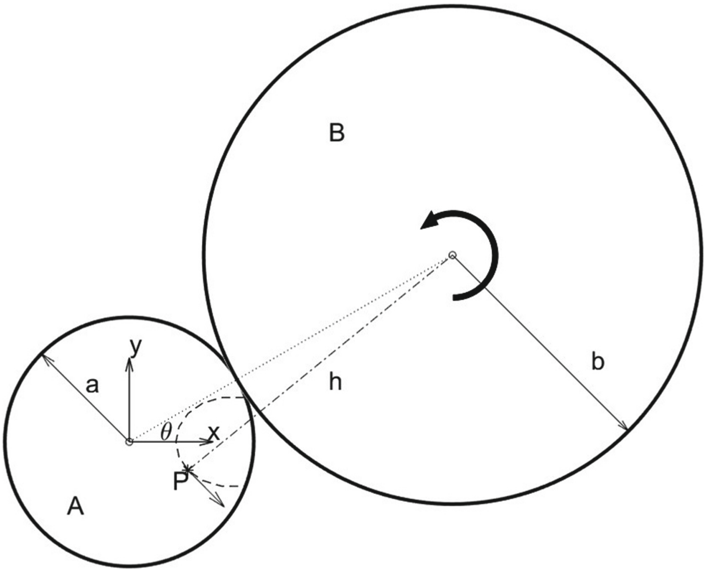


Fig. 3. Trajectory of P on A forms an epitrochoid.

2.1 Roller End Scoring Model

The scoring occurs as distinctly shaped arcs with varying radius. A simplified 2D visualization is given in Fig. 2. Let *A* be a circle of radius *a*, representing a roller end with a reference coordinate system centred on the roller end, while circle *B* of radius *b* represents a path of rolling contact with the bearing race. Then, consider B rolling on the circumference of A with a particle in point *P*, fixed at a distance *h* from the centre of *B*. The parameter θ is defined as the angle between the x-axis and the line between circle centres.

Let γ represent the following relation:

$$\gamma = \frac{a+b}{b} \tag{1}$$

The position vector of *P* as a function of θ is given in Eq. 2:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\cos\gamma\theta \\ \sin\theta & -\sin\gamma\theta \end{bmatrix} \begin{bmatrix} b\gamma \\ h \end{bmatrix} \tag{2}$$

If θ is a function of time, the instantaneous velocity vector of P, indicated by an arrow, is given in Eq. 3:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\sin\theta & \sin\gamma\theta \\ \cos\theta & -\cos\gamma\theta \end{bmatrix} \begin{bmatrix} b\gamma\dot{\theta} \\ h\gamma\dot{\theta} \end{bmatrix} \tag{3}$$

Note that the sliding velocity is a harmonic function of θ, resulting in variable sliding velocity for a constant RPM. In a similar application, monitoring piston seals in a combustion engine, AE activity showed a linear relationship to sliding velocity (Douglas et al. 2006). As seen by this simplified model, the same behaviour could be utilized in establishing a fault signature for roller end wear detection in bearings. With a known fault signature, it is possible to determine the presence of a fault in recorded data through hypothesis testing, providing decision support for scheduling of maintenance.

3 Test Rig

To the authors' knowledge, no test equipment for emulating roller end damage in a controlled environment during operation exists, motivating the development. Figure 4 shows a sketch of the test rig (left) and the bearing test unit (right).

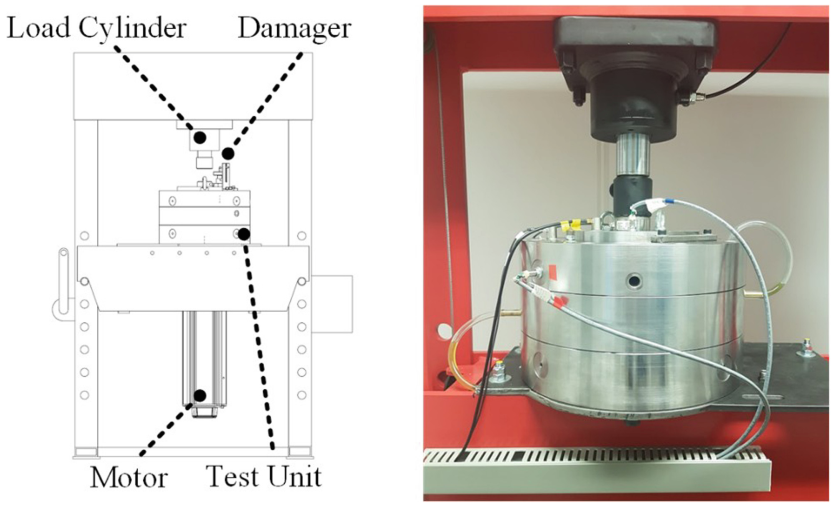


Fig. 4. Bearing test rig sketch (left), Bearing test unit (right)

The test bearing (designation 29230) is contained inside the bearing test unit, attached to the motor through a flexible coupling. Axial load is applied by a hydraulic load cylinder. A second, higher capacity bearing (designation 29336) transfers the load to the support structure, removing any axial load from the motor shaft. The test bearing is attached using an adapter, allowing for easy replacement and testing of other bearing types and sizes, up to 240 mm outer diameter. The load bearing and test bearing runs in two separate chambers, allowing different lubricants and preventing cross-contamination. The lubricant can also be circulated for monitoring purposes. The test rig is developed with two main test types in mind: artificial roller end damage and oscillating motion testing, described more in detail below.

3.1 Artificial Roller End Damage Test

The goal of artificial roller end damage end testing is to identify any characteristics of the failure mode. What separates this test rig from most other machines, is the capability to create damage while the bearing is running in the machine, emulating a particle in the lubricant. This allows control of when the scratch is made, as opposed to having loose particles in the lubricant. Figure 5 shows the mechanism for applying the damage. A hardened needle is inserted through a pre-drilled hole to scratch the roller. By varying the hole location and shape of the tip, analogous to changing the distance *h* in Fig. 3, the scratch location on the roller can be modified.

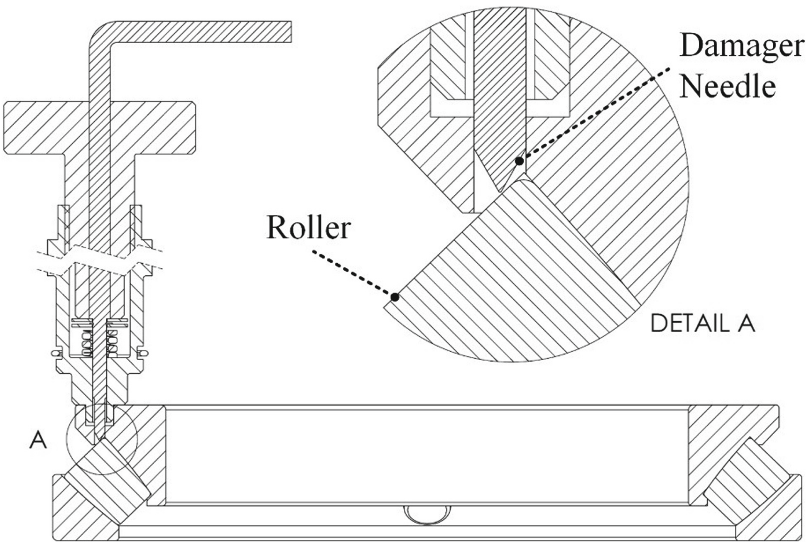


Fig. 5. Cross-section of artificial damage mechanism

3.2 Oscillating Motion Test

Heave-compensating winches is an example of equipment experiencing oscillating motion in parts of the operating cycle. Oscillating motion results in uneven distribution of stress on the bearing components, which is known to reduce fatigue lifetime (Houpert 1999). Oscillating motion is also known to cause fretting corrosion. Degradation of high loaded oscillating bearings was investigated by (Massi *et al.* 2014). For the context of this paper, oscillation implies that the rolling elements perform incomplete revolution before changing direction. The motor is a synchronous permanent magnet AC (PMAC) motor, capable of precise control of shaft torque and rotation at low speed.

3.3 Instrumentation

The test rig is prepared for accommodating a range of sensors. It is currently equipped with an axial load transducer, accelerometers and AE transducers. Additionally, the motor has an integrated encoder, making position, velocity, acceleration available from the Variable Frequency Drive (VFD) over Ethernet. Data is recorded using a National Instruments compactDAQ equipped with Analog Input (AI) and Digital Input/Output (DI/O) cards. An overview of the drive system, loading system and acquisition system is give in Table 1.

Table 1. System overview

Drive system	Loading system	Acquisition system
PMAC motor	Hydraulic	NI 9232-2 × 3 ch. 24 bit AI.
Direct drive	50 ton load capacity	NI 9205 - 16 ch, 16 bit AI
48 Nm rated torque	Manual control	NI 9401-8 ch, DI/O
3000 rpm limit	Force measurement	VFD Ethernet communication
		MISTRAS PCI-2 (AE specific)

Table 2 gives an overview of the sensors currently in use on the machine.

Table 2. Sensor specifications

Sensor type	Model
Accelerometer	TE Connectivity 805M1-0020-01
AE sensor	Physical Acoustics R15a
Force transducer	TECSIS F6210 500 kN

4 Discussion

This paper discusses roller end wear, in particular scoring on roller ends. The characteristic scoring shape is described mathematically, and a test rig is designed for replicating this failure mode through artificial roller end damage testing and oscillating motion testing. Through these tests, it may be possible to identify fault signatures for use in online condition monitoring systems. A data-driven approach allows the use of statistical tools for detection algorithms and bearing health assessment, providing decision-making assistance for maintenance actions. Compared to offline lubricant analysis, this gives more time to plan and execute maintenance.

Further, experiments with different needle shapes, speeds, lubricants and loads are possible. With minor modifications, the test rig can also accommodate different test bearings. While the test rig is prepared for vibration and AE measurements, other transducers can be retrofitted, including but not limited to motor current sensors, online lubrication monitoring and electrical methods for metal-to-metal contact detection.

Acknowledgement. The research presented in this paper has received funding from the Norwegian Research Council, SFI Offshore Mechatronics, project number 237896.

References

Bhuiyan, M.S.H., Choudhury, I.A., Dahari, M.: Monitoring the tool wear, surface roughness and chip formation occurrences using multiple sensors in turning. J. Manufact. Syst. **33**(4), 476–487 (2014). https://doi.org/10.1016/j.jmsy.2014.04.005

Crep, S., Bercea, I., Mitu, N.: A dynamic analysis of tapered roller bearing under fully flooded conditions Part 1: theoretical formulation. Wear **188**, 1–10 (1995)

Douglas, R.M., Steel, J.A., Reuben, R.L.: A study of the tribological behaviour of piston ring/cylinder liner interaction in diesel engines using acoustic emission. Tribol. Int. **39**(12), 1634–1642 (2006). https://doi.org/10.1016/j.triboint.2006.01.005

Fitzsimmons, B., Clevenger, H.D.: Contaminated lubricants and tapered roller bearing wear. ASLE Trans. **20**(2), 97–107 (1977). https://doi.org/10.1080/05698197708982822

Hemmer, M., Waag, T.I.: A Comparison of Acoustic emission and vibration measurements for condition monitoring of an offshore drilling machine. In: Proceedings of the Annual Conference of the Prognostics and Health Management Society, pp. 278–285 (2017). https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2017/phmc_17_031.pdf. Accessed 21 Mar 2018

Houpert, L.: Bearing life calculation in oscillatory applications©. Tribol. Int. **42**(1), 136–143 (1999). https://doi.org/10.1080/10402009908982200

Zhu, J., Yoon, J.M., He, D., Bechhoefer, E.: Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines. Wind Energy **17**, 1131–1149 (2014). https://doi.org/10.1002/we.1746

Jiaa, C.L., Dornfeld, D.A.: Experimental studies of sliding friction and wear via acoustic emission signal analysis. Wear **139**(2), 403–424 (1990)

Massi, F., et al.: Degradation of high loaded oscillating bearings: numerical analysis and comparison with experimental observations. Wear **317**(1–2), 141–152 (2014). https://doi.org/10.1016/J.WEAR.2014.06.004

Parker, R.J.: Large-Bore Tapered-Roller Bearing Performance and Endurance to 2.4 Million DN Apparatus and Procedure (1983). https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19830011863.pdf. Accessed 19 Feb 2018

Randall, R.B., Antoni, J.: Rolling element bearing diagnostics-a tutorial. Mech. Syst. Signal Process. **25**(2), 485–520 (2011). https://doi.org/10.1016/j.ymssp.2010.07.017

Taura, H., Nakayama, K.: Behavior of acoustic emissions at the onset of sliding friction (2018). https://doi.org/10.1016/j.triboint.2018.01.025

Wang, D., Tsui, K.L., Miao, Q.: Prognostics and health management: a review of vibration based bearing and gear health indicators. IEEE Access **6**, 665–676 (2018). https://doi.org/10.1109/ACCESS.2017.2774261

Zhu, X., Zhong, C., Zhe, J.: Lubricating oil conditioning sensors for online machine health monitoring – a review. Tribol. Int. **109**, 473–484 (2017). https://doi.org/10.1016/j.triboint.2017.01.015

Generation of a Tachometer Signal from a Smart Vibration Sensor

Eric Bechhoefer[1(✉)] and Rune Schlanbusch[2]

[1] GPMS Inc., Cornwall, VT, USA
eric@gpms-vt.com
[2] Teknova, Grimstad, Norway
rune.schlanbusch.2008@ieee.org

Abstract. Because of the bandwidth limitation of the engine controller, all helicopter gearboxes change speed over time. This change in speed necessitates the resampling of the data, based on a tachometer signal, to facilitate shaft, gear and bearing analysis of the rotating equipment. Without resampling, the quality of vibration analysis is degraded, and many mechanical faults would be missed. Further, interfacing with existing tachometer can be both expensive, and in some cases change certification requirements. The ability for a smart sensor to acquire vibration data, extract the shaft speed from the vibration data, and then process the data allows vibration-based fault detection capability at a lower cost, weight and reduced installation complexity than previously possible. Reducing cost, weight and installation complexity will expand the business case for condition monitoring, improving safety and reliability in industrial and transportation systems. This paper demonstrates a two-step process to recover a tachometer signal from vibration data that is of higher quality than raw tachometer data. Statistics are generated from known fault cases to demonstrate the effectiveness.

1 Introduction

Stewart (1977) introduced the time synchronous average (TSA) into gearbox analysis, and revolutionized condition-based maintenance on helicopters. His analysis, such as FM0/1/2/3/4, the "difference signal", the envelope analysis, etc., are used by most, if not all, Health and Usage Monitoring Systems (HUMS) currently available. Stewart's design is heavily dependent on a tachometer input to allow resampling of vibration data for the analysis. Further, Stewart recognized the importance of tachometer jitter, noting that there was a limit to the number of revolutions to take on the TSA (e.g. at some point, the TSA was not consistent).

The need to interface with a tachometer adds cost and complexity for every application. Further, there may be cases, such as glandless pumps (Boiler Circulator Pumps), where due to heat and pressure it is impractical or infeasible to install a tachometer sensor. In other situation, such as monitoring gas turbine engine, interfacing with the existing tachometer for the power turbine or compressor turbine, may change certification requirements (adding cost) and again increase system cost and weight. In other cases, the shaft simply may not be available for measuring the speed by a tachometer. These are just a few examples where adding a tachometer is impractical.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 758–771, 2020.
https://doi.org/10.1007/978-3-030-48021-9_84

Hence, the ability to derive a tachometer signal from vibration can reduce both cost and weight of a system and may facilitate new application where it is impossible to integrate a tachometer. The actual cost will of course vary due to the specific case as mentioned above, typically add 20% cost for a basic install (adding $2–5k on the total system cost) but adding impact of change in manufacturing process and testing may add up to $200k, while for recertification a full authority digital engine control (FADEC) system would cost millions of USD.

For helicopter applications, weight and cost of HUMS integration can be reduced using a bused, smart system (Bechhoefer et al. 2012). The use of a data bus greatly reduces interconnects, while smart sensors increases the scalability of the system. The smart vibration sensor also allows for local processing of the vibration data to reconstruct a tachometer signal on the sensor, so that the TSA and other tachometer dependent analysis can be performed in situ.

2 Tachometer Signal from Vibration

The ability for a smart sensor to acquire vibration data, extract the shaft speed from the vibration data, and then process the data, allows vibration-based fault detection capability at a lower cost and weight. This also results in reduced installation complexity over traditional tachometer-based system. Reducing cost, weight and installation complexity will expand the business case for condition monitoring, improving safety and reliability in light helicopters.

Cost is an important driver in the decision for operators to implement a HUMS. While it is well recognized that HUMS improves operational availability and reduces maintenance cost (through an installed rotor track and balance function, opportunistic maintenance, exceedance monitoring, etc.) the decision to install a HUMS for many operators is driven by cost. Removing a tachometer sensor and interface by adding this function to a smart sensor will improve the value proposition of a system to a customer.

2.1 The Need for a Tachometer

Spectral content of vibration is measured using the Fast Fourier Transform (FFT). The FFT is used in vibration-based diagnostics to: determine the magnitude and phase of component's vibration (such as shafts, gears or bearings), which can be indicative of wear and failure. Additionally, many common vibration analyses (residual or difference analysis, narrowband analysis (Bechhoefer et al. 2013), (Lebold et al. 2000), (McInerny et al. 2004) use the FFT for ideal filtering of the signal, or to perform a Hilbert transform of the signal (amplitude and frequency analysis).

The Fourier transform (and FFT) assumes that the signal under analysis is infinite in time. This assumption fails for real signals but can be mitigated using a window function (such as a Hamming or Hanning window). The other assumption in using the FFT is that the signal is stationary. Stationarity implies that the conditions of the signal do not change, i.e. the shaft rate is constant. This assumption in practice is violated; rotating machines have variation in their shaft rate. This change in rotational rate is due to changing load and the limits of the feedback control bandwidth.

The failure of stationarity results in "spectral smearing" of energy associated with a shaft. This in turn results in not measuring, accurately, the energy associated with a fault frequency.

2.2 Resampling Algorithms

The TSA resamples the vibration associated with a shaft or gear, in the spatial domain, such that vibration associated with each shaft order, in the Fourier domain, represent one frequency bin. For example, consider a system in which the shaft rate is such that for a given vibration sample rate, the acquisition system on average collects 800 samples per revolution. The TSA would resample the 800 samples to the 1,024 data points. The value 1,024 is the next highest radix 2 value. Radix 2 values are used because the simplest implementation of the FFT is based on powers or 2, i.e. radix 2 values.

Consider that if the load on the shaft is reduced, the shaft rate increases, and the measured vibration is reduced to 780 samples for the shaft revolution, as it takes less time for the shaft to make one revolution, hence the fewer samples. The 780 samples are resampled to 1,024 points. At some future time, say that the load on the shaft may increase, slowing the shaft, resulting in measuring 820 samples for one revolution, again the data is resampled to 1,024 points. For every revolution, the resampled data is summed point by point. After n revolutions, each of the 1,024 points of is divided by n, essentially time synchronously averaging the vibration data

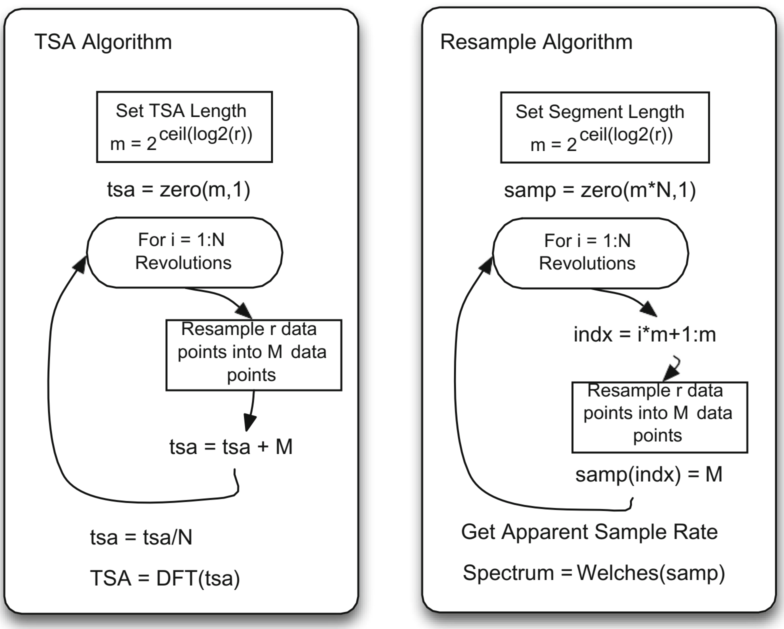


Fig. 1. Resampling algorithms for vibration based diagnostics.

Assume in this case that there is a 37-tooth gear on the shaft. The gear mesh energy of the gear would be a frequency of 37 times the shaft rate. In the Fourier domain, energy associated with the shaft rate would be in bin 2, and the gear mesh energy would be bin 38, and the second harmonic of that gear would be in bin 75 ($37 \times 2 + 1$, as the first bin is the DC energy). The TSA also reduces non-synchronous vibration by 1/r, where r is the total number of shaft revolutions which were used to construct the TSA. In this way, the TSA corrects for variation in the shaft rate and improves signal to noise.

The Time Synchronous Resampling (TSR) similarly resamples (i.e. up-samples) the vibration to correct variation in shaft speed. The apparent sample rate is the ratio of the total resampled time domain, vibration data set length divided by original data set length, multiplied by the original sample rate. A schematic comparison for TSA and TSR are shown in Fig. 1.

Both the TSA and TSR use a tachometer signal to calculate the time over which a shaft completes one revolution. The time taken for any shaft to complete one revolution can be calculated even if the tachometer is not associated with a given shaft. This is done by considering the shaft ratio from the shaft with the tachometer, to the shaft under analysis, then interpolating based on that tachometer signal.

2.3 The Tachometer Signal

In implementation, the tachometer signal is the rising edge of a voltage trigger from the passing of a shaft key phasor (i.e. a stationary point of the shaft). The tachometer signal is then converted to time. This time is accrued for each pass of the key phasor. In an architecture where the tachometer signal is measured using an analogue to digital converter (ADC), the resolution in time of the rising edge is 1 over the sample rate of the ADC. For condition monitoring purposes, the sample rate for a high-speed shaft could be 100,000 samples per second. In another architecture, the tachometer signal inputs into a voltage comparator. When the tachometer signal crosses zero (or some low voltage offset), the comparator goes high. The output of the comparator is monitored by the microcontroller using a general-purpose input/output (GPIO) pin. When the microcontroller senses the GPIO pin going high, it records the time. The resolution of time on the microcontroller is typically much higher than an ADC. For example, in a system using a 12 MHz clock, the microcontroller might run at 96 MHz, but the counter for time in the microcontroller would run at 48 MHz. The tachometer resolution in time would then be 2.0822e−8 s, or 480 times greater resolution than the ADC architecture.

Vibration signals from rotating equipment are sinusoidal and, by definition, synchronous with the shaft. However, the nature of vibration makes it impossible to use the vibration signal without significant signal process. Measured vibration is the superposition (i.e. addition) of many signals in the time domain. For example, consider a simple gearbox with an input shaft, and output shaft and a gear pair. The input shaft turns at 30 Hz, and has a 32-tooth gear, the output shaft has an 82-tooth gear with a speed of 11.707 Hz. The gear mesh frequency is 30 * 32 = 960 Hz. It is likely that the gear mesh frequency will have side bands because of any shaft imbalance being modulated onto the gear mesh. This can be proved using the trigonometric identity

$cos(a) \times cos(b) = 0.5[cos(a-b) + cos(a+b)]$ where in this example, cos(a) is 960 Hz, and cos(b) is 30 and/or 11.707 Hz shaft. Additionally, if the shaft is bent or bowed, there will be a 2x shaft vibration component. Other manufacturing defects such as the gear not being mounted perpendicular to the shaft, or not centering the shaft on the gear (i.e. eccentricity) adds additional tones.

3 Tachometer from Vibration

Bonnardot et al. (2005) described a method of using an acceleration signal of a gearbox to perform the TSA. This was further developed by Combet and Gelman (2007). These methods call for band pass filtering of the signal around a gear mesh, then in (Bonnardot et al. 2005) taking the Hilbert transform to estimate phase, and using phase to estimate zero crossing time. In (Combet and Gelman 2007), the signal was band pass filtered then demodulated, with the time resulting time domain signal used for the zero crossing.

Both solutions are sub optimal, due to the nature of using the FIR (Finite Impulse Response) band pass filter. Additionally, because modelling errors are present even after band pass filtering, the tachometer signal is corrupted by jitter. In this paper, we present a two-step method to reconstruct a tachometer signal from vibration. Step one is to ideally band pass filter and create an analytic signal in one functional procedure. Step two uses a jitter reduction model to remove noise (jitter) from the reconstructed tachometer signal not associated with changes in the shaft rate. While this system could be implement on any computer, it is ideally suited for implementation in a smart sensor, where cost and weight can be better addressed by lower cost condition monitoring system.

The pseudo code to recover a tachometer signal from vibration is:

Define the sample rate as sr The number of data points of vibration data, n = sr * n, where n is the acquisition length in seconds, then:

1. Calculate the next larger radix-2 length for the FFT: *nRadix = 2^ceil(log2(n)* where ceil is the ceiling function.
2. Calculate the low and high bandwidth index (b_{bwlow}, b_{bwhigh}), which are centered are a known gear mesh
3. Take the zero padded FFT of the vibration data
4. Zero the FFT from zero to b_{bwlow}, and from b_{bwhigh} to nRadix
5. Take the inverse FFT
6. Calculate the unwrapped argument of the signal from 1 to n time series
7. Normalize the time series of radians by the number of teeth of the gear (assuming 1^{st} harmonics)
8. Interpolate the number of indexes for every 2π rad.
9. Normalized to tachometer zero crossing times by sr.

A band pass filter is the convolution of a low pass filter with a high pass filter. These filters are implemented as FIR filters to improve their stability. Unfortunately, even rather large tap filters have poor filter response. Consider a case of a wind turbine gearbox, with approximately 29 Hz shaft speed and a 32-tooth gear, giving 928 Hz.

The bandwidth of the filter is set from 910 Hz to 950 Hz, to exclude the 29 Hz of the high-speed shaft (i.e. 29 * 32−29 = 899 Hz and 29 * 32 + 29 = 957 Hz). The filter response for this case, using a 120 tap FIR filter is seen in Fig. 2.

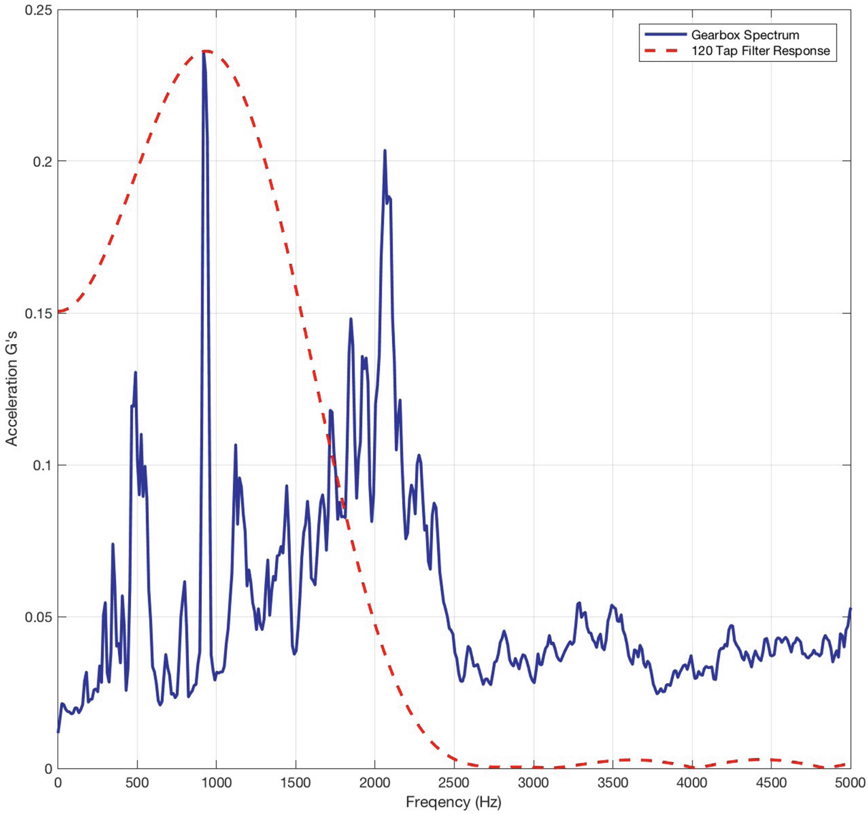


Fig. 2. Spectrum from a wind turbine high-speed shaft, and a 120 tap FIR band pass filter.

Note that the bandwidth (50% power, 3 dB) of this filter is 1,670 Hz. The filter does not reject the spectral content at 500 Hz, or at 1,121 Hz. These additional tones distort the desired analytic signal, reducing the quality of the resulting tachometer signal.

Consider a process in which the developed analytic signal uses an ideal filter. The analytic signal is defined for real valued signal *s(t)* as Eq. 2.

$$S(f) = F\{s(t)\} \tag{1}$$

Where F is the Fast Fourier Transform, then:

$$\begin{aligned} S_a(f) &= S(f), f = 0 \\ S_a(f) &= 2S(f), f > 0 \\ S_a(f) &= 0, f < 0 \\ S_a(t) &= F^{-1}(Sa(f)) \end{aligned} \tag{2}$$

where *S(f)* is as noted the Fourier transform of *s(t)*

For a signal which is sampled at 97,656 samples per second, for six seconds, the total length of $s(t)$ is n = 585,936 data points. As noted, there are advantages to using radix 2 lengths for the FFT. By zero padding the FFT to next larger radix two value, 2^20 = 1,048,576, the index representing the cut-off frequency for the band pass frequencies are: b_{bwlow} = 910 Hz/97,656 * 1,048,576 = 9,771, and b_{bwhigh} = 96/97,656 * 1,048,576 = 10,308. Then one can define the band pass analytic signal as:

$$\begin{aligned} S_a(f) &= 2S(f),\ bwlow \le i \le bwhigh \\ S_a(f) &= 0, f < bwlow, f > bwhigh \\ Sa(t) &= F^{-1}(S_a(f) \end{aligned} \tag{3}$$

There is no need to multiply by 2, as the argument (i.e. the angle) which is of interest, is the arctangent ratio of the imaginary parts of *sa(t)* and the real parts of *sa(t)*. The idealized band pass function rejects all signals not associated with the desired pass band. Figure 3 is a zoomed-in view of the spectrum in Fig. 2, comparing the pass band of the idealized filtered realized using Eq. 3.

The FIR filter does not reject the 500 Hz and 1,150 Hz tone. The idealized filter captures only the signal associated with the desired gear mesh tone. This allows for higher SNR and improved reconstruction of the tachometer signal from vibration. This tachometer signal is recovered from the arctangent of the analytic signal.

It is important to note that the arctangent function returns radians between 0 to π and –π to 0. But the evolution of the angle represents the incremental increase in phase for each sample in time. For example, the phase of the analytic signal for three cycles is: 2π * 3 or 6π. The arctangent of that signal will be −π to π for 3 cyclesi. The result of the arctangent must be unwrapped to capture these increases in phase vs. time. Unwrapping of the phase angle requires keeping track of the previous angle and current angle. The current angle is added to the previous angle, except when the returned arctangent goes from π to −π. In this case, π is added to the returned value to correct for the case when the returned value is between –π to 0.

After unwrapping the phase angle, the units are in radians per sample. While the FFT and inverse FFT operated on the radix 2 length (in this case 1,048,576), the argument and phase angle computation are performed only on the original sample length, n. Note that this time series of radians is for the gear mesh. To converts to radians per revolution of the shaft, the time series is divided by the number of teeth in the gear, in this example, 32.

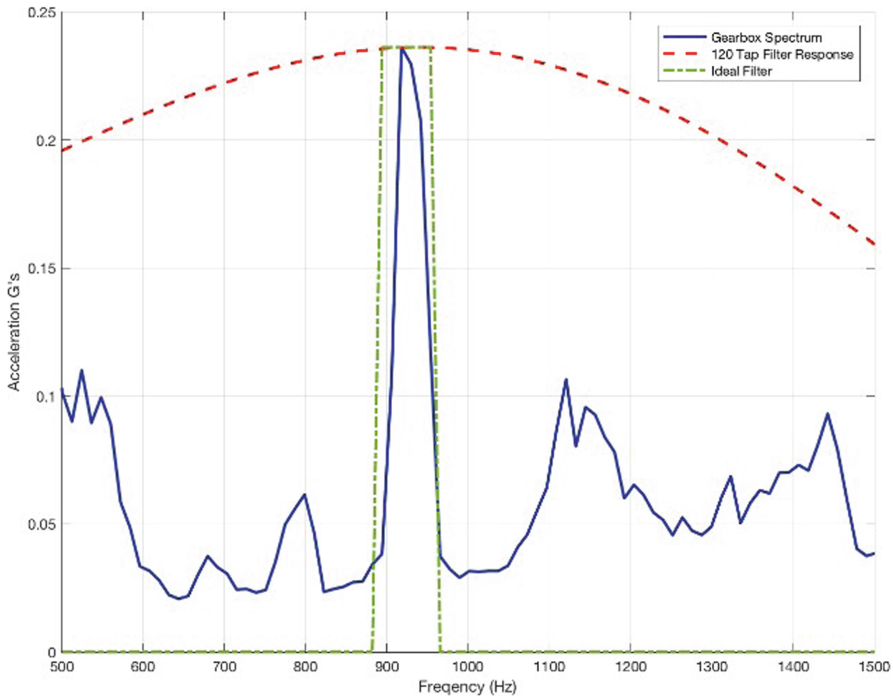


Fig. 3. FIR band pass filter vs. the idealized filtered realized using Eq. 3.

However, it may be that the strongest gear mesh tone is the 2nd or 3rd harmonic, in which case the pass band is adjusted accordingly, and the time series of radian angle is divided by $2n_{\mathrm{bccbO}}$ where n_{bccbO} is the number of teeth for the 2nd harmonic, etc.

The resulting time series represents the radian angle of the shaft, were each index advances the angle in time by dt, or $1/s$. Every 2π rad represent one shaft revolution. Because one is interested in the time, exactly every 2π, a form of interpolation is needed. For example, consider that the index just prior to 2π is 6.282780795474 (or 0.0004 less than 2π) at array index 3,395, while at index 3,396 the radian value is 6.284629142378, or 0.0014 greater than 2π. It is then a simple matter to interpolate between the index 3,395 and 3,396 for a radian value of 2π. In this case, the interpolated value is: 3395.21885053316.

This interpolation gives the number of indexes for each revolution. Note that this estimate of the tachometer zero cross signal is corrupted by noise. Combet et al. reports that the phase error standard deviation is related to the local signal to noise at the mesh harmonic k as:

$$\sigma[\delta] = 1/\sqrt{2}\,10^{-SNR/20} \tag{4}$$

Typical measured SNR is 6 to 8 dB. This suggests that the standard deviation of the phase error would be 6 to 10°. This phase is zero mean, but as it is non-zero: it will add jitter to the reconstructed tachometer signal.

3.1 Controlling Jitter

In (Bechhoefer and He 2014), it was shown that tachometer jitter contained a low frequency component associated with the engine control unit, and random, higher frequency components. Bechhoefer and He (2014) improved gear fault detection by using a zero phase, low order IIR, backward/forward filter. As noted previously, both the FIR and IIR filters bandwidths are defined by the 3 dB reduction in signal energy. The filter does not remove all signals above the bandwidth, and in fact reduces some signal energy below cut-off (up to 50%).

The idealized filter using the FFT processing is zero phase as well. This improves the analysis compared to the forward/backward process through the following procedure

1. Take the pseudo derivative of the tachometer signal
2. Calculate the radix-2 length of the pseudo derivative signal of length n
3. Zero pad the array from to the radix-2 length
4. Calculate the bandwidth index of the FFT
5. Idx = floor(bandwidth * radix-2 length/2);
6. Bandwidth is a normalized value, typically 0.12
7. Take the real FFT of the zero padded derivative signal
8. Set the real and imaginary parts of the FFT from Idx to the radix-2 length
9. Take the inverse real FFT.
10. Reconstruct the tachometer signal by taking the pseudo integral of the signal

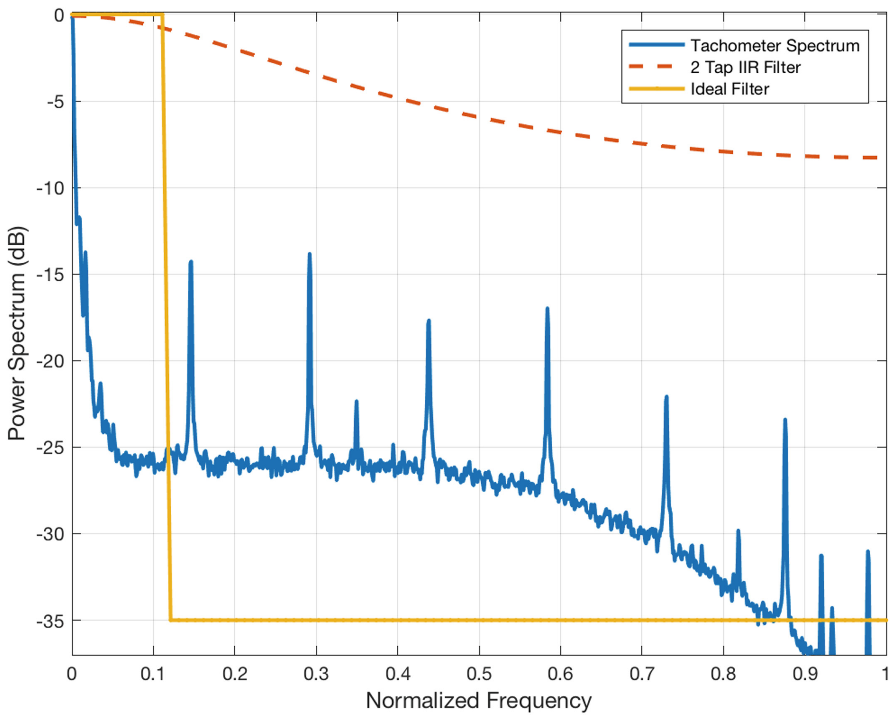


Fig. 4. Tachometer pseudo derivative spectrum compared to the IIR and ideal low pass filter.

Figure 4 demonstrates the effectiveness of the ideal low pass filter in removing jitter from the tachometer signal. The tachometer was from a wind turbine with 8 pulses per revolution. The tachometer signal was resampled, and the spectrum taken.

Note that the tachometer spectrum is primarily DC, with peaks associated with the unequal spacing of the targets. The idealized filter retains more low frequency response than the IIR filter, and effectively removes all spectral power (jitter) above the cut-off frequency.

4 Case 1: High Speed Pinion on a Wind Turbine

This machine was found, through vibration-based diagnostics, to have a cracked tooth on the high-speed pinion. The acquisition length was six seconds, sampled at 97,656 samples per second. The high-speed shaft is approximately 30 Hz.

This system was equipped with a tachometer, which is used as a reference to compare against the vibration-based tachometer signal. The installed tachometer was a Hall sensor using the GPIO architecture to capture zero crossing time. The tachometer target had eight pulses per revolution, but time was calculated using each 8th timing mark (i.e. one revolution). Figure 5 compares shaft speed derived from the Hall sensor

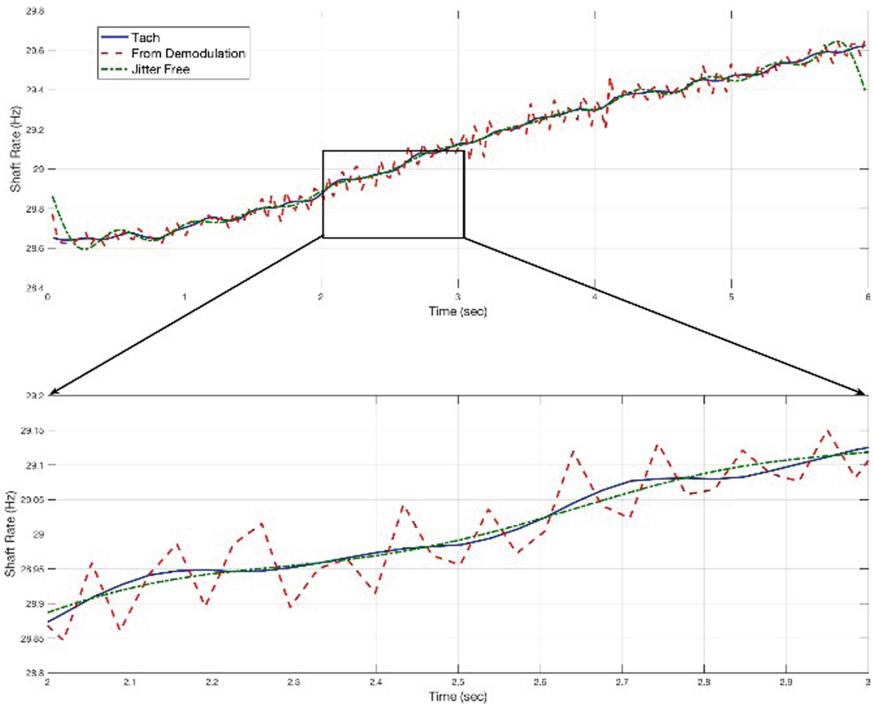


Fig. 5. Shaft rate from tachometer signal in comparison with jitter free tachometer from vibration.

base tachometer signal vs. the vibration-based tachometer signal vs. the final processing using the new jitter reduction technique. Wind turbine analysis is particularly difficult as there is a high degree of variability in the torque.

The bottom graph is a zoomed from the 2nd to the 3rd s in the acquisition. In general, after removing jitter using the new process, the vibration-based tachometer signal is not significantly different from the tachometer derived from the Hall sensor. Figure 6 depicts the TSA and the spectrum of the analysis.

The difference between the TSA is effectively only phase angle. From the spectrum, we can calculate the SNR to be approximately 6. Statistics derived from the TSA are given in Table 1.

Clearly, the gear fault was detectable with either a tachometer, or via the tachometerless vibration method.

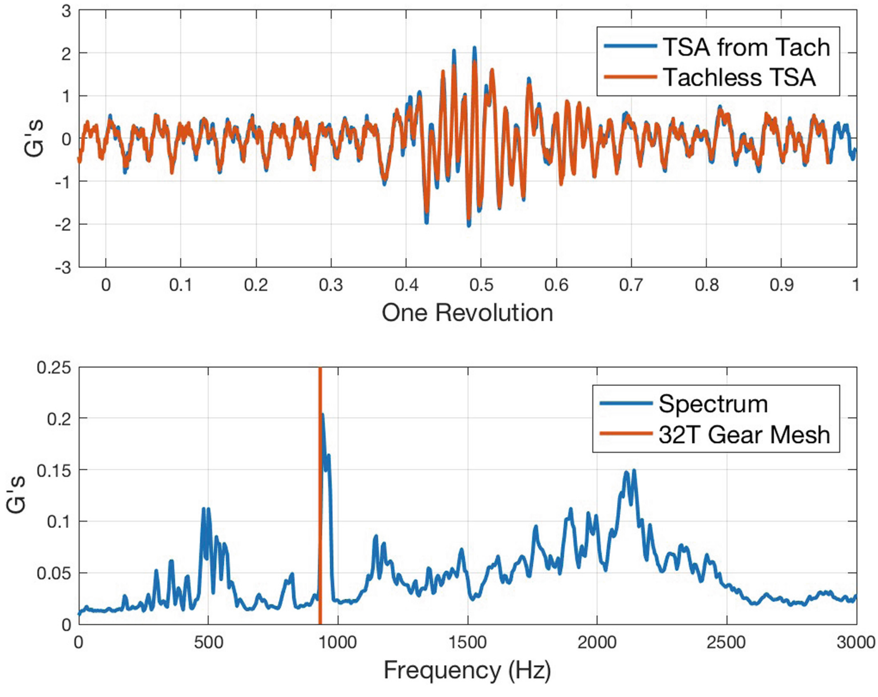


Fig. 6. Comparison of the TSA derived from a tachometer and vibration sensor.

5 Case 2: Tail Rotor Intermediate Gearbox

This was taken from the public domain data set, donated by NAVAIR AIR 4.4.2. The sample rate was 100,000 samples per second, using the ADC architecture for recording zero crossing data. The tachometer sensor was a variable reluctance speed sensor, using a 22-tooth gear as a target. The shaft rate for the target gear was approximately 3,000 rpm (i.e. 50 Hz), with a ratio from this shaft to the tail rotor drive shaft of 7.3:1. As seen in Fig. 7, the tachometer signal is noisy (high jitter), with a mean shaft rate of

68.5 Hz. The zoomed view shows the tachometer from vibration and with less jitter than the VR sensor tachometer signal, while the jitter free tachometer signal is smooth, capturing a slowly evolving change in shaft rate.

The SNR for this example is approximately 12 dB, due to the 65 g mesh frequency (Fig. 8). The resulting TSA, aside from a difference is phase, is indistinguishable from the TSA using a tachometer from the VR sensor.

Table 1. Tach vs. Tach for standard vibe statistics.

Analysis	Tach	Tach from vibe
SO1	0.0100 g	0.0104 g
SO2	0.0013 g	0.0016 g
SO3	0.0019 g	0.0018 g
TSA RMS	0.5091 g	0.4828 g
TSA P2P	2.0887 g	1.8430 g
FM0	4.278	4.00
AM RMS	0.100 g	0.099 g
AM Kurtosis	4.242	4.217
FM RMS	0.428 rad	0.426 rad
FM Kurtosis	4.995	4.844

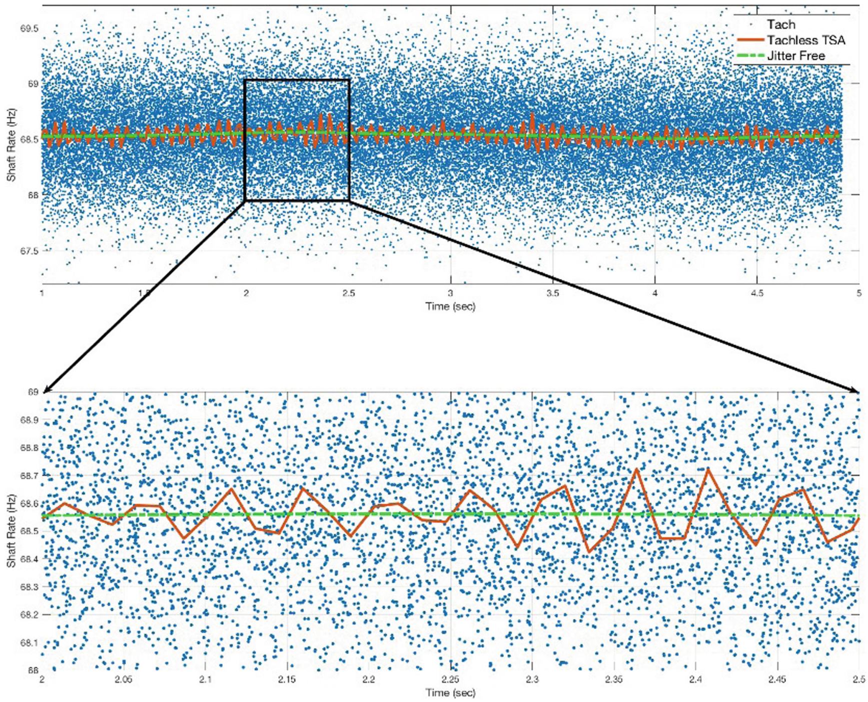


Fig. 7. VR tachometer vs. tach and jitter free tach from vibration.

Statistics derived from the TSA for the tail rotor input gear are given in Table 2.

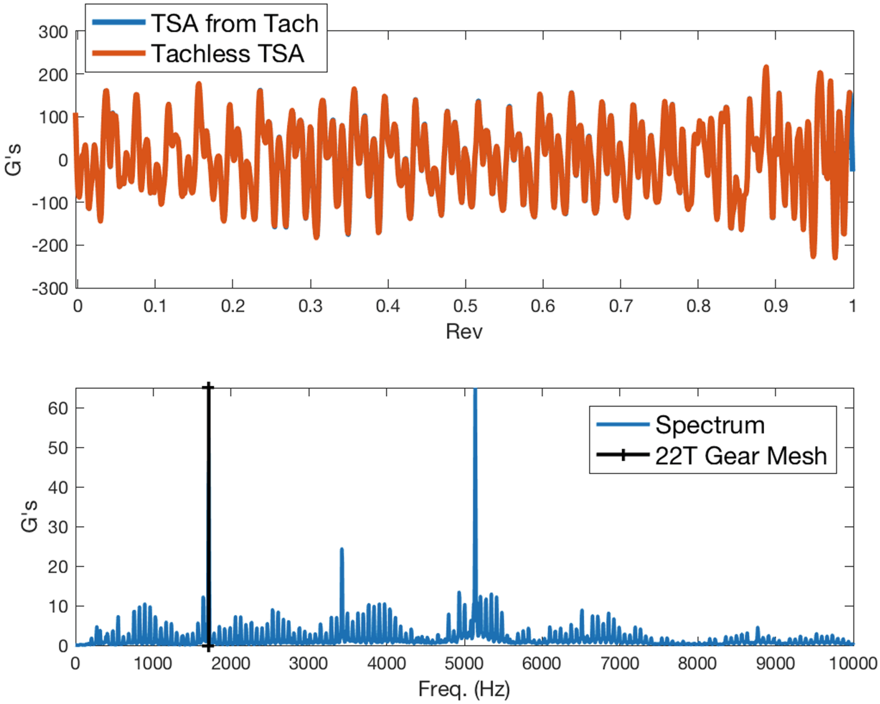


Fig. 8. Comparison of the TSA derived from a tachometer vs. vibration sensor.

Table 2. Tach vs. Tach for standard vibe statistics.

Analysis	Tach	Tach from vibe
SO1	0.043 g	0.043 g
SO2	0.283 g	0.282 g
SO3	1.855 g	1.85 g
TSA RMS	81.646 g	81.591 g
TSA P2P	222.77 g	223.02 g
FM0	16.94	15.38
AM RMS	17.06 g	17.06 g
AM Kurtosis	4.126	4.126
FM RMS	3.037 rad	3.048 rad
FM Kurtosis	2.46	2.45

6 Conclusions

A method to reconstruct the tachometer signal from vibration was demonstrated using a two-step process. The first step uses an idealized band pass filter implemented with a complex FFT to create an analytic signal, which is needed to derive the tachometer signal for the shaft under analysis. The second step uses another idealized low pass filter on the pseudo derivative of the reconstructed tachometer signal to remove jitter. This processing is run on a smart vibration sensor, which facilitates improved vibration analysis on rotating equipment where in the past the addition of a tachometer would be prohibitive due to cost, weight, certification requirements or being physically impractical.

The quality of the resulting analysis is compared favourably with traditional tachometers on real world fault data. In fact, the tachometer from vibration signal may have less noise, resulting in improved fault detection, than traditional tachometer sensor-based system.

References

Bechhoefer, E., Augustin, M., Kingley, M.: Architecture for a light helicopter HUMS. In: Proceedings of the AHS Forum, vol. 68 (2012)

Bechhoefer, E., He, D.: Improving gear fault detection by reducing tachometer jitter. In: Proceedings of the AHS Forum, vol. 72 (2014)

Bechhoefer, E., Van Hecke, B., He, D.: Processing for improved spectral analysis. In: Proceedings of the PHM Society Annual Forum (2013)

Bonnardot, F., El Badaoui, M., Randal, R.B., Daniere, J., Guillet, J.: Use of the acceleration signal of a gearbox in order to perform angular resampling (with limited speed fluctuations). Mech. Syst. Signal Process. **19**(4), 766–785 (2005)

Combet, F., Gelman, L.: An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor. Mech. Syst. Signal Process. **21**(6), 2590–2606 (2007)

Lebold, M., McClintic, K., Campbell, R., Byington, C., Maynard, K.: Review of vibration analysis methods for gearbox diagnostics and prognostics. In: Proceedings of the 54th Meeting of the Society for Machinery Failure Prevention Technology (2000)

McInerny, S.A., Sun, Q., Hardman, B.: Public domain algorithms for bearing and gear fault detection. In: Proceedings of the 11th International Conference on Structure Borne Sound and Vibration (2004)

NAVAIR AIR 4.4.2: Helicopter Drivetrain Vibration Condition Monitoring Algorithms. http://web.archive.org/web/20150514044141, http://qsun.eng.ua.edu:80/cpw_web_new/in-dex.htm

Stewart, R.M.: Some useful data analysis techniques for gearbox diagnostics. In: Proceeding of Meeting on Application of Time Series Analysis, Institute of Sound and Vibration Research (1977)

Temperature Measurements as a Method for Monitoring Ropes

Shaun Falconer[1(✉)], Ellen Nordgård-Hansen[2], and Geir Grasmo[1]

[1] Department of Engineering and Science, University of Agder, Grimstad, Norway
{shaun.falconer,geir.grasmo}@uia.no
[2] NORCE, Grimstad, Norway
enor@norceresearch.no

Abstract. Due to an increasing demand for operation at sea depths as low as 3000 m and under, the use of fibre ropes for offshore application in deep sea lifting and mooring is increasing. Consequently, improved knowledge is required regarding these ropes' thermo-mechanical properties and how these properties change as the rope is being used. This paper presents a 2D model of heat transport in the axial and radial directions along a 28 mm diameter fibre rope typically used for offshore applications. The model is combined with temperature measurements during heating and cooling of the rope, using both thermocouples and a thermal camera. Measurements are performed both on a new rope and on a used that has been through a high number of bending cycles. This allows for determining how the rope's thermal conductivity changes with use, which is a requirement for the use of temperature measurements as method for monitoring the rope's health.

1 Introduction

Fibre ropes have found use in various applications offshore including mooring and in general sailing. An interesting development is the use of fibre ropes in offshore construction cranes as an alternative to steel wires. The greater buoyancy of HMPE fibre rope opens the possibility of reaching water depths between 3000 and 4000 m, while at the same time presenting potential reduction in running costs through use of smaller vessels and cranes when compared to their steel rope counterparts. During use in subsea construction cranes, fibre ropes will be subjected to repeated bending cycles over a sheave, leading to higher temperatures created through inter-strand abrasion. The inherent issue with HMPE lies with its thermal properties, with some manufacturers reporting a maximum working temperature range of 65 °C (OTS 2017). Therefore, it is imperative that temperature measurements are incorporated into any condition monitoring system. Ultimately, better understanding on how to monitor these mechanisms will lead to better remaining useful life estimations, which in turn will contribute to more cost-effective condition based maintenance procedures.

This paper deals with experiments using thermocouples embedded in both used and unused fibre rope samples in two different heating conditions: homogeneous and heterogenous. The heating and cooling phases of each respective condition were

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 772–779, 2020.
https://doi.org/10.1007/978-3-030-48021-9_85

recorded and the cooling phase of the heterogenous regime is compared with observations from an infrared (IR) camera. The differences in thermal properties between the used and unused fibre rope specimens are investigated through the resulting thermocouple readings. An axisymmetric 2D model for the thermal behaviour in the axial and radial axes of the rope are presented based on a heat equation simulation, where the model parameters are fitted to the results from the experiments.

Temperature measurements play an important role in current offshore standards relating to the use of fibre ropes. Guidelines given by DNV-GL (DNV-GL 2015) use the 3-T parameter, where temperature is combined with tension and time to give lifetime estimates of ropes. Linear relationships between time and tension at a given temperature can be derived from the logarithmic time to rupture as a function of tension.

In cyclic-bend-over-sheave (CBOS) motion, the internal abrasion inside the rope is caused by strands in contact with each other, which generates heat. Figure 1 shows a CT scan of a fibre rope before and after CBOS testing by (Davies et al. 2013), where it can be clearly seen that the fibres have fused and created a denser rope. The influence of this change in structure on the thermal properties of the rope is of interest.

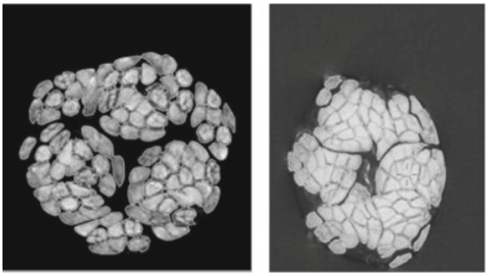


Fig. 1. X-ray tomography of a braided fibre rope before (left) and after (right) testing.

In terms of thermal condition monitoring techniques for fibre rope, there is a patent by (De Angelis 2002) where a conductive temperature element is embedded in the rope. Through this design it is proposed that a warning signal is sent to the operator should the maximum working temperature be exceeded.

Additionally, from a modelling perspective, (Oland et al. 2017) performed modelling of thermal properties of large diameter fibre ropes subject to CBOS motion.

2 Experiments

2.1 Equipment

The following equipment was used in the experiments:

- Six 28 mm Dextron 12 Plus fibre rope specimens (three used and three unused)
- Nabertherm Furnace (electrically heated, air atmosphere)

- Horst HTMC1 temperature regulator
- Fluke Ti400 IR camera
- K-type thermocouple wire
- Pico TC-08 data logger

The rope specimens were embedded with thermocouples at the positions shown in Fig. 2. Seven thermocouples were embedded in the rope at different positions and an eighth was used to measure the ambient temperature T_{amb}. The respective names and positions are named based on the system shown below, for example thermocouple "T_L_1" is the thermocouple placed in the left part of the region of interest at the centre of the rope corresponding to "L" and "1" respectively. The thermocouple readings were recorded using a data logger.

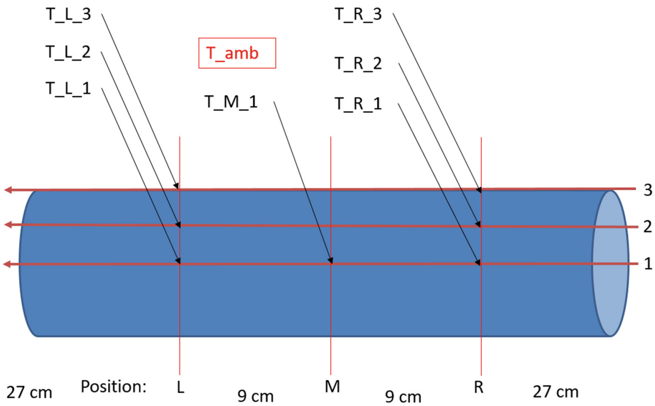


Fig. 2. Thermocouple set-up embedded in fibre rope.

The ropes used in the experiments are made up of three used and three unused specimens. The used specimens have been taken from a rope that was subject to testing in a CBOS machine after 11500 cycles at around 7% of the rope's maximum break strength. The portions of the rope used for analysis coincide with those in contact with the sheave during test operation.

2.2 2D Model

For modelling purposes, the rope is assumed to be an axisymmetric solid cylinder using the heat equation outlined in Eq. 1, neglecting radiation.

$$\rho c_p \frac{\partial T}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(k_r r \frac{\partial T}{\partial r}\right) + \frac{\partial}{\partial z}\left(k_z \frac{\partial T}{\partial z}\right) - h(T - T_{amb}) \tag{1}$$

Where ρ is density, c_p is specific heat capacity, k_r is the radial conductivity, k_z is the longitudinal conductivity, r is radius of cylinder, T is temperature, T_{amb} is ambient temperature and h is the heat transfer coefficient.

The heat loss term for convection is applied as a boundary condition at the exterior of the cylinder. Using the results of the experiments, the thermal properties of the model are adjusted to fit.

In the model, it is assumed there is an ambient temperature of 20 °C. The temperatures in the rope for the two types of testing (see below) are based on the steady states temperatures achieved during the respective heating phase. The rope is then cooled with a heat transfer coefficient defined on the boundary of the rope surface observed as time progresses.

2.3 Furnace Tests (Homogeneous Heating)

The furnace testing consisted of two phases: heating and cooling. Each rope specimen was placed in a furnace and kept until a stable temperature was reached. The temperature at the various thermocouple positions was recorded using a data logger. The furnace was set to achieve an ambient temperature of around 80 °C. Upon completion of the heating phase, the rope specimen was removed from the furnace and left to cool in still air on a rack. The same temperature recording process was performed during this phase of the testing. The furnace tests were performed to represent homogenous heating conditions and to specifically measure the radially dominated thermal properties of the rope. This process was performed on six fibre rope samples with three being used specimens and three being unused specimens.

2.4 Wrapping Tests (Heterogeneous Heating)

Like the furnace tests, the wrapping tests were performed in two phases: heating and cooling. Using a temperature regulator, the portion of the rope pertaining to the "R" position thermocouples was wrapped with heating bands from the temperature regulator and outer insulation. Once the thermocouple readings reached stable temperature, the wrapping was removed and the fibre rope specimens were left to cool on a rack in still air. These cooling phases were recorded using an IR camera as well as by thermocouples. The wrapping tests were performed to represent heterogeneous heating and to investigate the longitudinally dominated thermal properties of the rope. The process was performed on six fibre rope samples with three being used specimens and three being unused specimens.

3 Results and Discussion

3.1 Furnace Tests

Figure 3 highlights the temperatures recorded by the thermocouples at position "L" in both used and unused rope specimens during homogeneous heating (left) and cooling (right) phases of the homogeneous heating experiments for one sample of each rope type. These curves are representative for all cases. The results for the heating phase are taken from T_L_1 readings at 30 °C in the two cases and the changes in readings at T_L_1, T_L_2 and T_L_3 are observed for the subsequent 1500 s. From the graphs it appears that the readings for T_L_1 and T_L_2 only show minute differences from each other during the heating phase but it is noted that the temperature observed at position T_L_3 in the used case is higher than the corresponding thermocouple in the unused case.

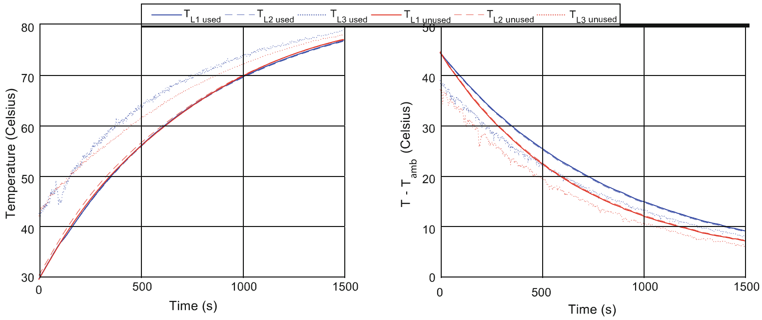


Fig. 3. Heating phases (left) and cooling phases (right) of used fibre rope specimen 3 and unused fibre rope specimen 3 respectively.

The positioning of the rope in the furnace has an influence on the temperature measured due to the air flow in the chamber resulting in these deviations, but from this for the used sample to exhibit similar thermal behaviour to the unused sample, a higher temperature is required. The behaviour could be attributed to the structural changes in the rope, such that the heat transfer coefficient of the used sample is lower than that of the unused sample. Essentially the structure is compressed such that less surface area of the strands is exposed to the atmosphere, resulting in a potentially lower heat transfer.

In addition to the air flow in the furnace, the thermocouple placement in the samples influence the temperature readings. While Fig. 2 highlights the positions of the various thermocouples inside the sample, there will be slight deviations on how they were attached in the samples, contributing to slight discrepancies in the results.

The cooling phase also revealed different thermal behaviour between the samples. The data was coordinated so the cooling phases in the difference between the temperatures recorded and the ambient temperature from 45 °C could be directly compared. The used sample is shown to cool at a slower rate than the unused sample. The readings shown by the T_L_3 thermocouples could be indicative that the used sample has a lower heat transfer coefficient to the surrounding air than the unused rope specimen.

3.2 Wrapping Tests

The heterogeneous heating phase on used and unused samples in the wrapping tests with the heat regulator was performed. The steady state temperatures achieved where then used as the starting values for the cooling model.

Figure 4 shows the correlation between the surface temperature reading T_R_3 and the IR camera for the cooling phases from 30 °C for used and unused specimens.

The IR camera readings for maximum and average temperature come from a small region of interest in the thermographic image centred on thermocouple T_R_3. In the used case, the maximum temperature reading from the IR camera shows good agreement with the T_R_3 reading before drifting slightly as time progressed. The IR camera readings in the unused case show slightly more drift but still some correlation with T_R_3 until around 700 s before totally diverging. The cause of this is unknown but it is apparent that in the time frame between 0 and 700 s that the used rope IR camera reading shows better agreement with its thermocouple equivalent. However, the results show that there is potential for an IR camera to give accurate readings of the outside temperature of a rope.

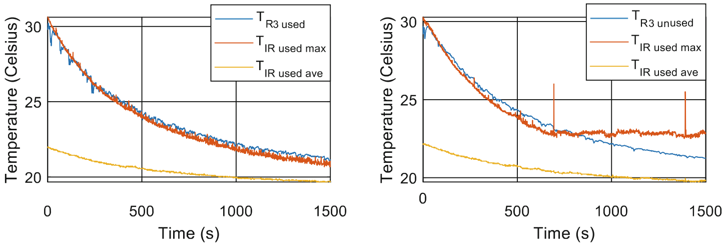


Fig. 4. Thermocouple reading T_R_3 for used fibre rope specimen 3 (left) and unused fibre rope specimen 3 (right) compared to IR camera readings during cooling phase.

3.3 2D Model

This model is used to compare the cooling phases of the used and unused fibre rope specimens in the furnace and wrapping tests. Figure 5 shows the results of the model and thermocouples cooling from 60 °C reading in the "T_L_1" thermocouple and the readings for the other sensors at the equivalent time for 1500 s. The models make use of different thermal properties for the used and unused cases, which have been fitted to experimental data and are highlighted in Table 1.

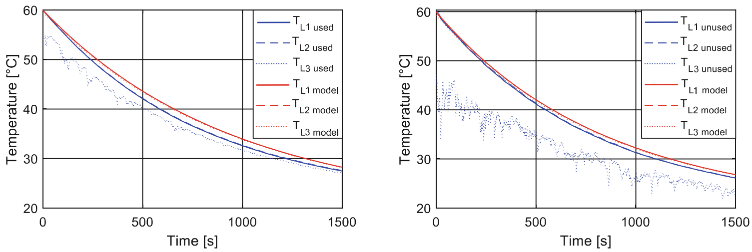


Fig. 5. Cooling phases for used rope specimen 1 (left) and unused rope specimen 1 (right) after furnace testing and compared to radial thermal models.

The unused case shows average agreement between the thermocouple readings and the model results for this period with a slight divergence with the progression of time.

Table 1. Thermal properties used in furnace cooling and wrapper cooling models.

Properties	Furnace cooling		Wrapper cooling	
	Used	Unused	Used	Unused
k_r [W/(m.K)]	8	6	8	6
k_z [W/(m.K)]			37	35
ρ [kg/m^3]	1019	970	1019	970
c_p [J/(K.kg)]	1850	1850	1850	1850
h [W/(m^{2}K)]	14	15	14	15
α [k_r/(ρ c_p)]	4.24×10^{-6}	3.34×10^{-6}	4.24×10^{-6}	3.34×10^{-6}

The unused case shows better agreement with the thermocouple readings but slightly diverges towards the end of the comparison period. To make the models agree with their respective thermocouple measurements, the thermal properties shown in Table 1 had to be altered. Using the assumption that the used specimens are denser and more compact that the unused specimens, this was reflected in the use of higher values for the thermal conductivity and density and a lower heat transfer coefficient. The results detailed in Table 1 were representative of the other rope samples tested.

Figure 6 and Table 1 show the results of the longitudinal model simulation and the thermal parameters used respectively. The model was simulated from the steady state temperatures recorded in the wrapping heating phase. The experimental and model data were compared in the cooling phase from 30 °C. Again, the model parameters were adjusted to fit the experimental data with slight deviations noted the agreement of the T_M_1 and T_L_1 parameters, however the T_R_1 showed good agreement in both models. The thermal properties used to fit the model again reflected that there are differences between the used and unused rope specimen samples.

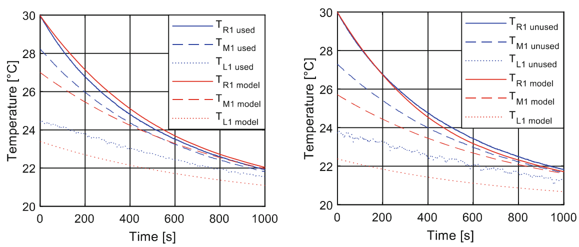


Fig. 6. Cooling phases for used rope specimen 2 (left) and unused rope specimen 3 (right) after wrapping tests compared to thermal model.

As the heat generated in ropes from the CBOS regime are a result of fibre to fibre motion resistance and abrasion, the highest temperatures would occur in this region. If the inside temperature of a rope can be estimated by its outside temperature in combination with IR camera monitoring it would form a vital component of a condition monitoring system and estimates of the rope's remaining useful life.

It must be noted however that this model is a simplified heat equation that does not fully account for the non-linear structure of the fibre rope samples. The samples will inherently contain different structures of the same fibres to bedding-in process of the rope as it is cycled over the sheave and as such will exhibit slightly different thermal responses in the cooling phase compared to the model.

4 Conclusions

From the experimental data it was clear that there are differences in the thermal properties of fibre ropes in their used and unused states. This knowledge highlights the importance of temperature measurements in condition monitoring of fibre ropes, as the rope structure would be subject to change as it progresses through the CBOS regime. It is also shown that this can be adequately modelled using a basic 2D finite difference model based on the standard heat equation. Additionally, the potential for IR camera usage has been highlighted. It is shown that accurate temperatures on the outside of a rope can be achieved as shown by the aligned experimental data from the thermocouples embedded in the rope specimens.

Future research directions potentially include testing on more rope specimens subjected to CBOS motion from intervals of higher number of cycles before failure. This would provide more understanding of the change in fibre rope thermal properties at specific time intervals during the progression of the CBOS regime. In addition to this, there is sufficient potential to improve the current 2D model, so it more accurately reflects the thermal response of the non-linearity in the fibre rope samples. Combining this through use of a thermal camera, there is potential for the inside rope temperature to be accurately measured based from the outside temperature.

Acknowledgements. The research presented in this paper has received funding from the Norwegian Research Council, SFI Offshore Mechatronics, project number 237896. The authors would also like to extend a special thanks to Gerrit Surup for his assistance during the experiments performed though this work.

References

Davies, P., et al.: Bend over sheave durability of fibre ropes for deep sea handling operations. In: Proceedings of the ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering, Nantes (2013)

De Angelis, C.: Synthetic fiber cable with temperature sensor. USA, Patenttinro US6392551 B2 (2002)

DNV-GL: DNVGL-RP-E305: Design, testing and analysis of offshore fibre ropes (2015)

Oland, E., et al.: Modelling the thermal properties of large diameter fibre ropes. In: OIPEEC, La Rochelle (2017)

OTS: Dextron 12 Plus Datasheet (2017)

Adaptive Canonical Variate Analysis for Machine Fault Detection and Identification

Xiaochuan Li(✉) and David Mba

Faculty of Technology, De Montfort University, Leicester, UK
{xiaochuan.li,david.mba}@dmu.ac.uk

Abstract. Canonical variable analysis (CVA) is one of the most commonly applied fault detection methods for industrial systems. However, conventional CVA is unable to handle the characteristics of time-varying processes. To address the challenge of implementing diagnostics in real-world applications with time-varying properties, the time-invariant CVA model is extended using first order perturbation theory, resulting in the improved adaptive CVA (ACVA) diagnostic model. Moreover, a bias-corrected exponentially weighted moving average approach is used in combination with the ACVA-based contributions to identify variables that are closely related to faults. The proposed technique is validated through condition monitoring data acquired from an operational centrifugal compressor. The results indicate that the proposed method can effectively detect and diagnose faults in real industrial systems under time-varying operating conditions.

1 Introduction

The growing interest in the reliability of industrial systems and continuing progress in the development of signal-processing techniques have motivated the development of advanced fault-detection methods for complex industrial systems. The complexity of large-scale industrial facilities hinders the development of first-principle dynamic models for health monitoring and fault analysis (Li et al. 2017). Existing condition-monitoring approaches for industrial processes are typically derived from routinely collected system-operating data. With the rapid growth and advancement in sensing and data-acquisition technologies, long-term continuous measurements can be acquired from different sensors mounted on machinery systems. However, the use of condition-monitoring data for reliable fault detection and diagnosis remains a challenge for researchers and engineers.

A number of multivariate statistical techniques have been developed based on condition-monitoring data for diagnostic health monitoring, such as filtering-based models (Guerra and Kolodziej 2017), multivariate time-series models (Serdio et al. 2014) and neural networks (Tran et al. 2014). Some of the key challenges in the implementation of these techniques are strongly correlated variables, changing operating conditions and inherent system uncertainty. Recent developments in dimensionality reduction techniques have yielded improvements in the detection of faults from highly correlated process variables. Conventional dimensionality reduction methods include principal component analysis (PCA) (Harrou et al. 2013), independent

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 780–790, 2020.
https://doi.org/10.1007/978-3-030-48021-9_86

component analysis (ICA) (Hyvärinen et al. 2004), partial least-squares analysis (PLS) (Kruger and Dimitriadis 2008). These basic multivariate methods perform well under the assumption that process variables are time independent. However, this assumption might not hold true for real industrial processes because sensory signals affected by noise and disturbances often show strong correlations between past and future sampling points (Jiang et al. 2015). Variants of these conventional multivariate approaches (Li and Qin 2001; Ruiz-cárcel et al. 2015; Yin et al. 2015) solve the time-independence problem have therefore been developed for dynamic processes monitoring.

Amongst the aforementioned multivariate diagnostic techniques, CVA is one of the most widely reported. There are many successful examples of the application of the CVA algorithm to model the dynamics of industrial systems (Juricek et al. 2004; Li et al. 2018; Pilgram et al. 2002; Stubbs et al. 2012). The requirement that process variables be IID and linear tends to limit the scope of CVA to linear processes operating under steady state conditions. Occasionally, problems associated with the effectiveness of the modelling tools can arise when the underlying assumptions are violated; for instance, the presence of nonlinear distortions and varying operating conditions. Therefore, it is necessary to develop an recursive monitoring approach based on adaptive canonical variate analysis for systems in which variations in the mode of operation are common. In this study, an adaptive canonical variate analysis (ACVA) approach based on the first order perturbation theory is developed to perform fault detection based on the abnormalities in the canonical state and the residual space. 2-D exponentially weighted moving average contribution plots for ACVA-based residual and state spaces are developed to identify variables that are closely associated with faults. The proposed technique is validated on industrial data captured from an operational centrifugal compressor.

2 Methodology

2.1 Canonical Variate Analysis

Given a data set $y_t \in \mathcal{R}^n$, where n is the number of variables, the past and future sample vectors $y_{a,t} \in \mathcal{R}^{na}$ and $y_{b,t} \in \mathcal{R}^{nb}$ are obtained from Eqs. 1 and 2:

$$y_{a,t} = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-a} \end{bmatrix} \in \mathcal{R}^{na} \tag{1}$$

$$y_{b,t} = \begin{bmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_{t+b-1} \end{bmatrix} \in \mathcal{R}^{nb} \tag{2}$$

where a and b are the length of the past and future sample vectors, respectively. The past and future sample vectors are then normalized to the zero-mean vectors $\hat{y}_{a,t}$ and $\hat{y}_{b,t}$. Then, the vectors $\hat{y}_{a,t}$ and $\hat{y}_{b,t}$ at different sampling times are rearranged according to Eqs. 3 and 4 to form the reshaped matrices $\hat{Y}_a$ and $\hat{Y}_b$:

$$\hat{Y}_a = \left[\hat{y}_{a,t+1}, \hat{y}_{a,t+2}, \ldots, \hat{y}_{a,t+N}\right] \in \mathcal{R}^{na \times N} \tag{3}$$

$$\hat{Y}_b = \left[\hat{y}_{b,t+1}, \hat{y}_{b,t+2}, \ldots, \hat{y}_{b,t+N}\right] \in \mathcal{R}^{nb \times N} \tag{4}$$

where $N = M - a - b + 1$ and M represents the total number of samples for y_t.

Subsequently, the convenience and cross-convenience matrices of $\hat{Y}_a$ and $\hat{Y}_b$ can be calculated as follows:

$$\Sigma_{a,a} = \hat{Y}_a \hat{Y}_a^T / (N - 1) \tag{5}$$

$$\Sigma_{b,b} = \hat{Y}_b \hat{Y}_b^T / (N - 1) \tag{6}$$

$$\Sigma_{b,a} = \hat{Y}_a \hat{Y}_b^T / (N - 1) \tag{7}$$

The objective of CVA is to find the maximum correlation between two sets of variables. The solution to this optimization problem can be obtained by performing singular value decomposition (SVD) on the Hankel matrix $\mathcal{H}$:

$$\mathcal{H} = \Sigma_{b,b}^{-1/2} \Sigma_{b,a} \Sigma_{a,a}^{-1/2} = U \Sigma V^T \tag{8}$$

The columns of $U = [u_1, u_2, \ldots, u_r]$ and the columns of $V = [v_1, v_2, \ldots, v_r]$ are called the left-singular and right-singular vectors of $\mathcal{H}$. Σ is a diagonal matrix; its diagonal elements are called singular values. The right-singular vectors in V corresponding to the largest q singular values are retained in the truncated matrix $V_q = \left[v_1, v_2, \ldots, v_q\right] \in \mathcal{R}^{na \times q}$. With the truncated matrix V_q, the na dimensional past vector $\hat{\mathrm{Y}}_{\mathrm{a}} \in \mathcal{R}^{na \times N}$ can be further converted to a reduced q-dimensional matrix $Z \in \mathcal{R}^{q \times N}$ as follows:

$$Z = \left[z_{t=1}, z_{t=2}, \ldots, z_{t=N}\right] = V_q^T \Sigma_{a,a}^{-1/2} \hat{Y}_a \tag{9}$$

where the columns of Z, $z_t = V_q^T \Sigma_{\mathrm{a,a}}^{-1/2} \hat{\mathrm{y}}_{\mathrm{a,t}}$, are the canonical variates at different sampling points. Similarly, the canonical residuals can be calculated as follows:

$$E = \left[\varepsilon_{t=1}, \varepsilon_{t=2}, \ldots, \varepsilon_{t=N}\right] = V_e^T \Sigma_{a,a}^{-1/2} \hat{Y}_a \tag{10}$$

where the columns of E, $\varepsilon_t = V_e^T \Sigma_{\mathrm{a,a}}^{-1/2} \hat{\mathrm{y}}_{\mathrm{a,t}}$, are the canonical residuals, and V_e contains the last $e = na - q$ columns of V in Eq. 8.

Motivated by the statistical indicators proposed in (Jiang et al. 2015), the statistics that provide information about the health of the system can be obtained as follows:

$$T_{d,t}^2 = z_t^T z_t \tag{11}$$

$$T_{e,t}^2 = \varepsilon_t^T \varepsilon_t \tag{12}$$

where $T_{d,t}^2$ represents the variations inside the canonical state space and $T_{e,t}^2$ represents the variations in the residual space. The probability density function of the health indicators in this study is calculated by using kernel density estimations (KDE), as proposed by Odiowei and Cao (Odiowei and Yi 2010).

2.2 Adaptive Canonical Variate Analysis Based on Efficient Adaptive SVD

Due to non-stationary process behaviour, many industrial processes have time-varying characteristics. As a result, the sample covariance matrices $\Sigma_{a,a}$ and $\Sigma_{b,b}$ and the cross-covariance matrix $\Sigma_{a,b}$ vary over time. Therefore, the exponential weighted moving-average method is employed in this investigation to update the covariance and cross-covariance matrices:

$$\Sigma_{a,a(t)} = (1-\beta) y_{a,t} y_{a,t}^T + \beta \Sigma_{a,a(t-1)} \tag{13}$$

Recursive modelling identification requires the computation of the nonzero singular values and the corresponding left and right singular vectors in Eq. 8 for each new observation. In this paper, the first order perturbation theory (Willink 2008) is adopted to realize an adaptive canonical variate analysis approach.

We denote z_p by $z_p = \left(y_{a,t} y_{a,t}^T\right) y_{a,t}$ and z_f by $z_f = \left(y_{b,t} y_{b,t}^T\right) y_{b,t}$. According to (Shang et al. 2016), the left and right singular vectors can be recursively updated as

$$\overline{u_{t,i}} = u_{t-1,i} + \frac{x_i^T}{\sigma_{t-1,i}} \sum\nolimits_{j=i+1}^{N} y_j u_{t-1,j} - y_i^T \sum\nolimits_{j=1}^{i-1} \frac{x_j}{\sigma_{t-1,j}} u_{t-1,j} \tag{14}$$

$$\overline{v_{t,i}} = v_{t-1,i} + \frac{y_i^T}{\sigma_{t-1,i}} \sum\nolimits_{j=i+1}^{N} x_j v_{t-1,j} - x_i^T \sum\nolimits_{j=1}^{i-1} \frac{y_j}{\sigma_{t-1,j}} v_{t-1,j} \tag{15}$$

where x_i and y_i are the i th element of vectors $x = \sqrt{1-\beta} U_{t-1}^T z_p$ and $y = \sqrt{1-\beta} V_{t-1}^T z_f$, respectively.

The series of vectors p_i^u and q_{i+1}^u can be defined for $i = 1, \ldots, N-1$, as

$$p_i^u = \sum\nolimits_{j=i+1}^{N} y_j u_{t-1,j} = p_{i-1}^u - y_i u_{t-1,i} \tag{16}$$

$$q_{i+1}^u = \sum\nolimits_{j=1}^{i} \frac{x_j}{\sigma_{t-1,j}} u_{t-1,j} = q_i^u + \frac{x_i}{\sigma_{t-1,i}} u_{t-1,i} \tag{17}$$

$$p_i^v = \sum\nolimits_{j=i+1}^{N} x_j v_{t-1,j} = p_{i-1}^v - x_i v_{t-1,i} \tag{18}$$

$$q_{i+1}^v = \sum\nolimits_{j=1}^{i} \frac{y_j}{\sigma_{t-1,j}} v_{t-1,j} = q_i^v + \frac{y_i}{\sigma_{t-1,i}} v_{t-1,i} \tag{19}$$

where $p_0^u = \sqrt{1-\beta} z_p$, $q_0^u = 0$, $q_0^v = \sqrt{1-\beta} z_f$, $q_0^v = 0$. Then the updated left and right singular vectors can be obtained as follows:

$$\overline{u_{t,i}} = u_{t-1,i} + \frac{x_i^T}{\sigma_{t-1,i}} p_i^u - y_i^T q_i^u \tag{20}$$

$$\overline{v_{t,i}} = v_{t-1,i} + \frac{y_i^T}{\sigma_{t-1,i}} p_i^v - x_i^T q_i^v \tag{21}$$

According to (Willink 2008), the updated singular vectors should be normalized to give the scaled left and right singular vectors $u_{t,i}$ and $v_{t,i}$ and singular values:

$$u_{t,i} = \frac{\overline{u_{t,i}}}{\left\|\overline{u_{t,i}}\right\|}, \quad v_{t,i} = \frac{\overline{v_{t,i}}}{\left\|\overline{v_{t,i}}\right\|}, \quad \sigma_{t,i} = \left|\sigma_{t,i}\right| \tag{22}$$

Given the scaled updated singular values and singular vectors, the new canonical variates and residuals can be calculated as per Eqs. 9 and 10.

2.3 Fault Identification

In addition to fault time detection, the canonical state space and residual space statistics are also used to calculate variable contributions. The contribution of variable n_k $(1 \le k \le n)$ obtained from the state space at time t is defined as follows:

$$\begin{aligned} c_{d,t} &= z_t^T z_t = z_t^T (V_q^T \Sigma_{a,a}^{-1/2} \hat{y}_{a,t}) = z_t^T \sum\nolimits_{i=1}^{na} \left(\hat{y}_{a,t} \left(V_q^T \Sigma_{a,a}^{-1/2} \right)^T \right)^T \\ &= \sum\nolimits_{i=1}^{na} z_t^T \left(\hat{y}_{a,t} \left(V_q^T \Sigma_{a,a}^{-1/2} \right)^T \right)^T \end{aligned} \tag{23}$$

where $\hat{y}_{a,t}$ denotes the column vector of $\hat{Y}_a$ at time instant t. In order to consider the correlations between past and future sampling points, the contribution for variable y_n based on the canonical state space is calculated by summing up all its past a observations:

$$c_{d,t}^{y_n} = \sum\nolimits_{i=1}^{a} z_t^T \left(\hat{y}_{a,t-i} \left(V_q^T \Sigma_{a,a}^{-1/2} \right)^T \right)^T \tag{24}$$

Similarly, the residual space contributions at time instant t are calculated as follows:

$$\begin{aligned} c_{e,t} &= \varepsilon_t^T \varepsilon_t = \varepsilon_t^T (V_e^T \Sigma_{a,a}^{-1/2} \hat{y}_{a,t}) = \varepsilon_t^T \sum\nolimits_{i=1}^{na} \left(\hat{y}_{a,t} \left(V_e^T \Sigma_{a,a}^{-1/2} \right)^T \right)^T \\ &= \sum\nolimits_{i=1}^{na} \varepsilon_t^T \left(\hat{y}_{a,t} \left(V_e^T \Sigma_{a,a}^{-1/2} \right)^T \right)^T \end{aligned} \tag{25}$$

The contribution of the variable y_n based on the residual space is calculated as follows:

$$c_{e,t}^{y_n} = \sum\nolimits_{i=1}^{a} z_t^T \left(\hat{y}_{a,t-i} \left(V_q^T \Sigma_{a,a}^{-1/2} \right)^T \right)^T \tag{26}$$

According to the literature (Jiang et al. 2015), CVA contributions can be excessively sensitive because of the inversion procedure $\Sigma_{\mathrm{a,a}}^{-1/2}$, resulting in incorrect diagnosis of faulty variables. To alleviate this sensitivity, a combination of residual and state space contributions is adopted for the identification of faulty variables as:

$$c_{de,t} = 0.5 * c_{d,t} + 0.5 * c_{e,t} \tag{27}$$

In order to implement an efficient monitoring method, we must consider the process data serial correlations. A bias-corrected exponentially weighted moving average approach is adopted in this study to improve the fault identification rate of the conventional combined contributions. The contribution of variable y_n based on the bias-corrected exponentially weighted moving average approach can be obtained as

$$c_t = \frac{(1-\delta)c_{de,t} + \delta c_{t-1}}{1-\delta^t} \tag{28}$$

$$c_{t-1} = \frac{\sum_{k=t-W}^{t-1} c_{de,t}}{W} \tag{29}$$

where $\delta = W - 1/W$ is the forgetting factor and W is the width of the moving window.

3 Case Study

To evaluate the performance of the proposed technique for fault detection and diagnosis, the model was tested using condition-monitoring data obtained from an operational industrial compressor (hereafter referred to as compressor A). The measured time series from compressor A consisted of 238 observations and 13 variables. For this study, all data were captured at a sampling rate of one sample per hour by the

Table 1. Measured variables of compressor A

ID	Variable Name	ID	Variable Name	ID	Variable Name
1	Speed	6	Radial vibration overall X1	11	Radial bearing temperature 2
2	Suction pressure	7	Radial vibration overall Y1	12	Active thrust bearing temperature
3	Discharge pressure	8	Radial bearing temperature 1	13	Inactive thrust bearing temperature
4	Discharge temperature	9	Radial vibration overall X2		
5	Actual flow	10	Radial vibration overall Y2		

machine's condition-monitoring system. Table 1 summarizes all measured variables for this compressor.

As shown in Fig. 1, the compressor is operated under healthy but time-varying conditions during the first 200 points of the time series. The operational speed of the machine changes dramatically at around the 10th, 50th and 170th sampling points. The readings of the four different bearing-temperature sensors start to increase at around the 200th sampling point; the machine continued to run until the 238th sampling point. At that time, site engineers shut down the compressor for inspection and maintenance. A data set captured under healthy operating conditions were used as the training data

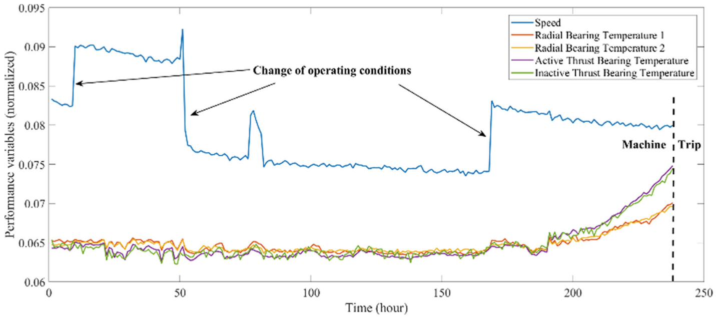


Fig. 1. Trend of speed and four different bearing temperature sensor measurements of compressor A.

for the ACVA diagnostic model, and the data captured throughout the entire monitored time series as shown in Fig. 1 were used to validate the trained model.

In order to obtain the lowest false alarm rate for both health indicators, it is necessary to adaptively determine the number of significant singular values q. Different methods can be used to decide the number of retained states q. The most commonly used methods are based on the Akaike information criterion (AIC) (Lee et al. 2006) and the dominant singular values (SV) in the diagonal matrix $\sum$ (Yang et al. 2012). In (Cárcel et al. 2007), CVA was implemented to test the false alarm rate for a healthy data set using different values of q in order to find the optimal state order that gives the lowest false alarm rate. The above-mentioned methods are not necessarily suitable for adaptive CVA because the information retained in the past state variates may not be able to fully represent the system dynamics in the future. In this study, the cumulative percentage variance was adopted to recursively determine the optimal number of retained states q. The cumulative variance (CV) is calculated as the first m eigenvalues divided by the sum of all eigenvalues:

$$CV_t = \frac{\sum_{i=1}^{m} \sigma_m}{\sum_{i=1}^{na} \sigma_m} \times 100\% \tag{30}$$

The optimal model order q was selected such that the cumulative variance can approximately explain 90% of the total variance. The value of forgetting factor β in Eq. 13 is set to 0.983 in this case. The upper control limit for normal operating conditions was calculated at the 99% confidence level with the aim of minimizing the false-alarm rate of the testing data set. The window length in Eq. 29 is set to 3.

Figures 2 and 3 show the results obtained in terms of fault detection for compressor A. The conventional CVA approach gives false alarms during the first 170 points of the time series. Compared with the results of CVA, the ACVA approach is able to distinguish normal deviations in operational conditions from real faults incurring dynamics anomalies and thereby significantly reduce the false alarm rate in normal conditions. Figure 4 illustrates the contributions of different process variables during the monitoring process with the variable number on the vertical axis and sampling time on the horizontal axis. Based on the information provided by the conventional CVA contribution plot, site engineers might mistakenly consider variables 1, 5, 11, 12 and 13 to be the root cause of the fault. However, the contribution plot based on the ACVA approach suggests that the variables contributing more to the detected fault are variables 8, 11, 12 and 13 because they show consistent strong bands of contribution over the entire degradation process, consistent with the monitoring data shown in Fig. 1. Table 2 summarizes the false alarm rate (FAR) and false identification rate (FIR) of the CVA and ACVA approach. The false alarm rates were computed as the percentage of sampling points for which the individual statistics exceed their threshold outside faulty conditions.

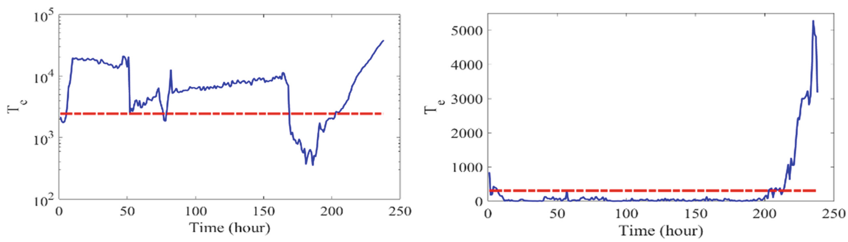


Fig. 2. Fault-detection results for compressor A. T_e^2 of ACVA (left) and CVA (right).

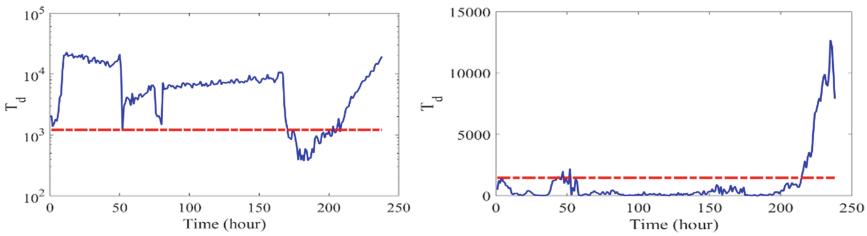


Fig. 3. Fault-detection results for compressor A. T_d^2 of ACVA (left) and CVA (right).

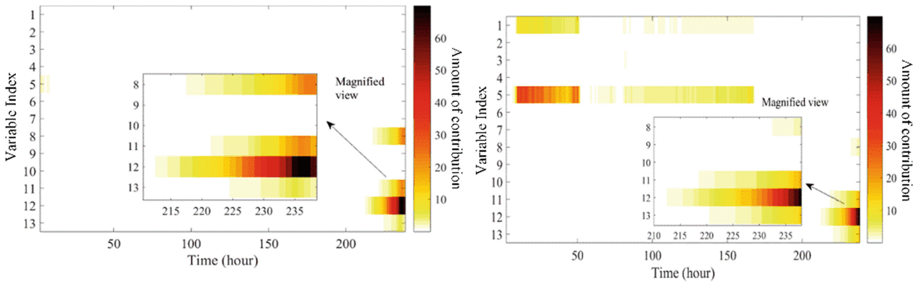


Fig. 4. Contribution plots for compressor A showing the contributions of the different variables over the entire times series. Bias-corrected exponentially weighted moving average contributions (upper) and CVA contributions (lower).

Table 2. False alarm rate (FAR) and fault identification rate (FIR) of the CVA and ACVA approach

	CVA		ACVA	
	T_e^2	T_d^2	T_e^2	T_d^2
FAR	79.5%	82.9%	**1.95%**	**1.95%**
FIR	83.34%		**98.07%**	

4 Conclusion

In this study, we developed an adaptive multivariate statistical monitoring tool for the monitoring of dynamic processes where variations in operating conditions are incurred. Condition monitoring data acquired from an operational centrifugal compressor were used to test the capabilities of the developed ACVA model for fault detection and identification. Fault detection was implemented by comparing the values of canonical state and residual space statistics with pre-determined thresholds. The faults were successfully detected by both health indicators. The statistics based on the ACVA approach are able to adapt to changes in operating conditions, leading to a lower false-alarm rate than the conventional CVA approach. Once a fault was detected, ACVA-based contribution plots were utilized to identify the variables most likely related to the specific fault. Overall, the ACVA-based contributions were very effective in identifying the root causes of the faults under study. The results showed that the bias-corrected exponentially weighted moving average contribution, which assembles the variations at multiple time instants, can clearly demonstrate the contributions of different process variables over the entire fault-propagation process and can provide greater insight into the root causes of the faults.

References

Cárcel, C.R., Cao, Y., Mba, D.: A benchmark of canonical variate analysis for fault detection and diagnosis. In: 2014 UKACC International Conference on Control, 9th–11th July 2014, Loughborough, U.K., pp. 329–356 (2007). https://doi.org/10.1016/S0167-4137(07)80012-6

Guerra, C.J., Kolodziej, J.R.: A data-driven approach for condition monitoring of reciprocating compressor valves. J. Eng. Gas Turb. Power **136**(1), 1–13 (2017). https://doi.org/10.1115/1.4025944

Harrou, F., Nounou, M.N., Nounou, H.N., Madakyaru, M.: Statistical fault detection using PCA-based GLR hypothesis testing. J. Loss Prev. Process Ind. **26** (2013). https://doi.org/10.1016/j.jlp.2012.10.003

Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. Wiley, Hoboken (2004)

Jiang, B., Huang, D., Zhu, X., Yang, F., Braatz, R.D.: Canonical variate analysis-based contributions for fault identification. J. Process Control **26**, 17–25 (2015). https://doi.org/10.1016/j.jprocont.2014.12.001

Juricek, B.C., Seborg, D.E., Larimore, W.E.: Fault detection using canonical variate analysis. Ind. Eng. Chem. Res. **43**(2), 458–474 (2004)

Kruger, U., Dimitriadis, G.: Diagnosis of process faults in chemical systems using a local partial least squares approach. AIChE J. **54**(10), 2581–2596 (2008)

Lee, C., Choi, S.W., Lee, I.B.: Variable reconstruction and sensor fault identification using canonical variate analysis. J. Process Control **16**(7), 747–761 (2006). https://doi.org/10.1016/j.jprocont.2005.12.001

Li, W., Qin, S.J.: Consistent dynamic PCA based on errors-in-variables subspace identification. J. Process Control **11**(6), 661–678 (2001)

Li, X., Duan, F., Loukopoulos, P., Bennett, I., Mba, D.: Canonical variable analysis and long short-term memory for fault diagnosis and performance estimation of a centrifugal compressor. Control Eng. Pract. **72**(January), 177–191 (2018). https://doi.org/10.1016/j.conengprac.2017.12.006

Li, X., Duan, F., Mba, D., Bennett, I.: Multidimensional prognostics for rotating machinery: a review. Adv. Mech. Eng. **9**(2), 1–20 (2017). https://doi.org/10.1177/1687814016685004

Odiowei, P.E.P., Yi, C.: Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. IEEE Trans. Ind. Inf. **6**(1), 36–45 (2010). https://doi.org/10.1109/TII.2009.2032654

Pilgram, B., Judd, K., Mees, A.: Modelling the dynamics of nonlinear time series using canonical variate analysis. Physica D **170**, 103–117 (2002)

Ruiz-cárcel, C., Cao, Y., Mba, D., Lao, L., Samuel, R.T.: Statistical process monitoring of a multiphase flow facility. Control Eng. Pract. **42**, 74–88 (2015). https://doi.org/10.1016/j.conengprac.2015.04.012

Serdio, F., Lughofer, E., Pichler, K., Buchegger, T., Pichler, M.: Fault detection in multi-sensor networks based on multivariate time-series models and orthogonal transformations. Inf. Fusion **20**, 272–291 (2014). https://doi.org/10.1016/j.inffus.2014.03.006

Shang, L., Liu, J., Zhang, Y.: Recursive fault detection and identification for time-varying processes. Ind. Eng. Chem. Res. **55**(46), 12149–12160 (2016). https://doi.org/10.1021/acs.iecr.6b02653

Stubbs, S., Zhang, J., Morris, J.: Fault detection in dynamic processes using a simplified monitoring-specific CVA state space modelling approach. Comput. Chem. Eng. **41**, 77–87 (2012). https://doi.org/10.1016/j.compchemeng.2012.02.009

Tran, V.T., Althobiani, F., Ball, A.: An approach to fault diagnosis of reciprocating compressor valves using Teager – Kaiser energy operator and deep belief networks. Expert Syst. Appl. **41** (9), 4113–4122 (2014). https://doi.org/10.1016/j.eswa.2013.12.026

Willink, T.J.: Efficient adaptive SVD algorithm for MIMO applications. IEEE Trans. Sig. Process. **56**(2), 615–622 (2008). https://doi.org/10.1109/TSP.2007.907806

Yang, Y., Chen, Y., Chen, X., Liu, X.: Multivariate industrial process monitoring based on the integration method of canonical variate analysis and independent component analysis. Chemometr. Intell. Lab. Syst. **116**, 94–101 (2012). https://doi.org/10.1016/j.chemolab.2012.04.013

Yin, S., Zhu, X., Member, S., Kaynak, O.: Improved PLS Focused on Key-Performance-Indicator-Related Fault Diagnosis. IEEE Trans. Ind. Electron. **62**(3), 1651–1658 (2015)

Modified Training and Optimization Method of Radial Basis Function Neural Network for Metrics Performance Guarantee in the Auto Association of Sensor Validation Tool

Marco A. D. Alves[1(✉)], Luigi Galotto Jr[1(✉)], João O. P. Pinto[1(✉)], Herbert Teixeira[2], and Mário C. M. Campos[2]

[1] Federal University of Mato Grosso Do Sul - UFMS, Campo Grande, Brazil
marcodual@hotmail.com,
{luigi.galotto,joao.pinto}@ufms.br

[2] CENPES Research Center of PETROBRAS, Rio de Janeiro, Brazil
{herbertteixeira,mariocampos}@petrobras.com.br

Abstract. This work presents the use of radial basis function artificial neural network to estimate the sensors measurements, exploring the analytical redundancy existent among different sensors in a process. However, in order to guarantee good performance of the network the training and optimization process was modified. In the conventional training algorithm, although the stop criteria, such as summed squared error, is reached, one or more of the individual performance metrics of the neural network may not be satisfactory. The performance metrics considered are Accuracy (training error), Sensitivity matrix (sensors propagated error to the estimations) and Filtering matrix (sensor propagated noise to the estimations). The paper describes the proposed method including all the mathematical foundation. A dataset of a petroleum refinery is used to train a RBF (Radial Basis Function) network using the conventional and the modified method and the performance of both will be evaluated. Furthermore, AAKR (Auto-Associative Kernel Regression) model is used to the same dataset. Finally, a comparison study of the developed models will be done for each of the performance metrics, as well as for the overall effectiveness in order to demonstrate the superiority of the proposed approach.

1 Introduction

Several works have been done in the last decades in order to improve the reliability of the industrial processes (Kristjanpoller *et al.* 2016), in controllers (Kettunen and Zhang 1997), actuators (Sharif and Grosvenor 1997) and sensors (Hines and Garvey 2006). Usually all equipments can be analysed using the processes measurements, but fault detection in sensors showed to have an extra complexity because these are the providers of the measurement. Sensors validation is an important subject in sensor based monitoring and control systems. In petroleum and natural gas industry, as in many others, reliable measurements are key to guarantee to the plant optimum and safe operation point and, in some cases, trustable billing. Therefore, many years of Research

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 791–797, 2020.
https://doi.org/10.1007/978-3-030-48021-9_87

and Development projects have been performed in order to improve diagnostic tools specially for the industry of petroleum and gas (Galotto *et al.* 2015). Many techniques using analytical redundancy in measurements have been widely tested and approved. The Auto-Associative Kernel Regression (AAKR) showed to have the best generalization (Galotto *et al.* 2006) in problems where the complexity of the process is hard to be modeled using the physical principles and there are enough amount of data to train, test and validate a memory based model. It is also easy to train or adapt to new data, since it is a memory based non-parametric technique. However, the main drawback of the AAKR probably is the high computational efforts necessary to evaluate each estimation.

The Radial Basis Function (RBF) Neural Network (Sundararajan et al. 1999) is one possible technique that could be applied in the same applications of AAKR. But the improvement of the training method is a key issue to allow RBF models reach similar performance metrics of AAKR models. That is because the main purpose of the auto-associative model is not to provide accurate estimations, but the possibility to estimate correctly even in fault conditions due to the existence of analytical redundancy, which is explored by the principle of auto-association. In other words, each sensor may be estimated using other sensors accordingly the existent relation among them.

This work presents the training of RBF models using performance metrics as targets, instead of only error. The proposed approach will be compared with the traditional training and with an AAKR model as a quality reference for the application of sensor fault detection.

2 Theoretical Background

2.1 Radial Basis Function Neural Network

Radial basis function (RBF) neural networks have an input layer, an intermediate layer and an output layer. The input layer represents the input space. The intermediate layer applies a nonlinear transformation of the input space, in order to get a hidden space with higher dimensions, through radial basis functions (e.g., Gaussian, multi quadratic or thin-plate spline). The output layer applies a linear transformation of the hidden space to generate the network outputs.

The response of the activation function of a RBF to a input vector x depends on the distance between x and the center of the function. The output of the activation function goes from 0 (for inputs far of the center) to 1 (for inputs near to the center).

Many methods have been proposed for the training of the RBF. Most of them are classified as hybrids due to they are composed in two parts: The first phase, the number of RBF and their parameters are established by non-supervisioned methods. In the second training phase, the weights of the neurons in the output layer are adjusted. Due to the outputs of the neurons in the intermediate layer is a linearly separable vector, the weights can be determined through linear models.

2.2 Performance Metrics

The quality of each model may be defined using some important performance metrics. Some of them are: mean square error and sensitivity. They can show if the model may improve the system response in normal conditions of operation and if the model can improve the response in fault conditions.

The mean square error in Eq. 1, compare the estimated output with the actual output without noise. In Eq. 1, n is the number of observations is the kth estimation of the signal j and ykj is the kth observation.

$$MSE_j^{norm} = \frac{\frac{1}{n}\sum_{k=1}^{n}\left(\hat{y}_{kj} - y_{kj}\right)^2}{\sigma_j^2} \tag{1}$$

This metric is well known and is useful to determine the accuracy of the model. This accuracy is fundamental to assure that the control system will keep the same response using the estimations of the model as a feedback signal in the control loop.

On the other hand, the sensitivity is a metric (Patton and Chen 1997; Hines and Garvey 2006; Galotto *et al.* 2008) to define the improvement of the model response in the presence of fault. Consider that is the k^{th} observation of x_i with artificially generated drift and is the k^{th} estimate observation of y_i to the same inputs with the drift. This drift may be additive or multiplicative changing the offset or the scale of the original signal. Thus, the denominator in Eq. 2 represents a simulated drift in the input i and the numerator is the propagation of the fault to the estimation of the signal j.

$$S_{ij} = \frac{\sum_{k=1}^{n}\left|\hat{y}_{kj}^{drift} - \hat{y}_{kj}\right|}{\sum_{k=1}^{n}\left|x_{ki}^{drift} - x_{ki}\right|} \tag{2}$$

Note the result is a matrix which shows the effect in the output j due to a fault in the input i. The optimum sensitivity should be zero so that the fault effects could be completely eliminated. When $i = j$ this is called robustness, and otherwise it is called spillover. As the models have a property to filter the noise, the same equation can be used, but a white noise is used instead of a drift.

3 The Training Method and Metrics

A cross-validation technique is frequently used to optimize the model based on the validation error. This is a good practice to minimize the modeling uncertainty, avoiding overfitting or biased solutions. However, the validation error is not the only desired target and the performance metrics could also be observed. Therefore, in order to overcome such limitation, the proposed approach includes directly these performance metrics in the weigh correction and stop criteria and as a cost function for the model optimization. The details of these procedures are presented in the next sections.

3.1 Proposed Cost Function

First, it has been proposed a preliminary cost function (L) depending on 4 designing parameters (α, β, γ and δ), which defines the weights for each performance metric respectively (*Accuracy*, *Robustness*, *Spillover* and *Filtering*), as shown in Eq. 3.

$$L = \alpha.Accuracy + \beta.Robustness + \gamma.Spillover + \delta.Filtering \quad (3)$$

This cost function has always non-negative values and the same goal for the best solutions independent on the designing parameters. The difference will be the priority given to each metric changing the minimum point during the modeling process.

3.2 Test Cases

Four comparative cases have been defined based on the cost function of Eq. 3. The solution will be changed depending on the priority given by the chosen design parameters. And the cases have been named accordingly the prioritized (focused) metric, as follows:

- AF: Accuracy Focused (original training method);
- RF: Robustness Focused;
- ARF: Accuracy and Robustness Focused;
- ARSFF: Accuracy, Robustness, Spillover and Filtering equally Focused;

In Table 1 the weights of each test case considering the performance metrics are presented.

Table 1. Weights of the considered test cases.

Case	Alfa α	Beta β	Gama γ	Delta δ
AF	1	0	0	0
RF	0	1	0	0
ARF	0,5	0,5	0	0
ARSFF	0,25	0,25	0,25	0,25

3.3 The Training Methodology

In this work the training algorithm of RBF was modified in order to use the cost function L, instead of the mean square error (MSE), as the cost function to be minimized in the training. In order to get the cost function with best performance, it was necessary to vary the values of spread (between 0.1 and 100) and the number to neurons (between 1 and $N - 1$, being N the number of inputs).

3.4 Real Data to Test the Models

Data from 35 sensors of a Liquefied Petroleum Gas pump of the Gas Treatment Unit of Petrobras were used, and this unit receives gas and condensate from three offshore platforms. The processing of this results in LPG (liquefied petroleum gas) and C5+ (also corresponding to the liquid gas fraction). The signals type are Temperature, Pressure, Flow, Rotation, Powers, Axial Displacement and Number of Starts. The received data were processed and grouped according to the linear correlation coefficient matrix and its function in the process. Data processing and sensor grouping are not part of this work, so they will not be detailed here. The function of each group and the number of instruments are presented in Table 2.

Table 2. Sensor groups of the test data

Group ID	Function	Number of sensors
1	Ungrouped or constant	4
2	Water flow and pressure	9
3	Bearings temperature	6
4	Windings temperature	5
5	Vibration	8
6	Lubricant oil pressure	3

4 Results with the Real Data

The models were trained for the six predefined groups (Table 2), using all the considered training cases in Table 1. Also, an optimized kernel (AAKR) model is presented as a reference to the expected goal to the performance metrics. Figure 1 shows the achieved Accuracy for all groups with the five training cases. Accuracy of AAKR is the large back bar for all groups.

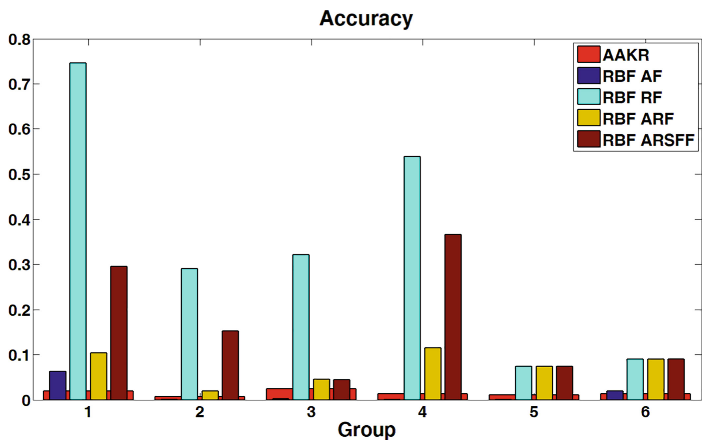


Fig. 1. Comparative between Kernel and RBF

It is noted that the RBF AF case is the best to reach similar or better accuracy results. However, Fig. 2 shows that RBF AF was one of the worst cases for most of groups, regarding the robustness and spillover performances. The RBF models showed to have better sensitivity when compared to AAKR. The RBF AF case is the usual training that may result in better accuracy and worst sensitivity as expected, since the other metrics are not observed during this process.

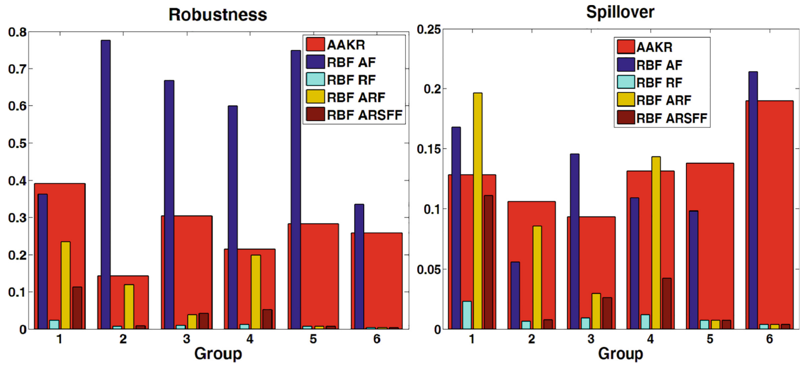


Fig. 2. Comparative sensitivity (robustness and spillover) between Kernel and RBF

The similar conclusion was observed with filtering metric. The average results for all groups are summarized in Table 3.

Table 3. Average results for all groups

	Acc.	Rob.	Spill.	Filtering	L	Neurons	Spread
AF	0,014	0,581	0,132	0,210	0,014	4,8	45,8
RF	0,343	0,010	0,010	0,011	0,010	1,0	3,2
ARF	0,075	0,100	0,077	0,082	0,087	2,2	3,9
ARSFF	0,170	0,038	0,033	0,036	0,069	1,3	3,7

The Fig. 3 shows graphically the difference of performance of all tests.

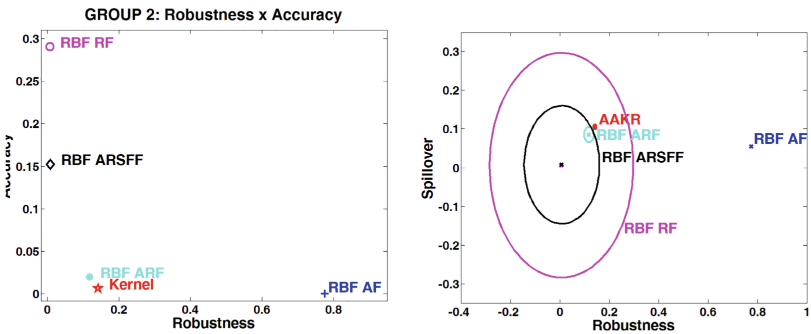


Fig. 3. Robustness x spillover map for each case with circle area as accuracy

The left chart shows the balance between Accuracy and Robustness. The right chart shows the Robustness and Spillover relation including a circle for each case with its area equal to the accuracy. Clearly, the AF case reach the best Accuracy, but the worst Robustness, while the RF case is the opposite. ARF approach was that with the nearest performance compared with the AARK model.

5 Conclusion

A novel modeling procedure for RBF neural networks based on different performance metrics have been proposed, observing more metrics of interest for sensors monitoring problems. The method has been compared with the traditional AAKR and the use of a combination of Accuracy and Robustness focus allowed the RBF to reach similar features of AAKR.

References

Galotto, L., et al.: Sensor compensation in motor drives using kernel regression. In: 2007 IEEE International Electric Machines & Drives Conference, pp. 229–234. IEEE (2006)

Galotto, L., et al.: Improvement of fault detection with partial auto-associative models. In: Kumar, U., Parida, A., Rao, R.B.K.N. (eds.) Monitoring and Diagnostic Engineering Management (COMADEM 2006) (2008)

Galotto, L., et al.: Data based tools for sensors continuous monitoring in industry applications. In: 2015 IEEE 24th International Symposium on Industrial Electronics (ISIE), pp. 600–605 IEEE (2015)

Hines, J.W., Garvey, D.: Development and application of fault detectability performance metrics for instrument calibration verification and anomaly detection. J. Pattern Recogn. Res. **1**(1), 2–15 (2006)

Kettunen, M., Zhang, P.: An embedded fault detection, isolation and accommodation system in a model predictive controller for an industrial benchmark process. Comput. Chem. Eng. **32**(12), 2966–2985 (1997)

Kristjanpoller, F., et al.: Reliability assessment methodology for multiproduct and flexible industrial process. In: Risk, Reliability and Safety: Innovating Theory and Practice, pp. 1101–1107. CRC Press (2016)

Patton, R., Chen, J.: Observer-based fault detection and isolation: robustness and applications. Control Eng. Pract. **5**(5), 671–682 (1997)

Sharif, M., Grosvenor, R.: Fault diagnosis in industrial control valves and actuators. In: IMTC/98 Conference Proceedings. IEEE Instrumentation and Measurement Technology Conference. Where Instrumentation is Going (Cat. No. 98CH36222), pp. 770–778. IEEE (1997)

Sundararajan, N., Saratchandran, P., Wei, L.Y.: Radial basis function neural networks with sequential learning: MRAN and its applications. World Scientific (1999)

Part IX: Asset Innovation, Business Models, and Co-value Creation

A Conceptual Framework for Innovation Assessment in Firms Based on 4Cs

Leila Beig(✉)

Business School, University of Stavanger, Stavanger, Norway
l.beig@stud.uis.no

Abstract. Companies are encountering competitive challenges which makes them be efficiently responsive to current markets while effectively preparing for new markets and opportunities. In this circumstance, innovation plays a critical role to heighten competitiveness capacity of a firm. But the question is "how firms' leaders can ensure that their firms are innovative, and they are taking the appropriate path which ends up with being uniquely innovative?" One way is to consider innovation measurement main factors which can provide a roadmap to define innovation related activities. Thus, the main purpose of this paper is to find the factors that should be measured and improved by managers to make sure that their organizations are innovative. In this paper, these factors will be identified and integrated into a holistic conceptual framework regarding four Cs so called by the author "innovation capacity, innovation capability, innovation competence and consequences of innovation".

1 Introduction

In the twenty first century, more than any other time in the firm's history, companies are under the global competitive pressures in a context that is widely dynamic and consequently makes them encounter a huge number of unpredictable changes. In this situation, firms are required to be efficiently responsive to current markets while effectively preparing for new markets and opportunities (Naman and Slevin 1993). (Tushman and Nadler 1986) believe that innovation plays a critical role to heighten competitiveness capacity of a firm. But the question is "how firms' leaders can ensure that their firms are innovative, and they are taking the appropriate path towards innovation?" Gorton believes that one way to find the answer is to measure innovation (Gorton 2000). (Saunila and Ukko 2012) showed that innovation measurement can contribute to a significantly better understanding of innovation when the measurement has been conducted properly.

E. Goldratt, the father of the Theory of Constraints and the writer of the book "the goal", specifically addresses the importance of metrics by this popular sentence: "tell me how you measure me I will tell you how I will behave" (Essmann 2009). This sentence implies that the way firms behave is to some extent dependent on the evaluation mechanism and factors involved. The main goal of this paper is to propose a conceptual framework for innovation assessment by looking at this concept from a more comprehensive and holistic point of view. The main features of conceptual frameworks include: "A conceptual framework is not merely a collection of concepts

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 801–811, 2020.
https://doi.org/10.1007/978-3-030-48021-9_88

but, rather, a construct in which each concept plays an integral role. It provides not a causal/analytical setting but, rather, an interpretative approach to a main concept. Rather than offering a theoretical explanation, as do quantitative models, conceptual frameworks provide understanding. Conceptual frameworks are indeterminist in nature and therefore do not enable us to predict an outcome" to (Jabareen 2009).

According to (Jabareen 2009) the procedure of conceptual framework encompasses of eight phases: mapping the selected data sources, extensive reading and categorizing of the selected data, identifying and naming concepts, deconstructing and categorizing the concepts, integrating concepts, synthesis, resynthesize, and making it all make sense, validating the conceptual framework and rethinking the conceptual framework. In this paper, except the last two phase the other ones have been conducted. Presenting the model at a conference, a seminar, or some other type of academic communities provides an excellent opportunity for researchers to discuss and get feedback. A theoretical framework representing a phenomenon will always be dynamic and may be revised according to new comments, literature, and so on.

For this purpose, the paper has been structured as following. In the next section the innovation assessment factors proposed in 14 selected papers will be reviewed briefly. Totally, 57 papers were reviewed and out of them, 14 were selected as the main references of this research. The criteria of choosing the papers were firstly, the availability of details of each component in the framework and secondly the citations and H score of the journals the papers were published in. In terms of time, papers published between the years 2000 to 2017 were reviewed.

Afterwards, the extracted factors were clustered into four categories including Capacity, Capability, Competence and Consequence of innovation so called 4Cs by the author. In some cases, several factors had the same meaning, but different terms were used by various authors, therefore one concept has been considered by the author as a representative of all the terms used. At the end, a conceptual framework has been proposed and building blocks have been explained. It should be mentioned that the main concentration of this paper is on "*what*" should be assessed not on "*how*" it is going to be evaluated. The study has the potential to contribute to better understanding of innovation measurement by taking the research issue one step further. The framework can also be used by managers to examine their innovation activities and become aware of the main factors which contribute to being innovative and identify areas for future improvement planning.

2 Literature Review

In these section, 14 related papers will be reviewed in terms of factors related to innovation assessment. In a couple of papers, a model or framework has not been proposed but the relationship between two or three entities have been examined. Those papers have been considered in this study as well, because the definition and explanation of factors constructing the main entities are useful and contribute to form the final conceptual framework.

According to (Adams et al. 2006), for innovation management measurement seven categories should be considered. Resources such as finance, human and physical resources, and new ideas for innovation are included in the first category. Knowledge management is the second category which shows an organization's ability to detect, acquire and apply external knowledge that can be useful to a firm's successful operation. Innovation strategy is another category encompassing strategic orientation and strategic leadership which is associated with resource allocation decisions to meet an organization's objectives. The forth category is organizational culture and structure. There are four more categories in this model including portfolio management, commercialization process, project efficiency as well as communications and collaboration tools (Adams et al. 2006).

In another study, for measuring innovation performance, six categories including external knowledge, structures and systems, regeneration, leadership, work well-being and employee skills have been proposed to (Saunila 2017). External knowledge refers to the exploitation of external networks and it measures external links such as collaboration with suppliers and customers. The structures and systems show internal processes and structures that support and reflect continuous innovation. Regeneration is about the organization's ability to learn from experience and to use that experience to create and develop innovations. Overall atmosphere of the organization that supports and motivates innovation, and a leadership culture that facilitates innovation has been also viewed as a critical factor in this model. In addition, Work well-being which refers to work climate for innovation development and know-how/expertise and skills of employees are among the other factors (Saunila 2017).

Innovation capability measurement has also been the focus of a research done by (Saunila and Ukko 2012). They categorized the factors into innovation potential including factors that affect the present state of innovation capability and reflect the potential of an organizations in producing innovative ideas, products and services; innovation process which refers to systems and activities that assist organizations to utilize their innovation potential and therefore enable innovations and innovation activities' results which can be product, service and process innovations (Saunila and Ukko 2012).

In a similar study by (Albaladejo and Romijn 2000), two main factors have been considered for measuring innovation capability in UK firms. They took a resource-based view and categorized the factors into internal sources which encompasses professional background of founder/manager(s), skills of workforce, and technological effort and external sources and external sources referring to networking and its proximity advantages as well as nature and extent of institutional support received.

Three other capabilities were introduced by (Momeni et al. 2015) for evaluating innovation capability. The first capability is structural capability which can be determined by managerial capacity, cultural capacity, communicative capacity and organizational knowledge capacity. The second one is personnel capability including factors such as idea generation capacity, opportunity detection capacity, individual knowledge capacity. Operational capability is the last one which is categorized into technological capacity and support capacity. According to this study technological capacity of a firm refers to "the ability of a firm in the use of technology and combination and recombination of parts and constituents, and the relationship among procedures, processes and techniques" and the supportive capacity of an organization is related to logistical and supportive process and the work place situation of an organization (Laforet 2011).

(Chang et al. 2012) studied radical innovation capabilities and proposed four other capabilities for measuring innovation capability. They include openness, integration, autonomy, and experimentation capability. Openness capability refers to activities such as participating in industrial networks like standard organizations and industrial forums. Integration capability is about applying the knowledge gained in previous projects to new projects, encouraging cross-functional learning and fertilization, upgrading and integrating technology capabilities, new product development and marketing, while autonomy capability is to renew product portfolios frequently, update necessary technological and market information, and to implement innovation strategy. Lastly, experimentation capability means adopting new ideas and developing them as reliable products, and commercializing proven concepts into market, developing methods and tools to improve R&D (Chang et al. 2012).

In another study by (Lawson and Samson 2001), besides the previous mentioned factors such as organizational structure, reward systems, supportive culture of ambiguity tolerance, empowered employees, and communication, four other factors have also been taken into account. 1- vision and strategy, which determine the configuration of resources, products, processes and systems that firms adopt to deal with the uncertainty existing in their environment. This factor was mentioned as a considerable factor in (Adams et al. 2006) as well. 2- harnessing the competence base which refers to ability to correctly and effectively direct resources to where they are required; 3- organizational intelligence which focuses on learning about customers competitors and the last one, creativity and idea management.

(Laforet 2011) looked at organizational innovation in a more comprehensive way from the first stages which drives innovation to the last steps which ends up with outcomes. They believe market environment and a firm's strategic posture can motivate innovation. But how innovative an organization can be, depends on a couple of factors including innovation capacity of a company which refers to availability of resources such as financial and skilled workforce; and collaborative structures and processes. In addition, they emphasized on the importance of innovation orientation of a company such as risk-taking attitude, and willingness to learn which can provide an appropriate atmosphere for innovation. Innovation strategy as a part of overall business strategy and outcomes of innovation in the form of business financial performance are also among factors being considered in this framework (Laforet 2011).

Similar to (Laforet 2011), willingness to take risks and exchange ideas, keen interest about expressing and exchanging information and sharing knowledge so called innovation orientation in (Laforet 2011) are among the factors which have been proposed in (Wana et al. 2005). They also considered organizational resources which address costs of instituting innovations and innovation funds as an important determinant of firm. In addition, communication channels, decentralized structure as well as a supporting culture which rewards innovation are also among other crucial indicators in this model.

(Carayannis 2007) share the same point of view as (Laforet 2011) and look at the whole picture. They provided a model called composite innovation index which consist of three main elements as follows. 1- posture which is an organization's position and addresses organizational culture, technological and relational knowledge, as well as market competitive position of a company. 2- Propensity is a firm's ability to capitalize on its posture. In a better word, it is an intangible reflection of process, routines and capabilities established within a firm. 3- Performance which is lasting result of innovation. It includes outcomes such as patents and new products, or revenues contributed by new products or impacts such as becoming a top innovator in a society (Ganter 2013).

In another framework by (Neely et al. 2001), more focus is on impacts and outcomes of innovation. For instance, business performance including factors such as return on investment, market share, competitive position versus direct competitors is one of the main building blocks. While lower costs, enhancement of existing products, introducing new products, better customer service, and value to customers have been considered as the outcomes of innovation, in this framework capacity to innovate including innovation culture, internal process capability, capability to understand market and technology trend; as well as external contextual environment encompassing public policies and infrastructure are among supportive factors (Neely et al. 2001).

In another study by (Jiménez-Jiménez and Sanz-Valle 2011), showing how innovation can lead to performance improvement through organizational learning was the main concentration. In this paper, learning refers to knowledge acquisition, knowledge distribution, knowledge interpretation, and organizational memory. The authors believe that through learning, different types of innovation such as product innovation, process innovation, as well as administrative innovation will be realized. Innovation itself has the potential to lead to better performance of a firm. Improved performance can be observed through quality product, better products' image, market share, profitability and productivity, as well as less personnel turnover and absenteeism (Jiménez-Jiménez and Sanz-Valle 2011).

The main idea of the paper written by (Rajapathirana and Hui 2017) is to find the relationship among three elements, innovation capability, innovation types, and firm performance. In their research, they showed that innovation capability which refers to organizational culture and using knowledge from different sources while involvement of workers and customers in the innovation management process can lead to various types of innovation such as marketing, product, organizational and process innovation which in turn can finally end up with better innovation, market and financial performance. While innovation performance can be measured by quality of new product or services, technological competitiveness, speed of introduce new products or service, and novelty of new product or service; market performance can be estimated by market share, customer satisfaction and total sales. In addition, return on investment, return on sales, and profitability are among sub factors for financial performance evaluation (Rajapathirana and Hui 2017). (Branzei and Vertinsky 2006) took a strategic view on product innovation and proposed a strategic pathway to bridge between product innovation capabilities and development strategies. In this study, acquisition capabilities, assimilation capabilities, transformation capabilities and deployment capabilities constitute innovation capabilitie. Acquisition capabilities are related to active monitoring different external sources of knowledge, looking for various partnerships and learning from different types of collaborators, including customers, suppliers, universities, and public support agencies. From the other hand, assimilation capabilities focus on redesigning the layout, updating prior operating routines, upgrading technology, and maintain production work force. While transformation capabilities are about creative cross distribution of relevant skills and resources across different areas of expertise, deployment capabilities include effective commercialization of innovative products. Three development strategies were also considered in this paper 1-Human capital development strategies, 2-Product development strategies, 3-market development strategies and 4-process development strategies (Branzei and Vertinsky 2006). The authors offered pathways between each capability and development strategy.

3 Innovation Assessment Conceptual Framework

In this section the conceptual model presented by author has been presented. for this purpose, based on the related concepts in literature review four categories has been proposed. The categories include, innovation capacity, innovation capability, innovation competence and innovation consequences. By Innovation capacity author means current state of a firm regarding the resources it holds and circulates to be innovative. Innovative vision and strategy, organizational culture, collaborative structure, leadership, skillful employees, innovation orientation (willingness to take risks, willingness to exchange ideas), openness, financial resources, and extent of institutional support are considered in this category.

As it has been showed in (Laforet 2011), it can be said that if an organization has a collaborative structure, it is more likely that it will be more innovative in comparison to a firm that does not have a collaborative structure. It is the same for the other factors.

The second component of the model is capability which refers to the ability of a firm in terms of innovation. In fact, this element is the answer to the following question "what are the important factors that make a firm able to use its capacity for being more innovative?". Through capability the author tended to look at the processes that facilitate the innovation process inside a firm and make them more capable of applying the resources in a better way and end up with being more successful in innovation. By integrating the related factors mentioned in the previous section, factors such as commercialization (including idea generation, opportunity detection, …), knowledge management (including knowledge acquisition, assimilation, transformation and deployment), regeneration/learning, innovation processes, management of technology, organizational intelligence, integration, autonomy, experimentation and learning has been placed in this category. As it has been illustrated in Fig. 1, spiral type of relationship has been proposed between innovation capacity and innovation capability. It means that capacity and capability are enforcing each other through time and they are transforming to each other. In a better word, while applying for example knowledge management in a firm, learning capability of personnel might increase and because of this change, one more skill in a specific field can be added to the innovation capacity of the firm and by this increased capacity, the firm will be able to do more innovation related activities and therefore, a new capability will emerge. It is similar to Nonaka's idea (Nonakaet al. 2000) about the relationship between tacit and explicit knowledge.

It can be said that not only all the factors inside the capacity and capability component might enforce and influence each other but also capacity and capability factors are transforming to each other dynamically in a way that sometimes we cannot make any border between them and differentiate capacity and capabilities. That is the beauty

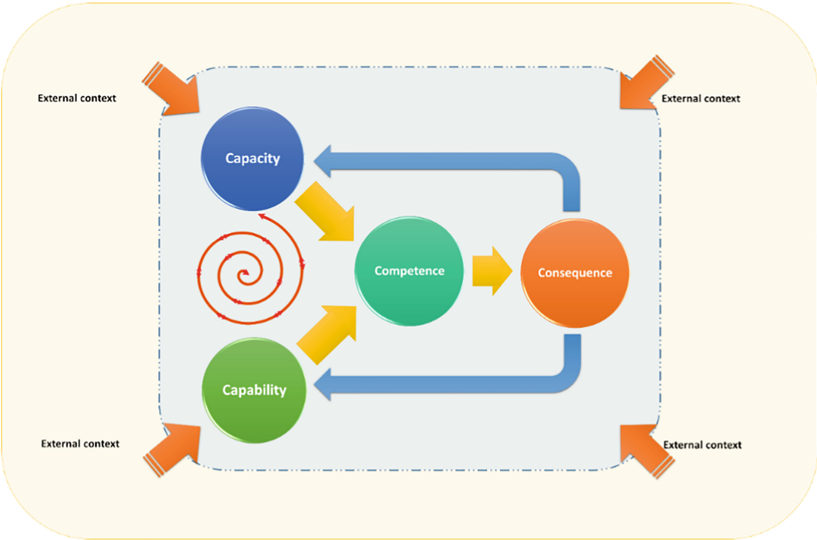


Fig. 1. Innovation assessment conceptual framework

of system dynamics, when one looks at a phenomenon (in this case innovation) as a system which is continuously changing as non-stop emerging patterns.

The third element of the framework is competence which shows what a firm acquires by applying its capabilities through capacities it owns. Different type of innovation such as product innovation, marketing innovation, organizational innovation and process innovation are considered in this part of the model. It shows that innovation capability and capacity can lead a firm to be competent in different types of innovation. But which factor in capacity and capability components contributes more in leading for example product innovation is not the main purpose of this paper. The mentioned question is important, and in some cases, researchers tried to find an answer to such questions.

The forth construct of the model is consequences of the innovation which means that what are the end results of innovation form financial, human resource and even society point of view. They can be categorized into business performance, marketing performance and human performance. As the framework shows these consequences feed the capacity and capability part. The important point in this loop is strategic decisions that managers make about what they are going to focus on and how they are going to invest on different projects. Because their decisions at this specific point, in terms of innovation, can make a positive or negative difference in a firm or even no difference from the previous state. We should be aware that a firm is not working in isolation, the environment factors, if we call them external context, are continuously changing and their alterations indeed influence a firm's behavior. The capability of monitoring the external context, opportunity detection and learning to change can be very useful in responding to environment dynamics. In previous studies theses kind of factors were mentioned in only two papers. Since they are important they have been kept by author in the proposed framework. The borders of the model are in the form of dashed lines which means that firms cannot prevent resource flow from outside into their firms and from inside out and open innovation is happening around organizations without prenotification.

To wrap up, for assessing innovation in a firm, looking at the whole picture and examining the inputs and process including capacity and capability, and competence including different types of innovation and the final consequences is recommended in order to make sure that the innovation engine is working properly. If we just look at the input and process and evaluate these two parts of the puzzle it means that we have necessary inputs, but it is not enough to conclude that we have product innovation for instance. Therefore, the competent and consequence part should be checked as well. Last not least, how to allocate the resources by applying the results from consequence part shows how a firm moves from one innovation state to a better state. It is worth mentioning that environment drivers most of time drive a firm towards innovations, therefore a firm should not forget to track these factors as well. Figure 1 illustrates the proposed conceptual framework.

In, the components of the framework and their definition as well as their frequency in literature have been presented. The culture has the most frequency followed by organization structure, internal resources and innovation knowledge management process. It can be concluded that these factors play a critical role in innovation management assessment (Table 1).

Table 1. Conceptual model components

Main category	Components	Definition	Reference(s)	Frequency
Innovation Capability	Commercialization capability/ integration capability/project management capability	Commercialization is the successful introduction of new products and services into markets. it includes issues such as market analysis and monitoring, reaching the customer and market planning, sales and distribution. Fine integration and alignment of corporate R&D units and existing lines of business is crucial to the commercialization of innovation	(Adams, et al., 2006) (Yuan-Chieh Chang, Huo-Tsan Chang, Hui-Ru Chi, Ming-Huei Che, Li-Ling Deng , 2012)	2
	Knowledge management (KM) capability/ Innovation process capability	KM ability is concerned with obtaining and communicating ideas and information that underlie innovation competencies, and includes idea generation, absorptive capacity which is the result of openness capability and networking. KM covers the management of explicit and implicit knowledge held by the organization. Detecting opportunities, and generating ideas can be a result of effective KM.	(Adams, et al., 2006) (Momeni, et al., 2015) (Yuan-Chieh Chang, Huo-Tsan Chang, Hui-Ru Chi, Ming-Huei Che, Li-Ling Deng , 2012) (Oana Branzei, Ilan Vertinsky, 2006)	4
	Organizational learning capability (Regeneration/experimentation)	The organization's ability to learn from experience and to use that experience to create and develop innovations. It comprises four sub processes including knowledge/idea acquisition, knowledge distribution, knowledge interpretation and, organizational memory.	(Saunila, 2017) (Daniel Jiménez-Jiménez *, Raquel Sanz-Valle , 2011) (Daniel Jiménez-Jiménez *, Raquel Sanz-Valle , 2011)	3
	Technological capability	Technological capability of a firm is defined as the ability of a firm in the use of technology and combination and recombination of parts and constituents, and the relationship among constituents, procedures, processes and techniques.	(Momeni, et al., 2015) (Lawson & Samson, 2001)	2
	Organizational intelligence capability	Organizational intelligence has been defined as "the capability to process, interpret, encode, manipulate and access information in a purposeful, goal-directed manner, so it can increase its adaptive potential in the environment in which it operates". Since knowledge and ideas are primary imports into the innovation process, intelligent firms can use this information to reduce the inherent uncertainty and ambiguity of innovation	(Lawson & Samson, 2001)	1
	Portfolio management capability	Allocating resources to projects to obtain the optimal balance in the product development portfolio, that is, arriving at a portfolio that optimizes the trade-off between returns and risks.	(Adams, et al., 2006)	1

(*Continued*)

Table 1. (*Continued*)

Main category	Components	Definition	Reference(s)	Frequency
Innovation Capacity	Innovation vision and strategy	alignment: Innovative firms are able to link their core technology strategies, with innovation strategy and business strategy. This alignment generates a powerful mechanism for competitive advantage.	(Lawson & Samson, 2001)	1
	Organizational culture/ leadership and innovation orientation	An organizational culture refers to overall atmosphere of the organization that supports and motivates innovation. An organization culture equipped to support risk-taking, freedom and self-management fosters individuality, as well as creativity and tolerance of failure. A leadership culture that facilitates innovation and willingness to take risks exchange ideas.	(Adams, et al., 2006) (Saunila, 2017) (Laforet, 2011) (Wana, et al., 2005) (Momeni, et al., 2015) (Lawson & Samson, 2001) (Yuan-Chieh Chang, Huo-Tsan Chang, Hui-Ru Chi, Ming-Huei Che, Li-Ling Deng , 2012)	7
	Collaborative/decentralized/ organic structure	Successful innovation requires an optimal overall formal business structure. The more permeable and organic the structure, the greater the potential for innovative ideas to spring.	(Adams, et al., 2006) (Saunila, 2017) (Laforet, 2011) (Wana, et al., 2005) (Lawson & Samson, 2001)	5
	Organizational resources/ inputs management/ Internal sources/ financial resources	The organizational resources include but not limited to number of personnel employed in R&D roles /number of people committed to the innovation task, total expenditure on R&D as a proportion of sales or revenues, skills, experience and education of human resource.	(Adams, et al., 2006) (Saunila, 2017) (Wana, et al., 2005) (Albaladejo & Romijn, 2000)	4
Innovation Competence	Product/service innovation	Product/service innovation is "introducing new product or service with the significantly improved performance characteristics to full fill the key customer needs better than the existing product" (OECD, 2005).	(Andy Neely, Roberto Filippini, Cipriano Forza, Andrea Vinelli, Jasper Hii, , 2001) (Daniel Jiménez-Jiménez ⁎, Raquel Sanz-Valle , 2011) (Rajapathirana, R. P. J., Hui, Y. , 2017)	3
	Process innovation	Process innovation is the "implementation of new or significantly improved production or delivery methods" (OECD, 2005).		
	Organizational innovation	Organizational innovation is "implementation of a new organizational method in the firm's business practice, organization or external relations".		
	Marketing innovation	Marketing innovation is "introducing new marketing methods involving significant changes in product design, product placement, and product promotion or pricing (OECD, 2005).		
Innovation Consequences	Business performance	Business performance includes factors such as return on investment, firm's competitive position, values to customers and lower costs.	(Andy Neely, Roberto Filippini, Cipriano Forza, Andrea Vinelli, Jasper Hii, , 2001) (Daniel Jiménez-Jiménez ⁎, Raquel Sanz-Valle , 2011) (Rajapathirana, R. P. J., Hui, Y. , 2017)	3
	Marketing performance	Marketing performance encompasses market share, profitability and productivity.	(Rajapathirana, R. P. J., Hui, Y. , 2017)	1
	Human performance	Human performance can be measured through two factors personnel turnover and absenteeism.	(Daniel Jiménez-Jiménez ⁎, Raquel Sanz-Valle , 2011)	1
External context	Innovation drivers/external sources	The factors that play a role in forming external context of an organization include technological change, changing of market competition at home and abroad, interactions in the value chain with customers, suppliers and service providers, and also external private and public agents including financial institutions and industry associations.	(Laforet, 2011) (Albaladejo & Romijn, 2000) (Andy Neely, Roberto Filippini, Cipriano Forza, Andrea Vinelli, Jasper Hii, , 2001)	3

4 Conclusion

In this paper, the author attempted to look at a firm from different angles, take a relatively comprehensive view and propose a multi-dimensional conceptual model for innovation assessment in a firm.

Literature review and previous studies shows that for innovation evaluation a couple of studies concentrated on input side of the phenomenon including factors such as R&D expenditures, financial resources, skillful human resource and the similar factors. In contrast, the output side was the main concentration of other studies. They looked at different type of innovations which various organizations offered. On the other hand, number of studies tried to look at the process side by looking at the activities related to innovation management. The author believes that these activities are like wheels of a bike if they work properly the bike moves forward otherwise it either woks, but no movement is observed, or it doesn't work appropriately in terms of being innovative. Most of what is happening inside a firm is similar to tacit knowledge that is difficult to be transferred from one firm to another easily and cannot be measured simply. Inputs and outputs can be measured by several indices but assessing the process part is very difficult because each organization might have its unique way of innovation management based on its context which has been formed through time. Some researchers believe that there isn't any single way of innovation that can be prescribed for all organizations. the importance of capacity and capability conversion and strategic decision that managers for applying the loop from consequences to capacity and capability and continuing the innovation journey has been considered in the proposed model as well.

By concentrating on the important factors for evaluating innovation in a firm, the purpose of this paper is to make the managers aware of the factors that should be checked as a routine to ensure that their firm is moving toward the right path. The proposed model has not been tested and validated, therefore, to this point it is only author's thoughts based on literature review and unfortunately the comments of the expert of this area has not been inserted yet. The relationship among factors inside each category has not been examined and was not the scope of this paper but can be considered as the future works.

Acknowledgement. I would like to thank Professor Rune Dahl Fitjar and Dr. Marte Cecilie Wilhelmsen Solheim for helpful comments on this paper.

References

Adams, R., Bessant, J., Phelps, R.: Innovation management measurement: a review. Int. J. Manag. Rev. **8**(1), 21–47 (2006)

Chang, Y.-C., Chang, H.-T., Chi, H.-R., Che, M.-H., Deng, L.-L.: How do established firms improve radical innovation performance? The organizational capabilities view. Technovation **32**, 441–451 (2012)

Albaladejo, M., Romijn, H.A.: Determinants of innovation capability in small UK firms: an empirical analysis. In: ECIS Working Paper Series, vol. 200013. Eindhoven Centre for Innovation Studies, Eindhoven (2000)

Neely, A., Filippini, R., Forza, C., Vinelli, A., Hii, J.: A framework for analysing business performance, firm innovation and related contextual factors: perceptions of managers and policy makers in two European regions. Integr. Manuf. Syst. **12**(2), 114–124 (2001)

Carayannis, E.: Measuring firm innovationess: towards a composite innovation index built on firm innovation posture, propensity and performance attributes. Int. J. Innov. Regional Dev. **1**, 90–107 (2007)

Jiménez-Jiménez, D., Sanz-Valle, R.: Innovation, organizational learning, and performance. J. Bus. Res. Issue **64**, 408–417 (2011)

Essmann, H.E.: Toward innovation capability maturity. Stellenbosch University, South Africa (2009)

Ganter, A., Hecker, A.: Deciphering antecedents of organizational innovation. J. Bus. Res. **66**, 575–584 (2013)

Gorton, M.: Overcoming the structure – agency divide in small business research. Int. J. Entrepreneurial Behav. Res. **6**(5), 276–292 (2000)

Jabareen, Y.: Building a conceptual framework: philosophy, definitions, and procedure. Int. J. Qual. Methods **8**(4), 49–62 (2009)

Laforet, S.: A framework of organisational innovation and outcomes in SMEs. Int. J. Entrepreneurial Behav. Res. **17**(4), 380–408 (2011)
Lawson, B., Samson, D.: Developing innovation capability in organizations: a dynamic capabilities approach. Int. J. Innov. Manag. **5**(3), 377–400 (2001)
Momeni, M., Nielsen, S.B., Kafash, M.H.: Determination of innovation capability of organizations: qualitative meta synthesis and delphi method, Denmark: In Proceedings of RESER2015 - Innovative Services in the 21st Century (2015)
Naman, J.L., Slevin, D.P.: Entrepreneurship and the concept of fit: a model and empirical tests. Strateg. Manag. J. **14**(2), 137–153 (1993)
Nonaka, I., Toyama, R., Konno, N.: SECI, Ba and leadership: a unified model of dynamic knowledge creation. Long Range Plann. **33**, 5–34 (2000)
Branzei, O., Vertinsky, I.: Strategic pathways to product innovation capabilities in SMEs. J. Bus. Ventur. **21**, 75–105 (2006)
OECD: Oslo manual proposed guidelines for collecting and interpreting technological. Innovation data (Paris) of radical innovation: Insights from pharmaceuticals. J. Mark. **67**, 82–102 (2005)
Rajapathirana, R.P.J., Hui, Y.: Relationship between innovation capability, innovation type, and firm performance. J. Innov. Knowl. **3**, 44–55 (2017)
Saunila, M.: Understanding innovation performance measurement in SMEs. Meas. Bus. Excell. **21**(1), 1–16 (2017)
Saunila, M., Ukko, J.: A conceptual framework for the measurement of innovation. Baltic J. Manag. **4**(7), 355–375 (2012)
Tushman, M., Nadler, D.: Organizing for innovation. Calif. Manag. Rev. **28**(3), 74–92 (1986)
Wana, D., Ongb, C.H., Leec, F.: Determinants of firm innovation in Singapore. Technovation **25**, 261–268 (2005)

Open Innovation in Mature Industries an Example from Maritime Industry in Norway

Tatiana Iakovleva[1(✉)], Marina Solesvik[2], and Odd Jarl Borch[3]

[1] Stavanger Business School, University of Stavanger, Stavanger, Norway
tatiana.a.iakovleva@uis.no
[2] Maritime Innovation and Management, Nord University Business School, Nord University, Bodø, Norway
marina.solesvik@hvl.no
[3] Nord University Business School, Nord University, Bodø, Norway
odd.j.borch@nord.no

Abstract. Open innovation has received increasingly attention in scientific research (Chesbrough 2002; van de Vrande et al. 2009). It showed to be a fruitful strategy to employ in highly dynamic and changing environment (Chesbrough 2003a, b). Since its emergence, evidence to support open innovation model was taken mainly from the so-called 'high technology' industries, such as computers, information technology, and pharmaceuticals. However, more resent work Chesbrough and Growther (2006) confirmed the relevance of this approach in more mature industries. However, little is known so far under which conditions open innovation serves as appropriate strategy in such mature industries. This paper investigates innovation process in maritime industry, which is characterised by long-term trust-based relationships between involved partners, and substantial financial investment. Using example from a shipping company in Norway we show how the innovative action of the firm may depend on the combined influence of entrepreneurial orientation (Lumpkin and Dess 1996; Rauch et al. 2009) within the firm and cooperative links towards knowledge providers. We investigate innovation process that led to construction of a ship that was specifically designed to operate under the harsh Arctic conditions. Findings indicate that open innovation approach is particularly suitable for innovation process in mature industries when product requirements are dictated by extra demanding conditions. We contribute to the domain of open innovation by specifying that firm entrepreneurial orientation, collaborative links with customer and partner are particularly suitable when the level of uncertainty is high.

1 Introduction

It is widely acknowledged that stimulation of innovative activity is crucial for the competitive advantage and growth of both companies and regions. In most countries, a broad number of policy instruments that stimulate R&D activities, science and technology are at hand (Jensen et al. 2007). It is now widely recognized that firms can

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 812–819, 2020.
https://doi.org/10.1007/978-3-030-48021-9_89

considerably increase their innovation capacity if the act over the firm boundaries (van de Vrande et al. 2009), the concept known as open innovation. Open innovation can be defined as the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and to expand the markets for external use of innovation (Chesbrough et al. 2006). At the heart of the open innovation model is the recognition that today, competitive advantage often comes from inbound as well as from outbound connections. Inbound connections is the practice of leveraging the discoveries of others: companies need not and indeed should not rely exclusively on their own R&D. Outbound open innovation suggests that rather than relying entirely on internal paths to market, companies can look for external organizations with business models that are better suited to commercialize a given technology (Chesbrough 2002).

Since its emergence, evidence to support open innovation model was taken mainly from the so-called 'high technology' industries, such as computers, information technology, and pharmaceuticals (Chesbrough 2003). However, later work of Chesbrough and Growther (2006) confirmed the relevance of this approach in more mature industries. In industries other than high-tech, when companies look outside for technologies to extend or defend their core business, they minimize risk by investing in technology that is often proven in other applications (as opposed to 'new to the world' technologies). This paper shows how the innovative action of the firm in the mature shipbuilding industry may depend on open innovation concept.

2 Open Innovation as a Suitable Strategy in Mature Environments

The conceptual argument suggests that firms benefit from highlighting newness, responsiveness, and a degree of boldness. These results hold up across different nations, industries and other contextual variables (Iakovleva 2005; Grande et al. 2011). Scholars generally define innovation as the development and commercialization of new ideas in organizations manifested in terms of a new product, service or method of production or a new market, organizational structure or administrative system (Foss et al. 2011a, b; Damanpour and Wischnevsky 2006). In his classic treatment on the subject of innovation, Schumpeter (1934, 1942) defined innovation as 'new combinations' of existing knowledge and resources, arguing that innovation thus defined the source of economic and social change. Without innovative efforts by the entrepreneurial individuals, society would in his view be stagnant. Due to labor mobility, abundant venture capital and widely dispersed knowledge across multiple public and private organizations, entrepreneurs can no longer afford to innovate on their own, but rather need to engage in alternative innovation practices (van de Vrande et al. 2009; Chesbrough 2003). However, mature environment demands significant efforts to stay "on the water" as competition is hard and resources often are just enough to continue the known line to products and services. It is at that point thinking in terms of open innovation might provide new ways of solving this challenge (Gans and Stern 2003). Open innovation comprises both outside-in and inside-out movements of technological ideas (Lichtenthaler 2008). We may expect firms in mature industry to rely on both inbound and outbound open innovation simultaneously (van de Vrande et al. 2009).

3 Method

3.1 Case Study Approach

This exploratory study was positioned within an interpretive research paradigm. Single case study method (Yin 2003) will be used to explore the research questions related to open innovation. This technique enables the analyst to get deep insights into the mechanisms which stand behind the selection mechanism for open innovation. A qualitative case study method is appropriate because the aim of this study is to generate fresh and deeper insights into the process of partner selection related to an open innovation.

3.2 Case Selection

We selected a case of development of a unique revolutionary ship which uses LNG and hydrogen power. There is only one ship in Norway under development of this type. We studied the process of how open innovation process helps to satisfy demand for the high-end sophisticated vessels during offshore operations.

3.3 Data Collection and Analysis

In 2015, four semi-structured interviews were carried out among employees of a Norwegian shipping company. The interviews lasted from 60 to 120 min with the project managers responsible for the Arctic project in the shipping company, and the shipyard.

In order to triangulate information collected from face-to-face interviews, additional data sources were used (e.g., information from reports, company web pages, the Internet more widely and from trade/technical magazines). By combining several modes of data collection, in-depth description of partner selection process was obtained.

Narrative accounts relating to the development of entrepreneurial competencies were analyzed. Comments of interviewees were consistently coded, and most frequently reported partner selection criteria were identified. An iterative analysis relating to within-case analysis was conducted (Eisenhardt 1989). Data were compared with existing theory and the data was allowed to talk.

4 Case Illustration

Simon Møkster Shipping AS is an offshore shipping company located in Stavanger, the oil capital of Norway. Captain Simon Møkster established the company in 1968. The company is still owned and managed by the Møkster family. The company owns 25 offshore vessels. There are 665 employees in the company, 32 persons work in the main office, and the rest of employees are in the sea. Though many Norwegian offshore shipping companies operate both in the Norwegian offshore sector and worldwide, Simon Møkster Shipping AS decided only to develop their operations alongside the

Norwegian coast, one of the most challenging sea areas in the World. They focused on adding competence on operating further up North. The Arctic contains as much as 25% of the remaining oil and gas resources in the world. The high Arctic encompasses the regions north of the Arctic Circle where cold weather may cause severe ice and icing conditions. It includes Alaska, the North Eastern Part of Canada, the Greenland coast, the Barents Sea, North and East Russia, and the North East Passage (Northern Sea Route) through Northern Russia. The stakeholder of the Norwegian petroleum industry move their operations gradually towards the north into the Norwegian Sea, the Barents Sea and the rest of what is termed Arctic waters (Borch and Batalden 2015). Operations in this region require vessels that are tailor-made for a harsh climate and an area with limited infrastructure. This calls for vessels with ice-strengthening, high degree of functionality, and well-equipped for multi-purpose action.

Recently, the Møkster shipping company wanted to win contracts to support offshore oil and gas operations in the Goliath field in the Barents Sea with standby vessel and emergency response and rescue vessels. The conditions for offshore operations in the Barents Sea are different from the familiar conditions of the North Sea.

«It depends upon how long north you are… Challenges begin when you are so far north that you have ice. Then you really need a special competence. How to navigate when it is ice on the sea… for the ship there are ice-specifications than…. Also if you are long north, you have the distances… helicopter cannot rich boats. For example when we were in Murmansk, we had these challenges… it is difficult if not impossible to get things you need. Than we have to take on board a lot of repair details, just in case… the same with food, we have to take on board as much as we can. This was a 3-month contract. We got problems with different antennas, TV, and the correction signal».

The conversion of standard offshore vessels to ice-class vessels and equipping them with the winterization package is not the optimal solution (Berg et al. 2012). Thus, the company's management decided to invest into the fleet of tailor-made vessels for the Arctic. Since the Arctic market is quite new and such ships have not been developed before by the Norwegian designers, the design appeared to be very innovative in terms of functionality, capacity and environmental friendly operation. Simon Møkster has developed several boats to serve this need, staring from Stril Luna, a supply-boat that was ready in 2014 for Statoil, than Stril Barents, which was delivered in February 2016 for ENI, and finally they just got a new contract to deliver a novel supply-boat equipped for both requiring functions for north operations to be used in Goliat field also fro ENI.

For the development of these boats, two options were available in relation to innovation, i.e. either to order a complete new project to a ship design company and not to be involved into the innovation development process or to engage into open innovation and to be an active participant in the R&D. The shipping company opted for the latter opportunity even though it had a small administration with limited capacity.

The shipping company decided to cooperate with the ship design and shipbuilding company to develop a tailor-made vessel. They evaluated several candidates looking into design companies. Due to technology newness and the lack of R&D in this area, the company searched for partners that could understand every aspect of the value chain including designing, building, equipping and running this type of vessels.

«Often it is cooperation with designer firm. We can spend a lot of time together… We communicate with them all the time, we come with suggestions that we would like it like this or that, change this or that aspect; we have this and that ideas. It is very inspiring process. I have been involved twice in it, and yes, it is very inspiring».

For the last vessel, the Vard company was selected in order to develop and build the vessel that should support operations of the oil company ENI. Vard unites both ship design firm and a shipyard in the same corporation, in addition to equipment and industry service. From the shipping company they included their top levels competence including the CEO, CTO and operation manage following the process closely and scrutinizing the suggestions from the design company. They involved their most experienced operative personnel bringing them to shore from their vessels to work on the details. The operating personnel together with middle management cooperated tightly with the designers to bring the different units together. Information was constantly exchanged between partners, and the designers had to reveal their knowledge as to best practice in the field and the limitations of different constructions.

"We had a lot of meetings – before and during the process, and also afterward when we did evaluation and analyzed what was successful and what could be improved. So we used a lot of time for that".

In addition, the cost of building and running the vessels was a critical issue.

"It was price ... and ... price and availability of the yard. We get a design, it costs something. Then we must get a price from a shipyard with delivery time on what it costs to build this design".

Therefore, not only the technical aspects had to be considered during the process. The financial and operating staff had to be included and the designers, the yard and the equipment producers confronted both with functionality and cost issues. One challenge was to learn what details the oil company would demand. The oil company participated marginally in the new vessel development due to market rules as to open competition. However, there was a systematic evaluation of data from other contracts with oil companies and the tacit knowledge acquired by senior staff and vessel management.

"The customer says we should have such a boat. It comes with a specification, which you will deliver, and so we have to be creative, and sometimes we get for example, two designers, and then we are working on the drawing there, and according to what we think is best. The specification is a guideline, but then you have to try to think about what can make us win, other than the price. Because that's the reason ... and price also depends on what you out in it. So once we have arrived at the design and the price, we sit down and calculate… how much you will be left with, how much do you think it will cost, class, how much maintenance do we get, it's more expensive in the Barents Sea than here - yes. Then you enter a very math and calculate ...".

After the first stage of development, an offer was given to the oil company on time with the necessary specifications. The Møkster company competed with several other concepts, but won due to special details on functionality, environmental friendly solutions combined with competitive price. The new advanced vessel with unique characteristics was delivered to owners in 2015.

"So, our customers are our partners ... at the same time we know that they can choose many others, but steel they are our major partners. So we have to deliver a high quality product all the way. We have demanding customers who place a lot of demands on us on delivery".

5 Conclusion

The case discussion illustrates how a shipping company became involved in the open innovation process (Iakovleva 2013; Solesvik and Gulbrandsen 2013, 2014) by first choosing a design partner with broad value chain insight, and then to take part in the whole process of design and development with their own staff. In this case, an important factor for involving a company into open innovation was the lack of knowledge about market characteristics and customer needs. The context of Arctic is specific (Borch and Solesvik 2016) and not much expertise was accumulated in the area of offshore operations in the Arctic with harsh weather conditions and long distances to the shore. Møkster brought in their most experienced personnel to interact with the designer and refine the tacit operation knowledge into formalized knowledge related to the functionality of the vessel, and to have the necessary tailor-make and functionality guaranteed. This included a time consuming representation at the shipyard following the building process with two-three persons at the site, being constantly in a dialogue with the different sub-contractors and installations. Second, there were demanding customers to consider. In this industry, the oil companies continuously look for vessels with increased productivity, safety standards and efficiency. The competition in this mature market with several large suppliers is fierce. One of the ways to win a competition game is to be one-step ahead compared to rivals. Innovativeness that supplies the crew and shore staff with market context competence is an important factor that can help the firms to be ready for the next step up the innovation ladder. Many companies prefer to concentrate on their own core competencies (Borch and Solesvik 2013; 2014). Møkster decided to lift their core competence into the cooperating design (Solesvik 2011), equipment and construction companies. This is in line with the findings of recent research that stresses the popularity of multi- firm network organizational form in the contemporary business life (Fjeldstad et al. 2012). According to multi-firm network concept, firms concentrate on their core competencies and collaborate with other firms in order to get other firms' core competencies to achieve project's goals. In the case of Simon Møkster Shipping AS, the company brought both their strategic apex and the most competent middle management staff into the process to a larger degree than major part of competitors. The company contributed with the expertise of its employees in Arctic waters operations, and was creative in their price and contract strategy towards ENI as their customer. Møkster then succeeded in not only emphasizing the technological innovation, but also innovation related to management and marketing. The management group of the firm supported intra-firm collaboration collecting feedback from the sea personal related to operations in the Arctic seas and which construction features should be taken into account in the design phase. The company shared the salaries and scarce time of key personnel with the ship designers and yard personnel. In return, the shipping company employees involved in this open innovation project acquired new insight on complex construction under uncertainty, and the designer's knowledge on competitor's best practice. Some other firms and organizations shared their competencies as well, i.e. classification society, equipment suppliers, and others. This connection may take a formal contract approach with loose couplings, or it may become a long-term partnership with strong ties based on trust and reciprocal exchange as in the cluster thinking. In this case, strong cluster mechanisms were present in the region serving as a platform for specific cooperative arrangements. As suggested by the competence-based view, the firms can stretch and develop their competencies if necessary (Solesvik 2018). In our case, collaboration using an open innovation approach helped the shipping company in a way to overcome its liability of smallness to develop a competitive edge in R&D and innovation.

References

Berg, T.E., Berge, B.O., Borgen, H., Hänninen, S.: Intervention vessel for Barents Sea operation. In: ASME 2012 31st International Conference on Ocean, Offshore and Arctic Engineering, Rio de Janeiro, Brazil, 1–6 July 2012, pp. 571–578. American Society of Mechanical Engineers, New York (2012)

Borch, O.J., Batalden, B.-M.: Business-process management in high-turbulence environments: the case of the offshore service vessel industry. Maritime Policy Manag. **42**(5), 481–498 (2015)

Borch, O.J., Solesvik, M.: Partner selection for innovation projects. In: ISPIM Americas Innovation Forum 2014, Montreal, Canada, 5–8 October 2014 (2014)

Borch, O.J., Solesvik, M.Z.: Collaborative design of advanced vessel technology for offshore operations in Arctic waters. In: Luo, Y. (ed.) Cooperative Design, Visualization, and Engineering, pp. 157–160. Springer, Berlin (2013)

Borch, O.J., Solesvik, M.Z.: Partner selection versus partner attraction in R&D strategic alliances: the case of the Norwegian shipping industry. Int. J. Technol. Mark. **11**(4), 421–439 (2016)

Chesbrough, H., Growther, K.: Beyond high tech: early adopters of open innovation in other industries. R&D Manag. **36**(3), 229–236 (2006)

Chesbrough, H.: Graceful exits and foregone opportunities: xerox's management of its technology spinoff organizations. Bus. History Rev. **76**(4), 803–808 (2002)

Chesbrough, H.: Open Innovation: The New Imperative for Creating and Profiting from Technology. Harvard Business School Press, Boston (2003)

Chesbrough, H., Vanhaverbeke, W., West, J.: Open Innovation: Re-searching a New Paradigm. Oxford University Press, Oxford (2006)

Damanpour, F., Wischnevsky, J.D.: Research on innovation in organisations: distinguishing innovation-generating from innovation-adopting organizations. J. Eng. Tech. Manag. **23**(4), 269–291 (2006)

Fjeldstad, Ø.D., Snow, C.C., Miles, R.E., Lettl, C.: The architecture of collaboration. Strateg. Manag. J. **33**(6), 734–750 (2012)

Foss, L., Iakovleva, T., Kickul, J., Oftedal, E., Solheim, A.: Discontinuous innovations - the case of Norwegian petroleum industry. In: Minola, T., Cassio, L., Massis, A. (eds.) Entrepreneurship. Technology and Change. Edward Elgar Publisher, Amsterdam (2012)

Grande, J., Madsen, E.L., Borch, O.J.: The relationship between resources, entrepreneurial orientation and performance in farm-based ventures. Entrep. Reg. Dev. **23**(3), 89–111 (2011)

Iakovleva, T.: Entrepreneurial orientation of Russian SME. In: Vining, T., Voort, R. (eds.) The Emergence of Entrepreneurial Economics, pp. 83–97. Elsevier Science, Amsterdam (2005)

Iakovleva, T.: Open innovation at the root of entrepreneurial strategy. A case from Norwegian oil industry. Technol. Innov. Manag. Rev. **3**(4), 18–22 (2013)

Schumpeter, J.: The Theory of Economic Development. Harvard University Press, Cambridge (1934)

Schumpeter, J.: Capitalism, Socialism, and Democracy. Harper, New York (1942)

Solesvik, M.: Collaborative knowledge management: case studies from ship design. Int. J. Bus. Inf. Syst. **8**(2), 131–145 (2011)

Solesvik, M.Z.: Partner selection in green innovation projects. In: Berger-Vachon, C., Gil Lafuente, A., Kacprzyk, J., Kondratenko, Y., Merigó, J., Morabito, C. (eds.) Complex Systems: Solutions and Challenges in Economics, Management and Engineering, pp. 471–480. Springer, Berlin (2018)

Solesvik, M.Z., Gulbrandsen, M.: Partner selection for open innovation. Technol. Innov. Manag. Rev. **3**(4), 11–16 (2013)

Solesvik, M., Gulbrandsen, M.: Interaction for innovation: comparing Norwegian regions. J. Entrep. Manag. Innov. (JEMI) **10**(3), 7–28 (2014)

Solesvik, M., Kondratenko, Y., Kondratenko, G., Sidenko, I., Kharchenko, V., Boyarchuk, A.: Fuzzy decision support systems in marine practice. In: IEEE International Conference on Fuzzy Logic, Naples, Italy, 9–12 July 2017, pp. 1–6. IEEE, New York (2017)

van de Vrande, V., de Jong, J., Vanhaverbeke, W., de Rochemont, M.: Open innovation in SMEs: trends, motives and management challenges. Technovation **29**(6–7), 423–437 (2009)

Responsible Innovations in Healthcare Sector

T. Iakovleva(✉)

Professor of Entrepreneurship in Stavanger Business School,
University of Stavanger, Stavanger, Norway
tatiana.a.iakovleva@uis.no

Abstract. Present paper describes the process of the development of responsible innovations on the firm level on the example of a case from a in healthcare sector in Norway. The rapid global diffusion of information and communication technologies has greatly improved access to knowledge. At the same time, communication is cheap, information is a commodity, and global trade increases technological diffusion. As a result, firms and users, including those outside of industrialized nations, get early exposure to the latest technologies and information. General-purpose technologies such as mobile phones and 3D printers enable individuals to solve local needs and customize products. The combined effect of these changes is having a profound impact on the innovation landscape. Meanwhile, the healthcare sector is facing unprecedented challenges, which are magnified by budgetary constraints, an aging population and the desire to provide care for all. This article addresses the question of *how responsible research and innovation may contribute to developing solutions to grand societal challenges such as healthcare*. A broad definition of the concept of responsibility in the context of innovation is adopted in this paper. Responsibility is thus seen as a collective, uncertain and future-oriented activity. This opens the questions of how responsibilities are perceived and distributed and how innovation and science can be governed and stewarded towards socially desirable and acceptable ends.

1 Introduction

One mounting European challenge are the cost and quality of healthcare services. However, in this area, new technology holds potential of providing cheaper care at higher quality. However, in many cases, technologies with a competitive advantage and a defined need, fail. Evolutionary theory of the firm (Hodgson 1998) argues that in many cases, innovations find boundaries before their establishment. There is a selection mechanism though system of norms and rules written to “trap or release” an innovation. Thus, even though innovations can provide economic growth through better solutions, they might not be able to reach the market. The Responsible Research and Innovation (RRI) perspective address this issue through proposing an innovation process where societal actors (researchers, citizens, policy makers, business, third sector organizations, etc.) work together during the whole research and innovation process. They have to negotiate which types of research and innovation to pursue. RRI includes multi-actor and public engagement in research and innovation, enabling easier access to scientific results. RRIs may also be a result of corporate social responsibility

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 820–827, 2020.
https://doi.org/10.1007/978-3-030-48021-9_90

(CSR) and may take place in every area where negative externalities are the result of human commercial activity (Laudal 2011). RRI is driven by political, environmental and social factors. As a result, many stakeholders are involved that have different success criteria (Owen, Bessant and Heinz 2013).

2 Responsible Research and Innovation in Healthcare: A Challenged Promise

The RRI concept acknowledges the power of research and innovation as a mechanism for genuine and transformative societal change to shape our collective future. In the words of (Owen, Macnaghten and Stilgoe 2012), 'Research and innovation must respond to the needs and ambitions of society, reflect its values and be responsible … our duty as policy-makers (is) to shape a governance framework that encourages RRI'. As pointed by (Benneworth et al. 2014), RRI is far closer to a new paradigm in public engagement with science, a deepening of relationships and responsibilities of societal stakeholders for grant- ing scientists and innovators a 'license to practice' (Asheim and Gertler 2005).

In their pioneering work, (Stilgoe, Owen, and Macnaghten 2012) suggests four integrated dimensions of responsible innovation process: Anticipation, reflexivity, inclusion and responsiveness. As authors argue, these dimensions are important characteristics of a more responsible vision of innovation. Anticipation involves systematic thinking aimed at increasing resilience, while revealing new opportunities for innovation and the shaping of agendas for socially-robust risk research (Martin 2010). Reflexivity, at the level of institutional practice, means holding a mirror up to one's own activities, commitments and assumptions, being aware of the limits of knowledge and being mindful that a particular framing of an issue may not be universally held. Inclusion, that can be referred as user-driven, open (Lynch 2000), open source (Chesbrough and Crowthe 2006), participatory (Raymond 1999) and networked innovation (Buur and Matthews 2008) all suggest the possibility of including new voices in discussions of the ends as well as the means of innovation. Finally, responsiveness means a capacity to change shape or direction in response to stakeholder and public values and changing circumstances.

Bringing the above mentioned dimensions in the innovation discussion in health and welfare sectors might allow to look into the challenges of developing innovations in a new way. Health sector, at least in some countries, is a highly regulated sector of economy. Private actors should comply with several rules and regulations, especially during the commercialization process. Bring innovations in this sector is a difficult task. From one side, innovations are highly appreciated and needed by users. From the other side, users and customer are sometimes different, and budgetary constraints as well as security issues challenge the diffusion of innovative solutions. Welfare sector is less regulated, but the same challenge of users-customers often is in place. Focusing on anticipation, reflectivity, inclusion and responsiveness seems to be necessary in order to foster commercialization of innovations in these sectors. Thinking about health and welfare sector, it is important to notice that the four elements of firm RRI framework should include interactions with users, potential customers and other stakeholders.

Therefore, the RRI framework can be applied to get the organization/individual introducing an innovation to think more extensively around the proposal and explore it along dimensions which would help ensure its value. The four dimensions challenge the company to ask itself

- How far have we thought through the future implications of what we are trying to do? (Anticipation)
- How far do we challenge ourselves, discuss and debate what we are trying to do, look at it from different perspectives? (Reflection)
- How far do we try to include the views and perspectives of our planned end users and other possible stakeholders? (Inclusion)
- How far is there flexibility in our design to change it in response to the above questions? (Responsiveness)

The same type of questions can be applied for any actor/stakeholder in the value chain, so that principles of anticipation, reflection, inclusion and responsiveness being embedded into the decision-making process. Moreover, these processes or concepts are not mutually exclusive or sequential, rather they should be seen as interacting and adding value to each other. Thus, any innovation project or process should integrate in some degree these elements, and should be seen and embedded into its governance.

3 Taking Responsible Innovations to Market

The challenge with responsible innovations may be that they are driven from a policy or interest – perspective and not from a market point of view. This leads to an extra challenge when it comes to commercialization. In the welfare technology/telemedicine area, for instance, one observes that the user and the customer are separated. The user may be an individual with need of care and with little influence power, while the customer may be the public health sector which is resource constrained and do not know enough about the user. Thus, it is important to explore the process of taking responsible innovations to the market. What drives RRI to market and how different is the mechanism of commercialization for responsible innovation? Which stakeholders influence that process and how does it actually happen? In Europe, the current overarching strategy, Europe 2020 effectively frames social innovation as a mechanism for responding to an array of the non-economic elements of these challenges. A range of dialogues, often under the name of sustainable in- novation, have sought to find alternative and better ways to meet existing needs and to more effectively address the unintended consequences of industrial development upon society (STEPS 2010; Howaldt and Schwarz 2010). A research by (Weber and Rohracher 2012) maintains that a new policy for transformative change is emerging focused more in the role of research, technology and innovation towards societal challenges rather than economic growth. Given these emerging conditions social innovation has become identified with new forms of self-management and innovative bottom-up initiatives proposed to help groups and communities cope with marginalization and deprivation (Boyle and Harris 2010).

Outbound open innovation suggests that rather than relying entirely on internal paths to market, companies can look for external organizations with business models that are better suited to commercialize a given technology (Iakovleva 2013). Extending this mind set one can see how user-driven ideas and solutions might be incorporated into innovation products and services, which allows addressing the Grand Challenges of the 21st century (Murray, Caulier-Grice, and Mulgan 2010).

4 Example of Responsible Innovation in Healthcare

In this paper we present an example of responsible innovation in healthcare, the private company called Laerdal medical AS established in Norway in 1940. Within its 60 years of development this company demonstrated the deliberate focus on responsibility, which guide its business development. Åsmund S. Laerdal setup his company in 1940, creating series of books and wooden toys for children (Tjormsdal 2015). A few years later, Laerdal began experimenting with soft plastics, developing plastic toys for kids. Laerdal branched into the medical industry when the Norwegian Civil Defense requested the company to develop plastic imitation wounds for training. Then, in 1954, when Laerdal's son was found lifeless in a creek, he saved the boy by shaking him vigorously and making him vomit water. This event, coupled with his exposure to a mouth-to-mouth life saving technique at a medical conference, stimulated Laerdal's interest in developing a tool for the education and training of this method, and propelled him into the production of the world's first resuscitation mannequin. From then on, the company's mission changed from simply inspiring children, to helping prevent untimely deaths. Since then and until now the company mission presented on its webpage is "saving more lives together". Re-establishing itself as Laerdal Medical, the company began producing a variety of life-saving medical products from ventilators and suction units to full- sized training mannequins.

In 2000, research revealed in the UN's Millennium Development Goals four and five showed large figures for birth-related deaths in under-developed countries; Laerdal adjusted the scope of its mission to include addressing this issue. Thus, the non-profit sister company, Laerdal Global Health, was formed dedicated to saving babies and mothers in low-resource areas.

Not long after, Laerdal companies conducted an internal review and expanded its vision, declaring a goal to help save more lives - 500,000 a year - by 2020. The aim is to help save 100,000 lives in high-resource areas by increasing cardiac survival rates and patient safety. For the remaining 400,000 lives, the company will focus on addressing birth related fatalities and helping children under five survive in low resource areas. With offices in 24 countries, employing roughly 1400 people, the private, family-owned business has become a dominant player in the medical industry. In addition to Laerdal Medical and Laerdal Global Health, the corporation has also invested in the SAFER-Simulation Center, an in-house equipment training facility as well as the Laerdal Foundation, which offers aid for practically oriented research to improve life-saving and acute care work.

Open Innovation and Responsibility

Laerdal vision is about "mutual trust and opened doors" which the company follows by expanding beyond national borders and being curious in other cultures. For Laerdal it is essential to collaborate with other companies both in Norway and abroad in order to save more lives, simplify their products and programs, embrace customer responsiveness, and deliver quality in all they do. The company clearly operates with open innovation process. The open innovation aims to improve Laerdal's ability to innovate new ideas for products and solutions by exploiting ideas and solutions from other consumers and partners. For Laerdal it is about gathering knowledge and about learning. By listening to their consumer's needs, they redefined healthcare quality through changes in training and procedures such as the revolutionary low dose high frequency training program.

Laerdal has not always operated with open innovation, it started with a more closed innovation concept, manufacturing only in-house. Later the firm transitioned to a more open concept which included external suppliers in the manufacturing process. Moreover, Laerdal relied only on their own engineers in order to gain more knowledge and continuously developing new and better products and services. They later realized the importance of including external knowledge and began including partners and users in the product development process.

Laerdal does operate with open innovation, but does not share all the information. They maintain a bit of traditional closed innovation strategy in terms of concealing sensitive information such as sales numbers, market position and financial status in all countries as well as their skin formula trade secret.

The composition of Laerdal is another example its open innovation strategy. The firm is diverse with multicultural operations in 24 countries; 40 different nationalities can be found in the headquarters office alone. And it's network of alliances, partners and customers expands through more than 75 countries. Combining the competencies of the numerous cultures both inside the company and in its increasing network of alliances and partners, enables Laerdal to cultivate its absorptive capacity and knowledge base - two elements proven crucial for innovative firms (Cohen and Levinthal 1990). By combining both an open and closed innovation strategy Laerdal is able to benefit from the additional knowledge's sources without the risk of losing any competitive advantage.

Responsible Innovation

It's easy to see social innovation is integrated in Laerdal's business strategy. The company's mission "helping save lives" makes this abundantly clear. The company sprouted from social entrepreneurship; Åsmund Laerdal mobilized his network and personal resources to acquire the materials required to address society's need for children's inspiration. As Laerdal grew his business and expanded his mission, the company made the necessary transition into third party social innovation. And through its involvement with diverse national institutions such as American Heart Association, Red Cross and European Resuscitation Council, the company has contributed to public-sector innovations, such as CPR training school curriculum requirements.

Since inception, all of Laerdal's products and services have been created in order to improve quality of life - not only through the products themselves, but also by keeping them affordable and ensuring greater access to healthcare. Laerdal's CPR Anytime is a 30 min, thorough training program priced at less than $30, making it highly-affordable to schools and other institutions around the world. (Tjomsland 2015). According to Cole Edmonson, Chief Nursing Officer at Texas Health Presbyterian Hospital in Dallas, implementation of Laerdal's RQI program has saved the hospital nearly $500,000 in annual CPR related training and education expenses. (Tjomsland 2015). Additionally, Laerdal is able to increase the number of lives it helps to save by diversifying in both physical products and service solutions.

While Laerdal is ultimately focused on creating societal value, it conducts both commercial and philanthropic business through Laerdal Medical (LM) and Laerdal Global Health (LGH), respectively. This diversified strategy ensures the company's overall success and sustainability. Though LGH receives some contributions from partners and the Laerdal family, its continuation relies heavily on support from Laerdal Medical, which invests a majority of its commercial proceeds into LGH. (Westnes et al. 2018). Conversely, Laerdal Medical is able to continuously succeed at developing useful and innovative products by utilizing research form LGH programs. (Tjomsland 2015).

5 Conclusion

Responsibility has always been an important theme of research and innovation practice. It has been framed differently in different settings. From the scientist point of view, responsible innovation often limited to production of reliable knowledge, For the society, however, innovations have to address societal needs and solve problems that modern society meets. One of such needs associated with available and safe healthcare service, as well with the prevention measures to prolong well being of population.

The health care sector meets future challenges and demands that available resources that are often publicly funded used with care. There is a need to find new ways of doing more for less. New technologies and especially digitalization brings a possible solution to this challenge. The trend of aging population is difficult to ignore, and society will soon need a health care where less people do more tasks. This will allow to reduce burden of financing the public procurement of such services. According to (Røtnes and Staalesen 2010), there is a high demand for high quality services and treatment. Innovation, in medical treatment and equipment as well as processes and methods, is essential to provide effective care for citizens.

The suggested framework of responsible innovation can be a way to guide governance developments in these sectors in order to enable social learning and em- power social agency. Much needed upscaling of innovations in welfare and healthcare sectors should be done with responsiveness in minds of those who take decisions. Incorporating user needs and wants, constantly reflecting of the feedback, meeting anticipation seems to be the right means to achieve accept of the innovations in the society, and thus open up for upscaling and dissemination. Therefore, the framework presented above could be used by firms trying to take innovations to market, as well as by the society. It will allow to create more opportunities for engagement with range of publics that could enable researchers and innovators to hone and refine their sense of responsibility to society in the way in which they go about designing and conducting their innovations.

6 Acknowledgments

The author acknowledges the grant from Norwegian Research Council nr 247716/O70 for the research project "Digitalize or Die: Dynamic Drivers of Responsible Research and Innovation in Health and Welfare services".

The author acknowledge work of students of Stavanger University, Anette Horpestad, Kristina Kraabøl, and Ceilia M. Vrnak, contributing by Laerdal case story.

References

Asheim, B. T. and Gertler, M. S. (2005) "The Geography of Innovation: Regional Innovation Systems", in: Fagerberg, J., Mowery, D. & Nelson (eds.) The Oxford Handbook of Innovation. Oxford: Oxford University Press, 2005, pp. 291–317

Benneworth, P., Amanatidou, E., Schachter, M.E., Gulbrandsen, M.: Social innovation future: beyond policy panacea and conceptual ambiguity, position paper for European Forum for studies of Policies for Research and Innovation (2014). http://doc.utwente.nl/94038/1/benneworth%20paper%20H.pdf

Benneworth, P., Coenen, L., Moodysson, J., Asheim, B.: Exploring the multiple roles of Lund University in strengthening Scania's regional innovation system: Towards institutional learning? Eur. Plan. Stud. **17**(11), 1645–1664 (2009)

Boyle, D., Harris, M.: "The challenge of coproduction", Discussion paper. London, UK: NESTA. CE, 2010 (2013)

Buur, J., Matthews, B.: Participatory innovation. Int. J. Inn. Manag. **12**, 255–273 (2008)

Chesbrough, H.W., Crowther, A.K.: Beyond high-tech: early adopters of Open Innovation in other industries. R&D Manag. **36**(3), 229–236 (2006)

Cohen, W., Levinthal, D.: Absorptive capacity: a new perspective on learning and innovation. Admin. Sci. Q. **35**, 123–133 (1990)

Geoghegan-Quinn M.: (2012). http://ec.europa.eu/commission_2010-2014/geogheganquinn/headlines/speeches/2012/documents/20120423-dialogue-conference-speech_en.pdf

Hodgson, G.: Evolutionary and competence-based theories of the firm. J. Econ. Stud. **25**(1), 25–56 (1998)

Howaldt, J., Schwarz, M.: SI: Concepts, research fields and international trends, Report of ESF, EU, and Aachen University, 02 April 2013 (2010). http://www.internationalmonitoring.com/fileadmin/Down-loads/Trendstudien/Trendstudie_Howaldt_englisch.pdf

Iakovleva, T.: Open Innovation as the Root for SME's Entrepreneurial Strategies. Technol. Inn. Manag. Rev., 18–22 April 2013

Laudal, T.: Drivers and barriers of CSR and the size and internationalization of firms. Soc. Responsib. J. **7**(2), 234–256 (2011)

Lynch, M.: Against reflexivity as an academic virtue and source of privileged knowledge. Theory Cult. Soc. **17**, 26–54 (2000)

Martin, B.: The origins of the concept of 'foresight' in science and technology: an insider's perspective. Technol. Forecast. Soc. Change **77**(2013), 1438–1447 (2010)

Murray, R., Caulier-Grice, J., Mulgan, G.: The Open Book of Social Innovation, NESTA and the Young Foundation (2010)

NoU:11, Innovasjon i Omsorg. Helse og Omsorgsdepartmenetet, p. 90 (2011). ISSN 0333-2306 ISBN 978-583-1099-7

Owen, R., Macnaghten, P.M., Stilgoe, J.: Responsible Research and Innovation: from science in society to science for society, with society. Sci. Public Policy **39**(6), 751–760 (2012)

Owen, R., Bessant, J., Heinz, M.: Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Wiley, UK (2013)

Raymond, E.: The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary. O'Reilly and Associates Inc, Sebastopol (1999)

Røtnes, R., Staalesen, P.D.: New methods for user driven innovation in the health care sector, ECON Pöyry AS (2010). http://www.nordicinnovation.org/Global/_Publications/Re-ports/2010/New%20methods%20for%20user%20driven%20innova-tion%20in%20the%20health%20care%20sector.pdf

STEPS: Innovation, Sustainability, Development: A New Manifesto. STEPS Centre, Brighton (2010)

Stilgoe, J., Owen, R., Macnaghten, P.: Developing a framework for responsible innovation. Res. Policy **42**, 1568–1580 (2012)

Tambo, T., Hoffmann-Petersen, N., Bejder, K.: Architecting for connected healthcare – a case of telehomecare and hypertension. In: Saha, P. (ed.) Enterprise Architecture for Connected Government: Practices and Innovations. IGI Global, Hershey (2012)

Tjomsland, N.: Saving more lives - together. The vision for 2020, 1st edn. Laerdal Medical, Stavanger (2015)

Weber, K.M., Rohracher, H.: Legitimizing research, technology and innovation policies for transformative change Combining insights from innovation systems and multi-level perspective in a comprehensive "failures" framework. Res. Policy **41**, 1037–1047 (2012)

Westnes, P., Horpestad, A., Kraabøl, K., Vrank, C.M.: Interview with Laerdal Medical AS (2018)

The Role of Stakeholders in the Context of Responsible Innovation

Luciana Maines da Silva[1(✉)], Cláudia Cristina Bitencourt[1], Kadigia Faccin[1], and Tatiana Aleksandrovna Iakovleva[2]

[1] Unisinos Business School, UNISINOS University, São Leopoldo, Brazil
lucmasi@gmail.com, lucianamaines@unisinos.br
[2] UiS Business School, Stavanger University, Stavanger, Norway

Abstract. The participation of multiple stakeholders in the innovation process is one of the assumptions of Responsible Innovation (RI). This partnership aims to broaden visions, in order to generate debate and engagement. The present study's aim, based on a meta-synthesis, is to evaluate how stakeholder participation in RI takes place. Thus, qualitative case studies were identified that investigated the participation of stakeholders in responsible innovation. Those studies have shown that, although participation is achieved when innovation is already in the process of being implemented or already inserted in the market, it serves as a basis for modifications, both in the developed product and in the paradigm of innovation. Based on the concept of Responsible Innovation and its dimensions, the role of stakeholders in the context of innovation is restricted to consultative participation. The agents that stimulate their participation are academic researchers and researchers linked to multi-institutional projects. We have noticed that the studies favour the participation of multiple stakeholders like policymakers (including funding agencies, regulators and executives), business/industry representatives (internal or outsourced innovation departments and/or some R & D base), civil society organizations (such as foundations, associations, social movements, community organizations, charities, media), as well as researchers and innovators (affiliates of various institutions and organizations at different levels). One point that stands out is the change of vision of one stakeholder over the other. Although the difficulty is pointed out in the dialogue, it is possible, by inserting them collectively into the discussion, that the different stakeholders will develop a better understanding of the different points of view. The present study has discovered that RI is treated as a result and not as a process.

1 Introduction

It is the social and environmental concerns of research and innovation that have resulted in one of the most recent terms associated with innovation. Responsible innovation (RI) is primarily related to responsible research, which has recently become a part of management and business studies (Owen et al. 2012; Stilgoe et al. 2013). The main studies address RI from the perspective of governance (Stilgoe et al. 2013) and public policy (Owen et al. 2012).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 828–835, 2020.
https://doi.org/10.1007/978-3-030-48021-9_91

One of the main assumptions of RI is the participation of several stakeholders in the innovation process, with the objective of broadening visions, purposes, issues and dilemmas for wide and collective deliberation through the processes of dialogue, engagement and debate (Owen et al. 2013). The intention of this participation is to develop greater democratic accountability in the innovation lifecycle (Eden et al. 2013). It may be in the early stages (Burget et al. 2017), when innovation has already been developed or is already in the marketplace (Stilgoe et al. 2013; Owen et al. 2013).

The way this participation is conducted does not follow a pattern and may vary according to the nature and flow of information between those responsible for the exercises and the participants. According to the flow of information, the effectiveness of an exercise can be determined by how efficiently the complete and relevant information is obtained from all appropriate sources, transferred to (and processed by) those responsible and combined to generate a response (Rowe and Frewer 2005), or reconfiguration of the initially proposed model.

Based on the premise of the need for inclusion (Stilgoe et al. 2013; Owen et al. 2013) of the various stakeholders, preferably in the early stages of innovation (Burget et al. 2017), and considering that, many times, innovators and stakeholders have different objectives, it is sought, through this study, to evaluate: **How does stakeholder participation in the process of responsible innovation occur?** To advance in the theoretical construction of the phenomenon of responsible innovation, as well as to answer the guiding question of the study, a meta-synthesis was performed (Hoon 2013). In the following sections, the theoretical reference that is the bases for the present research will be presented, as well as the method used. Finally, the discussion and the final considerations are presented.

2 Responsible Innovation

Responsible innovation evokes a collective duty of care: a commitment to rethink the purposes and impacts of innovation, as well as a reflection on how to make its pathways sensitive to uncertainty (Mejlgaard and Bloch 2012). In considering innovation as a process, RI denotes a guideline for anticipation, inclusion, responsiveness, and reflexivity (van Oudheusden 2014). These four dimensions imply a collective and continuous commitment to be (To be What?) (Stilgoe et al. 2013). The anticipation describes and analyzes the intended and potentially unintended impacts that may arise and is supported by methodologies that include forecasting, technology assessment and scenario development. Reflection considers the underlying purposes, motivations, and potential impacts of what is known and what is not known. Inclusion refers to diminishing the authority of experts as part of a quest for legitimacy. Finally, responsiveness means that RI requires an ability to change form or direction in response to stakeholder and public values and changes in circumstances (Owen et al. 2013).

One of the aspects highlighted in terms of responsible innovation is the continuous participation of different actors, mainly highlighted by the inclusion dimension, which seeks to engage stakeholders already in the early stages of innovation (Burget et al. 2017). Stakeholder engagement is described below.

In addition to the participation of entrepreneurs and organizational agents, RI requires the inclusion of several actors in the context of responsible innovation (Ratiu and Mortan 2013; Yenicioglu and Suerdem 2015), seeking to engage different stakeholders already in the early stages of innovation (Burget et al. 2017). When perceived as a space constituted by activities, actors and norms, RI requires the participation of individual researchers, research organizations (both public and private), research ethics committees and their members, users of research and innovation, civil society (e.g., schools and universities) and public bodies (from local authorities to regional structures) (Stahl 2013).

However, it is considered complex and diffuse to organize and coordinate participation-based roles, due to political bias, in order to persuade others to some form of action (Yenicioglu and Suerdem 2015), which can influence the innovation policy (Bakker et al. 2014).

3 Methodology

Aiming to answer the guiding question, the meta-thesis method was defined, based on qualitative case studies, which produces a new and integrative interpretation of findings that is more substantial than those resulting from individual investigations. This methodology allows the clarification of concepts and standards and results in the improvement of existing knowledge states and emerging operational models and theories (Finfgeld 2003). The focus was the analysis of the evidence in all studies as well as in the construction of the theory, in order to guarantee sensitivity to the contextual considerations of the primary studies (Hoon 2013). From the establishment of the research strategy, the remaining steps are described below.

The first stage of meta-synthesis that was proposed (Hoon 2013) concerns the conceptual framework of the theme, aiming to identify theoretical gaps that can be later filled. In light of the role of stakeholders in the IR process, the first stage was developed, and the guiding question was proposed, and its results are already presented in Sect. 2.

The next step concerns the location of relevant surveys. In this step, the articles that may be considered relevant for meta-synthesis are identified. To locate the set of existing qualitative case studies, on September 7, 2017, searches were carried out on the databases of the CAPES Periodical Portal, EBSCOHost and Web of Science, totaling 217 articles. We identified 104 conceptual articles, 10 editorials, 14 review studies, 12 quantitative studies and 61 qualitative studies. Finally, each of the 61 qualitative studies was classified as 03 ethnographic, 01 action research, 01 phenomenology, 01 field study and 55 case studies.

The inclusion and exclusion criteria were then established. The first criterion that relates to the restriction of meta-synthesis was limited to articles that depend on qualitative case studies, ensuring that there is no difference between the research method that primary researchers claim to have used and the actual approach used (Hoon 2013). Articles using illustrative cases are also excluded. We excluded 24 articles. The second criterion relates to full access to studies. 2 articles were excluded. The third criterion concerns the quality of the studies, related to rigor, the link between theory and practice, the context of the case, multiple sources of data. Seven articles were excluded. Finally, 15 articles that did not refer to the innovation process were excluded or there was no direct stakeholder participation in the case study. There are 7 articles that, according to analysis, contribute to the understanding of the role of stakeholders in the context of responsible innovation.

The remaining articles were published between 2015 and 2017, in different countries (Germany, Canada, Denmark, Scotland, the Netherlands, Norway and the United Kingdom) and sectors. All studies clearly describe the methods of analysis, applying research strategies consistent with best practices.

The next proposed step is the extraction, coding and categorization of the evidence of the studies. The empirical material that serves as the basis for the meta-synthesis should highlight the insights of the original researchers through the understanding and interpretation of the data by the meta-synthesis (Hoon 2013).

To obtain the necessary data, a coding form was developed, aligning the research questions.

The studies are focused on the biofuel sectors (Raman et al. 2014; Shortall et al. 2015), sanitation (Metze et al. 2017; Wright and Kaiser 2015) and health. In the case of the latter theme, the studies deal with diseases such as demeanor (Decker et al. 2017), public safety through neuroimaging (de Jong et al. 2016) new drugs (Demers-Payette et al. 2016). All of the studies have been written in developed countries such as Germany, Canada, Denmark, Scotland, the Netherlands and the United Kingdom. All the studies reached the proposed objectives and present results based on the participation of the stakeholders. Only the study by Raman et al. (2014) do not present clearly the contribution of the stakeholders. Some of them have also managed to generate results for specific stakeholders, such as the case of Decker et al. (2017), which, through workshops with several participants, presented possibilities for improving products for technology developers. The study by de Jong et al. (2016) pointed to a change in perception of one stakeholder group over another group. Finally, the study by Bremer et al. (2015), used stakeholder participation to develop policies. The authors, however, pointed out the need for participation of these same stakeholders in the initial stages of the process, corroborating with what is proposed by Stilgoe et al. (2013) and Burget et al. (2017).

In order to analyze the studies under review, it was not only necessary to analyze at a specific level in each case, but also to answer the questions and meta-synthesis research objectives. In order to understand the role of stakeholders in the context of responsible innovation, as proposed by Hoon (2013), a causal network was developed for each study. Causal network is the definition of the most important independent and dependent variables in a field study (shown in boxes) and the relationships between them (shown by the arrows) (Miles et al. 2013). From the objective of the study the following criteria of analysis were proposed: which stakeholders are involved, which are the motivating agents of their insertion, what type of innovation did they influence and what was the stage of innovation. It was found that all articles brought the concept of Responsible Innovation to the discussion. The concepts of Owen et al. (2013) and von Schom-berg95 (2012) were the most mentioned. And in all the studies, several stakeholders were involved, prioritizing the discussion and exchange of ideas between them, not just discussing individually with each group. The study by Decker et al.

(2017) included patients with dementia in their study, even though they had already pointed out the limitations that this contribution could have. Further analysis is carried out in the next step.

4 Synthesizing Stakeholder Participation

Innovation can take several forms, summarized in four dimensions of change: product, process, position and paradigm innovation (Bessant and Tidd 2009). Because it is a complex process, product/service development requires the management of several factors at different stages. In general, it starts from concept generation to product marketing (placing on the market), through project design and testing (Tidd and Bessant 2015). The present study demonstrated that, although participation is realized when innovation is already in the phase of implementation or already inserted in the market, it serves as a basis for modifications, either in the developed product (Decker et al. 2017) or in the paradigm of innovation (Shortall et al. 2015).

The agents that stimulate their participation are academic researchers and researchers linked to multi-institutional projects (Bremer et al. 2015). It is perceived that the studies favor the participation of diverse stakeholders (Ratiu and Mortan 2013). Policy makers (including funding agencies, regulators and executives), business/industry representatives (internal or outsourced innovation departments and/or some R & D base); (such as foundations, associations, social movements, community organizations, charities, media), as well as researchers and innovators (affiliates of various institutions and organizations at different levels). From stakeholders, it was noticed that only the educational community (formal and informal, from the Ministry at the school level) was not included (Smallman et al. 2015). One point that stands out is the change of vision of one stakeholder over the other. Although the difficulties in dialogue (Blok 2014) are pointed out, it is possible, by inserting them collectively in the discussion, that different stakeholders develop a better understanding of the different points of view (de Jong et al. 2016).

From the results of the studies, it can be concluded that the participation is carried out in the final stage of innovation, that is, when it is already in the market, but serves as the basis for modifications in the developed product. Decker et al. (2017) found that the involvement of technology developers helped these participants to begin to imagine more specific potential technical solutions and to evaluate them in relation to their future desirability. A change in the paradigm of innovation was already addressed by Shortall et al. (2015) where, when consulting the stakeholders, an alternative agriculture was proposed, predicting a sustainable production of multifunctional biomass in terms of a nutrient and energy cycle in the farm and local production in a smaller scale. However, it was concluded that participation is merely advisory, with no evidence that the suggestions resulted in improvements in innovations.

The main agents that stimulated participation were academic researchers, as verified in Raman et al. (2014), Shortall et al. (2015), Demers-Payette et al. (2016), and Metze et al. (2017), interested in analyzing studies on biofuels or health systems, for example under the light of RI. Studies such as those by Bremer et al. (2015) and Decker et al. (2017) are part of larger studies developed by multi-institutional projects such as

Alliance Project (http://alliance-project.eu/) and Pegasus (http://pegasus.ieep.eu/). Although both academic researchers and agents involved in multi-institutional projects can also be considered as stakeholders, in these studies, they neither participated, nor helped developed the research.

It has also been realized that there are several stakeholders involved. Research included biomass producers, government representatives, biomass industry representatives, academics, NGO representatives (Shortall et al. 2015), representatives of the Ministry of Economic Affairs, nature and environment foundations, environmental education foundations, OECD (Metze et al. 2017), stakeholders in agriculture and intermediate work related to farms, bioenergy science, research and industry, policy makers and NGOs, and a group of experts (five leading members of the bioenergy research community with expertise in life sciences, life cycle assessment, sustainability assessment and social sciences) (Raman et al. 2014), patients (with dementia), relatives of patients, professional caregivers, volunteers, (de Jong et al. 2016), scientists using neuroimaging technologies, security professionals, neuroscientists, social psychologists, security professionals, development practitioners and an ethicist (de Jong et al. 2016) researchers (engineers and designers), innovation managers (universities, health organizations and biomedicine companies) (Demers-Payette et al. 2016) and researchers, bioethics experts in aquaculture, industry representatives, seafood technology experts, lifetime patent attorneys, government representatives, animal preservation NGO representatives, veterinarians and fish breeders (Bremer et al. 2015).

As far as the context in which the surveys are carried out, they were all carried out in developed countries, such as Germany, Canada, Denmark, Scotland, the Netherlands and the United Kingdom. This verification leads to a question: does the phenomenon of responsible innovation still not occur in emerging markets? Or has it simply not yet been properly researched?

When analyzing the RI literature, especially one that presents empirical studies with the participation of stakeholders, a fundamental issue arises. The stakeholders' input is advisory, not participating in the initial stages of innovation, when it has not yet been developed and there is more scope for adjustment and adequacy of the needs of those directly or indirectly involved. From the perspective of management, the anticipation of the participation of those involved can contribute to making economic gains more profitable, while the use of resources, mainly non-renewable resources, are spent moderately. From this analysis, evidence was highlighted and will be discussed in the next section.

5 Final Considerations

The purpose of this article was to analyze the role of stakeholders in the context of responsible innovation. Through a meta-synthesis, seven qualitative case studies were analyzed. We found that academic researchers and agents involved in multi-institutional projects encouraged stakeholder participation in the phase in which the product had already been developed. This finding raises the question of whether companies do not promote stakeholder participation from the earliest stages of innovation, or if there are no studies describing inclusion. Thus, the present study has discovered that RI is treated as a result and not as a process. However, it is important that innovation be treated as a process, which will make the whole procedure responsible, not just in the final stage. Inclusion as a process allows different visions of stakeholders to be considered, making the process responsible and ethical, better discussed and generating an innovation based on co-creation.

In addition to theoretical advancement, this study contributes to the practice of innovation, demonstrating and seeking to inform innovators and managers about the importance of multi-stakeholder participation throughout the innovation process.

As a limitation, we highlight the use of seven qualitative case studies, especially when compared to the 217 studies initially selected. However, as commented by Hoon (2013), being very inclusive in relation to embedded studies implies the risks of reducing the range of interpretations of a phenomenon. Another limitation is the range of areas of study that use the concept of responsible innovation, with several methodological models and that may have contributed to the exclusion of research with relevant results. It is also worth noting that the selection took place in September 2017, and other research may have been published after that date.

The development of new research is recommended, preferably empirical, that may help to better understand stakeholder participation, as well as to evaluate it in the context of emerging markets.

References

Bessant, J., Tidd, J.: Inovação e empreendedorismo: administração. Bookman Editora, Porto Alegre (2009)

Blok, V.: Look who's talking: responsible innovation, the paradox of dialogue and the voice of the other in communication and negotiation processes. J. Resp. Innov. **1**(2), 171–190 (2014). https://doi.org/10.1080/23299460.2014.924239

Bremer, S., Millar, K., Wright, N., Kaiser, M.: Responsible techno-innovation in aquaculture: employing ethical engagement to explore attitudes to GM salmon in Northern Europe. Aquaculture **437**, 370–381 (2015). https://doi.org/10.1016/j.aquaculture.2014.12.031

Burget, M., Bardone, E., Pedaste, M.: Definitions and conceptual dimensions of responsible research and innovation: a literature review. Sci. Eng. Ethics **23**, 1 (2017). https://doi.org/10.1007/s11948-016-9782-1

de Jong, I.M., Kupper, F., Broerse, J.: Inclusive deliberation and action in emerging RRI practices: the case of neuroimaging in security management. J. Resp. Innov. **3**(1), 26–49 (2016). https://doi.org/10.1080/23299460.2015.1137752

Decker, M., Weinberger, N., Krings, B.-J., Hirsch, J.: Imagined technology futures in demand-oriented technology assessment. J. Resp. Innov. **4**(2), 1–20 (2017). https://doi.org/10.1080/23299460.2017.1360720

Demers-Payette, O., Lehoux, P., Daudelin, G.: Responsible research and innovation: a productive model for the future of medical innovation. J. Resp. Innov. **3**(3), 188–208 (2016). https://doi.org/10.1080/23299460.2016.1256659

Eden, G., Jirotka, M., Stahl, B.: Responsible research and innovation: critical reflection into the potential social consequences of ICT. In: Proceedings - International Conference on Research Challenges in Information Science (2013). https://doi.org/10.1109/RCIS.2013.6577706

Hoon, C.: Meta-synthesis of qualitative case studies. Org. Res. Methods **16**(4), 522–556 (2013). https://doi.org/10.1177/1094428113484969

Mejlgaard, N., Bloch, C.: Science in society in Europe. Sci. Public Policy **39**(6), 695–700 (2012). https://doi.org/10.1093/scipol/scs087
Metze, T., Schuitmaker, T.J., Bitsch, L., Broerse, J.: Breaking barriers for a bio-based economy: interactive reflection on monitoring water quality. Environ. Sci. Policy **74**(May), 1–7 (2017). https://doi.org/10.1016/j.envsci.2017.04.015
Owen, R., Macnaghten, P., Stilgoe, J.: Responsible research and innovation: from science in society to science for society, with society. Sci. Public Policy **39**(6), 751–760 (2012). https://doi.org/10.1093/scipol/scs093
Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., Guston, D.: A framework for responsible innovation. Resp. Innov. Manag. Resp. Emerg. Sci. Innov. Soc. **31**, 27–50 (2013). https://doi.org/10.1002/9781118551424.ch2
Raman, S., Mohr, A., Helliwell, R., Ribeiro, B., Shortall, O., Smith, R., Millar, K.: Integrating social and value dimensions into sustainability assessment of lignocellulosic biofuels. Biomass Bioenergy **82**, 49–62 (2014). https://doi.org/10.1016/j.biombioe.2015.04.022
Ratiu, P., Mortan, M.: Ethical aspects of decision-making in agri-food enterprises. In: Popa, I., Dobrin, C., Ciocoiu, C.N. (eds.) Proceedings of the 7th International Management Conference: New Management for the New Economy, pp. 206–215. Editura Ase., Bucuresti (2013)
Shortall, O.K., Raman, S., Millar, K.: Are plants the new oil? responsible innovation, biorefining and multipurpose agriculture. Energy Policy **86**, 360–368 (2015). https://doi.org/10.1016/j.enpol.2015.07.011
Smallman, M., Lomme, K., Faullimmel, N.: D2.2. Report on the analysis of opportunities, obstacles and needs of the stakeholder groups in RRI practices in Europe. In: RRITools (2015)
Stahl, B.C.: Responsible research and innovation: the role of privacy in an emerging framework. Sci. Public Policy **40**(6), 708–716 (2013). https://doi.org/10.1093/scipol/sct067
Stilgoe, J., Owen, R., Macnaghten, P.: Developing a framework for responsible innovation. Res. Policy **42**(9), 1568–1580 (2013). https://doi.org/10.1016/j.respol.2013.05.008
Stirling, A.: "Opening up" and "closing down": power, participation, and pluralism in the social appraisal of technology. Sci. Technol. Hum. Values **33**(2), 262–294 (2008). https://doi.org/10.1177/0162243907311265
Tidd, J., Bessant, J.: Gestão da inovação-5. Bookman Editora, Porto Alegre (2015)
Yenicioglu, B., Suerdem, A.: Participatory new product development–a framework for deliberately collaborative and continuous innovation design. Procedia Soc. Behav. Sci. **195**, 1443–1452 (2015). https://doi.org/10.1016/j.sbspro.2015.06.442

Different Perspectives for Combining Exploration and Exploitation Strategies

Leila Beig[1(✉)] and Atefeh Ghavamifar[2]

[1] Business School, University of Stavanger, Stavanger, Norway
l.beig@stud.uis.no
[2] ICT Research Institute, Tehran, Iran
a_ghavamifar@itrc.ac.ir

Abstract. In the twenty first century, more than any other time in the firm's history, companies are under the global competitive pressures in a context that is widely dynamic and consequently makes them encounter a huge number of unpredictable changes. Leaders of firms, on one hand face the immediate pressures of delivering value to increasingly sophisticated and globally diverse customers while accelerating the return on these efforts for financial stakeholders and on the other hand, strategic leaders must identify and prepare for disruptive technologies and emerging market opportunities over the long-term. Therefore, combining two different innovation strategies, exploration and exploitation is a huge challenge for managers and leaders of the companies. The literature review shows that this phenomenon forms a new type of organizations so called ambidextrous organizations. In these paper, based on related studies six perspectives have been described and compared. At the end, Laerdal Medical AS company has been examined as a case study to show how these perspectives work in practice in an organization.

1 Introduction

There are two strategies for innovation, exploitation strategy and exploration strategy. The process of taking advantage of what exists, allocating the resources to improve the existing products and processes and thinking about today's customers is what (March 1991) calls exploitation strategy. In contrast, exploration concentrates on future consumers and addresses the process of experiencing novel methods of doing things (March 1991). This is related with the opportunities of development beyond organizational limits and borders and, therefore, encompasses scanning the environment in which the firm seeks to absorb new knowledge (Marín-Idárraga et al. 2016).

When it comes to innovation, exploitation is associated with incremental innovation and exploration with radical innovation. As Marín-Idárraga et al. (2016) states, exploiting existing competences and knowledge and internal resources can lead to incremental innovations, given that they involve small alterations in technology and little deviation from the current product market experiences. Conversely, the exploration of new markets and technologies will produce radical innovations, as they go through fundamental and sometimes revolutionary changes in the firm's technologies and markets (Marín-Idárraga et al. 2016).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 836–845, 2020.
https://doi.org/10.1007/978-3-030-48021-9_92

From the marketing point of view, exploratory innovation strategies are good at opening up new markets, reach new customers, and generate new business opportunities. In contrast, exploitive innovation strategies typically concentrate on making small improvements to existing products and require much less R&D spending, and as a consequence serve as a more immediate and short-term source of income (Kollmann and Stockmann 2012).

Because of the benefits that organizations can get by implementing each strategy and in order to be pioneer in long term from one hand and be efficient in short term from the other hand, managers have a tendency to deploy both strategies and take advantages of both. This is the moment that challenges for different levels of authority in an organization emerges one after another when the managers have to be capable of balancing these two extensively different strategies in their organization.

To reconcile this dilemma, Judge and Blocker (2008) proposed the new concept "strategic ambidexterity", which is conceptualized as the ability to simultaneously pursue exploitation and exploratory strategies in ways that leads to enhanced organizational effectiveness.

"Organizational Ambidexterity" term was first introduced by Duncan in 1976 which refers to new form of organization with dual structures inside (Duncan 1976). Organizational Ambidexterity did not get a lot of attention and wide interest until a seminal paper about organizational learning and exploration and exploitation processes by March (1991). He described exploration and exploitation as two fundamentally different strategies, whereby exploitation addresses "refinement, efficiency, selection and implementation" and exploration focuses on "search, variation, experimentation and innovation" (March 1991).

Having both exploration and exploitation strategy at the same time is very challenging and makes a lot of tension and competition in resource allocation. Therefore, the purpose of this essay is to find the different perspectives which contribute to resolve related challenges to some extent. In the following section a literature review about ambidexterity, different definitions, strategies and logics will be presented.

2 Literature Review

Probing into literature about ambidexterity, ambidextrous organizations and organizational ambidexterity shows that there are different definitions for ambidexterity from a variety of point of views. From marketing standpoint, "ambidexterity is the capability to operate in both mature markets and develop new products and services for emerging markets". He and Wong (2004) and Im and Rai (2008) defines it form an organizational learning lens as "an organization which is capable of conducting two paradoxical things at the same time by requiring organizations and their people to have two heterogeneous but related skills simultaneously". Li et al. (2008) believe that from an innovation management perspective these kinds of organizations are "those like a juggler who can handle numerous balls at the same time, could rely on given resources to ensure radical innovations as well as incremental innovations simultaneously". Regarding the strategic management perspective, ambidextrous firms are capable of exploiting existing competences as well as exploring new opportunities (Lubatkin et al.

2006). Simultaneous implementation of exploration and exploitation is very challenging and strategic importance of balancing exploitation and exploration has been extensively discussed in the organizational structure and strategic resource allocation literature.

Tushman and O'Reilly (1996) recommended **separate structures** within the same organization as a solution for the conflict raised because of simultaneous implementation of exploration and exploitation strategies in an organization. They suggested small, decentralized and with loose processes for exploratory units and, in contrast, larger, centralized and with tighter processes as exploitative units. O'Reilly and Tushman (2004) highlighted the role of the top management team as the 'corporate glue' that keeps the organization together by handling the tensions that emerge between these two units.

In comparison to the previous solution, Raisch and Birkinshaw (2008) address **parallel structures**. Parallel structures provide a situation for a single business unit to switch between structures regarding to the need for exploration or exploitation (Devins 2010). Parallel structures have also called as "collateral" organizations in Zand (1974) or "shadow" organizations in Goldstein (1985).

Other strategies to managing exploration and exploitation tensions points out **temporal separation** whereby an organization switches between exploration and exploitation sequentially (Jansen et al. 2005). During temporal balancing the structure of the organization shifts from a mechanistic structure (focusing on centralization) to an organic structure (allowing decentralization) as organizations switch from exploitative to explorative strategies, respectively (Devins 2010). Two concepts were used in this method, punctuated equilibrium and 'organizational vacillation'. Punctuated equilibrium which is a term which originates from an evolutionary biology theory in organization refers as a movement between periods of exploration and periods of exploitation. Similarly, Boumgarden et al. (2012) refer to 'organizational vacillation' to describe firms' dynamic capability of temporally and sequentially alternating between periods of exploration and exploitation. In temporal balancing, time plays a crucial role in how organizational ambidexterity is conceptualized and how exploration and exploitation tensions evolve over time (Uotila et al. 2009).

Another approach to ambidexterity is **contextual approach** in which the systems and structures are more flexible, allowing employees to make their own decision as to how they divide their time between two various activities of exploration and exploitation (Birkinshaw and Gibson 2004). To improve this type of ambidexterity on the individual level, a much greater level of attention has to be paid to the human resources of the organization. Since top managers form organizational context through the performance systems, incentives and control mechanisms, hence they play a crucial role in shaping the context.

O'Reilly and Tushman (2004) defines **individual-level ambidexterity** based on the premise that ambidextrous organizations require to recruit ambidextrous individuals who are capable of handling both exploration and exploitation practices simultaneously. Birkinshaw and Gibson (2004) believes that Ambidextrous individuals "take the initiative and are alert to opportunities beyond the confines of their own jobs. They are cooperative and seek out opportunities to combine their efforts with others. They are brokers, always looking to build internal linkages and they multitaskers who are comfortable wearing more than one hat." Similarly, Mom et al. (2009) propose these properties for ambidextrous managers: multitaskers who are able to welcome conflicts, and combine and make new knowledge, skills, and expertise and make a new experience. Eisenhardt et al. (2010) explored that ambidextrous leaders use simple, cognitive strategies to address the complex tensions of exploration and exploitation.

There is another perspective to ambidexterity so called **Peripatric ambidexterity**, a term borrowed from genetics which was first used by Lubatkin et al. (2006). The main logic of this approach is that an ambidextrous leader as the focal point absorbs others of the same characteristics and they together collaborate and make a community that forms ambidextrous organization. In leadership literature this phenomenon is referred to 'human trait of ambidexterity' (Lubatkin et al. 2006). The followers of trait theory of leadership believe that leaders are born not made. In this theory, six attributes make leaders different from non-leaders including motivation, the tendency to lead, honesty and integrity, self-confidence, intellectual capabilities, and business knowledge and expertise (Germain 2012). This strategy is very similar to the previous strategy. They are both of the same assumption that individuals play a critical role in making an organization ambidextrous, the only difference is that the peripatric strategy emphasize more on the capability of an ambidextrous leader in absorbing the other members of management team while in the previous strategy the human resources were looked at from an individual window focusing on unique characteristics of these kinds of persons. Figure 1 illustrates proposed strategies by different scholars.

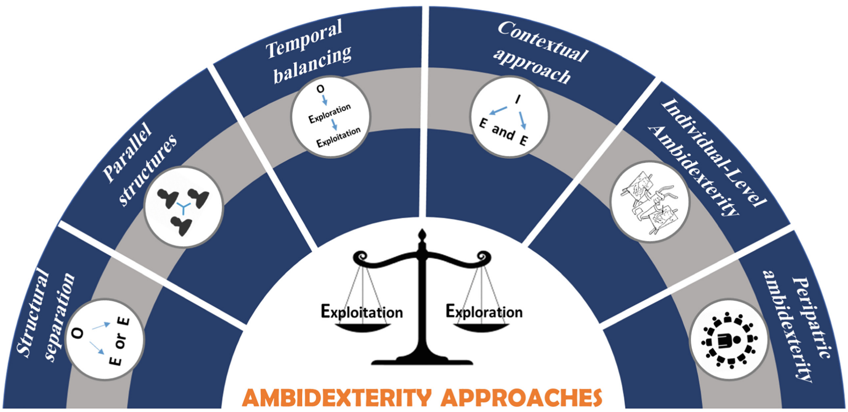


Fig. 1 Different strategies of ambidexterity

3 Case Study: Laerdal Medical AS

In this section Laerdal Medical AS was considered as a case study and we examined the way exploration and exploitation strategy are combined in this company according to above mentioned perspectives.

3.1 About Laerdal Medical AS

Åsmund S. Lærdal established his business in Stavanger in February 1940 as a publisher of cards and books, and later toys. With an eye for innovation, in 1960, Lærdal recognized the potential for using his expertise in producing plastic dolls for a new purpose - the development of a realistic training manikin to teach mouth-to-mouth resuscitation. Over the next 55 years, Laerdal Medical developed a broad range of products and programs to support resuscitation training and emergency interventions. In 2000, with a growing focus on increased patient safety, Laerdal broke ground in the field of medical simulation, introducing a series of relatively low cost realistic patient simulators, allowing for risk-free interactive training in emergency care. The Laerdal Foundation for Acute Medicine was established in 1980 in collaboration with the University of Oslo. Donations from Laerdal Medical have enabled it to support a large number of practically oriented research projects. In 2010, Laerdal Global Health (LGH) was established as a not-for-profit company to develop high impact, low cost training and therapy products aimed at helping save the lives of newborns and mothers in low-resource countries. Today, Laerdal Medical is a global company with more than 1400 employees in 24 countries dedicated to helping save lives through the advancement of resuscitation, emergency care, and patient safety. In collaboration with renowned researchers and prominent partners, Laerdal aims to continue to improve and innovate to reach their ultimate mission "helping save more lives" and their vision "No one shall die or disable unnecessary in connection with childbirth or as a result of sudden illness, serious accident or patient injury. They have set helping save 500,000 more lives, each year, by 2020 as their main goal" (Laerdal 2018).

3.2 Laerdal Journey to Ambidexterity

Laerdal is an innovative company pursuing incremental and radical innovation in their products and services. In terms of IP rights their portfolio includes 60 active patent families (430 cases), 18 active design families (130 cases) and 50 Trademark families (540 cases). They have been successful in changing their structure to make them able of taking the full advantages of combining the exploration and exploitation strategy. Below the findings of an interview with Petter Westnes, Corporate QA, RA and IP Director at Laerdal Medical manager of Laerdal IP department has been presented. The interview was based on seven main questions as it follows.

- **How was the structure of Laerdal before changes?**

Laerdal had two completely different and physically separated departments including product management department and product development department. The first one consisted of product specialists who focused on managing the products. Their responsibilities include putting together marketing packages, supporting the sales' companies in selling the products and services, as well as implementing the product into the market. They focus very much on helping to understand the customer segments and their requirements.

Product development department, in contrast, encompasses many specialized engineers and designers. for example, they had a team of electrical engineers, a team of mechanical engineers, a team of plastic engineers and a team of mechatronic engineers. According to this structure, product management department identified needs at the market for new products or for changes to existing products, and then they ask the product development department to either perform a project or perform changes to the product.

- **Why they started the journey to ambidexterity?**

In Laerdal they believe that any company should take both exploration and exploitation strategy for survival. Five years ago, they came to this point that the above-mentioned structure is not a very entrepreneurial approach. It was too complex and the distance between product development and product management departments which in fact required to work together was too big. Both physically because they typically located in different physical spaces in the office as well as mentally. There was no team feeling. Hence, they came to this point that there is a need to make smaller more entrepreneurial teams that were able to act quicker and getting closer to the customer to understand the customer needs. That is the major reason that they switched from product development group and product management group to a new structure including business units.

- **How is the structure at present?**

From five years ago, Laerdal has three various business units including Laerdal global health, patient care, and resuscitation. They focus on three different product areas. Within each business unit there are segment teams. Teams focusing on very specific user segment meaning customer groups or product groups. Within these teams they put together three different roles 1 - product managers, 2 - engineers and 3 - service designers which is based on a new role defined in the novel structure. In this structure, not only are the teams smaller and more focused but also customer groups or product groups are physically put together to create the team feeling. The purpose is to get closer to the customer groups as well as customer requirements and wants. In this way Laerdal will be able to have a quicker turn around and work quicker in a more entrepreneurial or innovative approach to both exploiting the existing customer and products and exploring the new opportunities.

- **What challenges did they encounter?**

It was a radical shift for Laerdal and of course any change is difficult for the people working in Laerdal for a long time because in the new structure they took away the mechatronic team, the mechanical team, and the engineering team so they instead of being specialist within their area and being part of the specialist group, they are supposed to be in a different kind of environment which is much more focused on customers and product groups and not only on their expertise.

- **Did they consider any changes in employing human resource?**

Laerdal is no longer just manufacturing physical products to sell to the customer but also it accompanies the customers through their whole journey they go. It provides service support and implementation, training, learning material, online content, etc. Since they add all these different aspects to the physical products, the service will get complicated, and therefore, they require different set of capabilities than just design or engineer a physical product and commercialization of that product. There is one type of employees they are hiring more in the last couple of years, service designers. They hire more and more of this position, people who look at the whole service delivery process to the customer, trying to understand the customer journey. The process has technical aspects as well as more commercial or marketing/implementation aspects. That is an important dimension and it is all about the way we deliver products to the market. The way we understand the journey that the customer is going through is becoming more and more important.

- **Did they consider any changes in the process of marketing and identifying customer needs?**

The team members are required to spend a certain amount of their time together with customers. Hence, Laerdal wants its people, for example the traditional engineers who just used to sat on their office and do engineer stuff, to spend time with customers. They believe that it is also an important part to become more entrepreneurial. They also believe that if we are supposed to design something that the customer needs, we really need to understand the customer and the only way is to spend as much time as possible with the customer.

- **What are the outcomes of restructuring?**

The outcome of this restructuring is, first the product and services Laerdal places in the market now are more fit for customers' purposes. They are able to have a greater fit with what the customer really requires and willing to pay for and having an impact on the customer. In addition, they are also better at delivering projects on time and on cost. In a better word, they can better identify and plan for existing customers' needs as well as new opportunities which they might create for customers or the customers might demand.

4 Conclusion

In this paper we attempted to find different strategies for combining exploration and exploitation strategies in companies. The main logic of these strategies as well as their main focus are completely different. A couple of perspectives such as structural separation, parallel structure and temporal balancing, consider these two different strategies, as contradictory poles. Hence, they look at them as situations making conflict in resource allocation and management. In contrast perspectives such as individual level ambidexterity and peripatric ambidexterity believe they are complementary poles, and contextual approach state that they are independent poles. Interpretation of the ambidexterity as a concept essentially is of importance because it determines individuals' and organizations' next movements, behaviors and consequently their decisions they make. Findings show that firms that consider them interrelated and complementary are more effective in reaching ambidexterity and resolving the probable problems.

As Table 1 shows the focus of the perspectives is also different. Some go through combining these two strategies by changing organizational design, some by adaptation through time, some considering contextual aspects of organization, focusing on characteristics of individuals or capability of making ambidextrous team.

Table 1. A comparison of different strategies for simultaneous exploration and exploitation in a frim

strategies	Definition	Main logic	Main Focus
Structural separation	Separate structures within the same organization: small, decentralized and with loose processes for exploratory units and, in contrast, larger, centralized and with tighter processes as exploitative units.	Contradictory poles	Organizational design
Parallel structures	A situation for a single business unit to switch between structures regarding to the need for exploration or exploitation.	Contradictory poles	Organizational design
Temporal balancing	The structure of the organization shifts from a mechanistic structure (focusing on centralization) to an organic structure (allowing decentralization) as organizations switch from exploitative to explorative strategies, respectively.	Contradictory poles	Organizational adaptation
Contextual approach	Simultaneous capability of the human resources toward the alignment and adaptability.	Independent poles	contextual aspects of organization
Individual-Level Ambidexterity	Recruiting ambidextrous individuals who are capable of handling both exploration and exploitation practices simultaneously.	Complementary	Characteristics of individuals
Peripatric ambidexterity	An ambidextrous leader as the focal point to absorb others of the same characteristics and they together collaborate and make a community that forms ambidextrous organization.	Complementary	Capability of making ambidextrous team

New structure designs can be helpful, but they depend on different factors in organizations. For example, if an organization applies a separated structure as a strategy, integration in different levels should also take into consideration because separated units in long term might convert into two totally different isolated islands that there is no relationship between them and no alignment to the organization's vision. As explained in contextual ambidexterity, providing an appropriate balanced context can support ambidexterity, otherwise, it might contradict the main organizational goal. Having an appropriate human resource development model including required trainings can have a considerable effect on promoting the personnel's capabilities in terms of ambidexterity. It might be a good idea to recruit ambidextrous individuals in an organization, but the important point is that a supportive culture and leader's support which makes a suitable environment for this kind of people to perform effectively. Establishing strong external networks with customers, supplier and other elements of company's value chain, from one hand, and internal networks from the other hand, is also recommended because it fosters knowledge sharing which is the beating heart of both innovation strategies (exploration/exploration). In this case absorptive capacity is plays a considerable role which means that if it is higher in an organization that organization can benefit more from using networks.

Laerdal, the case study of this paper, for becoming ambidextrous and combining these two strategies did changes in their structure design. They moved from structure separation which was the first perspective to business units of teams including technical and marketing members. They did these changes consciously to move them closer to customer, understand their needs better and fulfill their requirements in a more effective and efficient way not only the existing customer needs but also new customers that might emerge. They also did some changes in human resource they add one more category which they don't call ambidextrous individuals, but the role is a combination of technical and marketing roles. In this way the outcome was better fit for customers' purpose and more cost efficient and on time project management.

All in all, it can be concluded that following factors are very important in solving the challenges organizations encounter in simultaneous implementing of both strategies. Having a clear vision, a supporting organizational culture, recruiting ambidextrous individuals, having an appropriate human resource development model, having an integration team/steering committee in organizational design, monitoring regularly the alignment of strategies with the organization vision, make a competitive environment for funding of exploration or exploitations projects based on related performance indicators, having a team observing the environment dynamics and market orientation. It is worth mentioning that all of the above recommendations are based on literature review and should be defined as hypothesis and be tested in different organizations to be more reliable because each organization has its own culture, context, human resource, management and leadership style vision, mission, social interactions and knowledge assets which play a part in reaching the final results.

References

Birkinshaw, J., Gibson, C.: Building Ambidexterity into an Organization. MITSloan Management Review 45 (4), (2004)

Boumgarden, P., Nickerson, J., Zenger, T.R.: Sailing into the wind: exploring the relationships among ambidexterity, vacillation, and organizational performance. Strateg. Manag. J. **33**(6), 587–610 (2012)

Devins, G.: Structuring ambidextrous organizations: exploitation and exploration as a key for long-term success. In: Stadtler, A.S.L. (ed.) More than Bricks in the Wall: Organizational Perspectives for Sustainable Success, pp. 60–67. Springer Gabler, Berlin (2010)

Duncan, R.B.: The ambidextrous organization: designing dual structures for innovation. The Management of Organization **1**(1), 167–188 (1976)

Eisenhardt, K., Furr, N., Bingham, C.: Microfoundations of performance: balancing efficiency and flexibility in dynamic environments. Organ. Sci. **21**, 1263–1273 (2010)

Germain, M.-L.: Traits and skills theories as the nexus between leadership and expertise: reality or fallacy? Perform. Improv. **52**(5), 32–39 (2012)

Goldstein, S.G.: Organizational dualism and quality circles. Acad. Manag. Rev. **10**(3), 504–517 (1985)

He, Z., Wong, P.: Exploration vs. exploitation: an empirical test of the ambidexterity hypothesis. Organ. Sci. **15**, 481–494 (2004)

Im, G., Rai, A.: Knowledge sharing ambidexterity in long-term interorganizational relationships. Manage. Sci. **54**, 1281–1296 (2008)

Jansen, J.J., Bosch, V.D., Volberda, H.W.: Exploratory innovation, exploitative innovation, and ambidexterity: the impact of environmental and organizational antecedents. Schmalenbach Bus. Rev. (SBR) **57**(4), 351–363 (2005)

Judge, W.Q., Blocker, C.P.: Organizational capacity for change and strategic ambidexterity: flying the plane while rewiring it. Eur. J. Mark. **42**(9), 915–926 (2008)

Kollmann, T., Stockmann, C.: Filling the entrepreneurial orientation–performance gap: the mediating effects of exploratory and exploitative innovations. Entrep. Theory Pract. **38**(5), 1001–1026 (2012)

Laerdal. www.laerdal.com. Accessed from 30 July 2018

Li, C., Lin, C., Chu, C.: The nature of market orientation and the ambidexterity of innovations. Manag. Decis. **46**, 1002–1026 (2008)

Lubatkin, M.H., Simsek, Z., Ling, Y., Veiga, J.F.: Ambidexterity and performance in small- to medium-sized firms: the pivotal role of top management team behavioral integration. J. Manag. **32**, 646–672 (2006)

March, J.G.: Exploration and exploitation in organizational learning. Organ. Sci. **2**(1), 71–87 (1991)

Marín-Idárraga, D.A., Hurtado, J.M., Medina, C.C.: The antecedents of exploitation exploration and their relationship with innovation: a study of managers' cognitive map. Creativity Innov. Manag. **25**(1), 18–37 (2016)

Mom, T.J.M., Van Den Bosch, F.A.J., Volberda, H.W.: Understanding variation in managers' ambidexterity: investigating direct and interaction effects of formal structural and personal coordination mechanisms. Organ. Sci. **20**, 812–828 (2009)

O'Reilly, C., Tushman, M.: The ambidextrous organization. Harv. Bus. Rev. **82**, 74–81 (2004)

Raisch, S., Birkinshaw, J.: Organizational ambidexterity: antecedents, outcomes and moderators. J. Manag. **34**(3), 375–409 (2008)

Tushman, M., O'Reilly, C.: Ambidextrous organizations: managing evolutionary and revolutionary change. Calif. Manage. Rev. **38**, 8–30 (1996)

Uotila, J., Maula, M., Keil, T., Zahra, S.A.: Exploration, exploitation, and financial performance: Analysis of S&P 500 corporations. Strateg. Manag. J. **30**(2), 221–231 (2009)

Zand, D.: Collateral organization: a new change strategy. J. Appl. Behav. Sci. **10**(1), 63–89 (1974)

Value Creation Mechanisms of Modularisation in the Engineering Asset Life Cycle

Jarkko Pakkanen(✉), Tero Juuti, and Timo Lehtonen

Laboratory of Mechanical Engineering and Industrial Systems,
Tampere University of Technology, 33101 Tampere, Finland
{jarkko.pakkanen,tero.juuti,timo.lehtonen}@tut.fi

Abstract. Many companies offering physical assets have to adapt to different market requirements to maintain profitability. Product modularisation is a common solution for this challenge used by suppliers (manufacturers) of engineering assets. Modularisation enables greater product variety and increases commonality between product variants. Modularisation includes defining a set of modules, interfaces, modular architecture and configuration rules and constraints based on case specific partitioning logic. This paper reviews the main value creation mechanisms (VCMs) of product modularisation in the manufacturing industry, and studies what kind of VCMs are related to the main life cycle stages of engineering assets and how companies in case studies have incorporated VCMs. Key VCMs were identified based on the engineering asset life cycle, but other VCMs were considered to be important from a supplier's perspective. Suppliers should consider the whole life cycle when designing engineering assets and clarify which VCMs are the most important guiding principles for their product and make trade-offs when required.

1 Introduction

Engineering assets, such as equipment and buildings, are managed by engineering asset managers, and are needed to create financial assets, capability value and financial value (Amadi-Echendu et al. 2010). An engineering asset life cycle typically consists of commissioning, operations, maintenance, decommissioning and replacement (Haider 2007). From the supplier's perspective, engineering assets are considered to be physical equipment, in which case the stages of the asset life cycle become part of the product life cycle. This paper focuses on the supplier's perspective by including additional stages of the product life cycle in the asset life cycle. These other stages common in the manufacturing industry include product development, marketing and sales, production and transportation.

Product modularisation (referred to simply as modularisation) is a common product development approach that enables suppliers to effectively offer product variety. Clarifying the partitioning logic of the module system and defining modular architecture, including the set of modules, interfaces, and configuration rules and constraints, are the key aspects of modularisation (Pakkanen et al. 2016). Modular architecture and defined interfaces support the interchangeability and independence of modules that are the building blocks of product variety. Modularisation is typically

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 846–854, 2020.
https://doi.org/10.1007/978-3-030-48021-9_93

aimed at reducing the complexity in the supplier's operations and harvesting positive benefits in certain areas of the business (Andreasen 2011). Consequently, it is important to clarify the areas and asset life cycle stages where the benefits could be realised.

This paper presents the types of value creation mechanisms (VCMs) that influence modularisation and how VCMs relate to the main life cycle stages of engineering assets. VCMs are helpful for describing aspects of the life cycle that modularisation can affect. As a baseline for studying and linking VCMs, we specifically considered the asset life cycle from the supplier's perspective. The purpose of this kind of comprehensive review of the asset life cycle was to study what motivated suppliers to realise specific asset structures. Throughout the literature review, we focused on the following research questions (RQ):

- RQ1: What are the main VCMs of modularisation in the manufacturing industry?
- RQ2: How did case study companies treat VCMs related to essential life cycle stages of engineering assets in the target setting of a modularisation project?

Section 2 of the paper presents the literature review mainly focusing on Scopus and Google Scholar, which adequately covered many potential sources in this area of research. We used a number of different combinations of search terms and focused on journal articles, books, book chapters and conference papers related to modularisation. Out of hundreds of search results, the titles of papers and abstracts with the highest number of citations were used to narrow down the relevant publications to be studied in greater detail. Nonetheless, there is a risk that valid references were missed during the review because of the page limitation of this paper. Section 3 of the paper presents case studies revealing how companies viewed VCM's in modularisation. Along with further discussion, answers to the RQs are summarised in Sect. 4.

2 Value Creation Mechanisms of Modularisation

This section presents relevant VCMs of modularisation determined from the literature review. First, generic life cycle stages of the manufacturing industry are discussed, followed by the reasoning for defining the VCMs related to the modularisation of each life cycle stage. At the end of each life cycle subsection is a bulleted summary of the main VCMs related to each stage.

Incremental redesigning or new product development are needed when existing product variety contradicts business goals or the cost of offering product variety is unsustainable. Companies can implement redesigning by developing new products, product families or features, or by updating or correcting designs, changing commercial components or updating technologies. Product development is typically followed by marketing and sales, which relies on technical support and knowledge related to product variety. Production often includes procurement of components and materials, subcontracting, manufacturing, testing of parts and modules, assembly and final testing. Transportation is also an important aspect when considering modularisation because reaching the customer in a profitable way may require different logistics. Products might require commissioning in the operating site before being used. Operations are followed by maintenance, decommissioning and replacement. These are the seven stages of the manufacturing industry life cycle that are discussed in more detail below.

2.1 Product Development

Modular design requires investment because of the costs associated with formulating design rules, experimenting, designing modules and testing (Baldwin and Clark 2000). Consequently, R&D investment is one of the major mechanisms that needs to be considered for modularisation to succeed. Design reuse reduces the engineering effort needed for product design (Pulkkinen 2007), which results in benefits related to cost, quality and time. Reusing parts, sub-systems, modules or interfaces are some of the possible drivers for modularisation being able to reduce design efforts (Erixon 1998; Pakkanen et al. 2016). Reusing component designs reduces the development costs of new product variants (Sanchez 1999). Hence, design reuse can be done at different levels and it is an important VCM of modularisation.

Technology is another aspect of modularisation that is important. Technological evolution or technology push can set requirements for products whose technology develops fast (Erixon 1998). Future product elements should be considered in defining product family architecture (Harlou 2006). Encapsulated technologies are discussed in the modularisation context. Encapsulation of modules makes product management easier (Lehtonen 2007) and isolating design parameters from other parts of the design makes incorporation of new solutions simpler (Baldwin and Clark 2000). We summarised these findings in five VCMs related mainly to product development.

- R&D investment
- More capacity for new product development
- Design by reuse (parts, sub-systems, modules, interfaces)
- New technologies
- Encapsulated technologies

2.2 Marketing and Sales

Modularisation affects the product cost. Variety can create additional costs for several life phases, unless commonality of the products is considered (Andreasen 2011). Use of common components (Sanchez 1999) and part count reduction (Fixson 2006) are typical cost reduction tools. One factor in analysing the demand for offering variants based on modularity is brand effect (Hopp and Xu 2005). Enabling configurable products contributes to brand management (Tiihonen et al. 1999), but a modular product family may not be designed very often on a large scale within a company because it requires large investments. While designing a modular product family, the future attractiveness of the products must also be taken into consideration (Harlou 2006). Styling possibilities can be an important driver for modularisation for products whose purchasing decision strongly relies on trends and fashion (Erixon 1998). Modularisation supports matching products to customer needs through product configuration which defines product variants based on predefined rules and product elements (Pulkkinen 2007). Product configuration knowledge and human or software driven configurations have been discussed by several authors as a facilitator of sales, distribution tendering and product variant definition (Soininen 2000; Haug et al. 2012). Based on these explanations, the following VCMs are connected to marketing and sales.

- Product cost
- Brand impact
- Reactivity to market changes
- Product fit to customer needs, standards and local legislation
- Support for sales and distribution tendering or product variant definition

2.3 Production

Delivery-specific products are more challenging than standard products because there are fewer opportunities to learn lessons from repetition. Designing modular components simplifies and increases the predictability of production (Sanchez 1999; Baldwin and Clark 2000). Production includes many stages, such as part procurement. If common parts can be used in a product range, economies of scale in part sourcing may exceedingly reduce the total cost of production (Baldwin and Clark 2000). Late point product differentiation by allocating variations to only one or a few parts, and keeping the product generic or as "universal chassis" as long as possible, lowers the buffer inventories in production, and thus reduces the complexity of the manufacturing process and overall costs (Lee and Tang 1997; Erixon 1998; Sanchez 1999). The use of universal chassis may reduce the variety of parts to be inventoried and handled in assembly, and thus reduce product costs (Sanchez 1999). Existing assembly and operating environment reasons can favour distributed production in separate factories (Lehtonen 2007). Clearly defined building blocks and interfaces support this. Standardised interfaces of modular products enable delivering product modules to a customer site in which the product is to be assembled effectively (Lau Antonio et al. 2007). Separate testing of modules decreases the feedback time about the quality of modules compared to testing done in the main production flow (Erixon 1998). The reliability of key components can be improved over time because reusing parts enables improvements in materials and processing, which reduces warranty costs (Sanchez 1999). In product modularisation, companies aim to find proven solutions. Product configuration relates to managing product quality because it focuses on defining product variations based on certain predefined rules, limitations and standardised elements (at least within a company) and therefore, the risk of mistakes is minimised (Tiihonen et al. 1999; Juuti 2008). These aspects lead to a number of VCMs related to production.

- Improving controllability of production through transparency and predictability
- Component availability and number of supply sources
- Late point differentiation
- Relocating production to more favourable areas or environments
- Decreased ramp-up time and costs with distributed module testing
- Reduction in component quality issues
- Distributing or decentralising assembly
- Reduction in product quality issues

2.4 Transport

Products can be designed to match the relevant type of logistics. Product architectural decisions determine the needed packing space and protection requirements of the product in relation to logistics, including transportation (Fixson 2006). Standardised modules with less variety allow for lower logistics costs (Erixon 1998).

- Decreased transportation costs

2.5 Commissioning

Reductions in components and product quality issues were already discussed in other stages of the asset life cycle, and they are also very relevant in the commissioning stage. Increased commonality enables repetition in operational processes, which makes commissioning more effective (Juuti 2008).

- Reduction in component quality issues
- Reduction in product quality issues
- Effective commissioning by learning

2.6 Operations

Modularity in product use enables customers to reorganise elements of the final product to match their tastes or needs by making substitutions, augmentations or exclusions (Baldwin and Clark 2000). Accessibility to different types of interfaces strongly affects the ability to customise the product.

- Ability to reconfigure the product during use

2.7 Maintenance, Decommissioning and Replacement

Modular structures and standardised interfaces support maintenance. Modularisation can make it easier to replace possible damaged areas (Erixon 1998). Companies doing maintenance-oriented modularisation should focus on designing the product in a way that ensures the replacement of working components is avoided when maintenance is needed in other components (Umeda et al. 2000). Proper designing can create new business opportunities for companies by providing services that reduce the amount of waste by reusing the modules (Umeda et al. 2000).

- Reduced down time and maintenance costs by using replaceable modules
- Increased end-of-life value by reusing modules

3 VCMs in Case Studies in the Machinery Industry

Table 1 presents the VCMs that three companies highlighted during target setting of a modularisation project. Large Companies A, B and C operate globally in the machinery industry, but do not compete with each other. Names of the companies were kept anonymous to respect the sensitive nature of their business. The data was collected from 2010–16 by authors who participated as facilitators and consultants in design projects (Companies A and C) or interviewed key company personnel (Company B). During the projects, companies were not shown the list of VCMs, but authors analysed retrospectively these case studies according to their observations and notes. Therefore, the results could have been different if companies would have been asked to select VCMs they considered important to their business. When considering the key life cycle stages of engineering assets (commissioning, operations, maintenance, decommissioning and replacement) (Haider 2007), Company A highlighted that the ability to replace modules and to reuse specific modules when making new products would be valuable. Company A managed to reduce the number of items and different technologies for similar functions without losing flexibility in the commissioning phase to adapt to different customer environments. Rationalised items allowed the company to be more consistent with its product offerings because it could reuse the same items in different product variants. They reported a 25% cost savings in some variants. Company B had a different focus. They strived for quality improvements by using modularisation and thought it would be valuable if products could be reconfigured during operational changes. Company B argued positively that after a modular product family has been designed and a new operational model is fully implemented, the investment costs decrease every year; they did highlight that this whole process could take many years until achieving its full potential in cost savings, such as site works related to commissioning. Company C noted that modularisation could decrease product quality problems through proven and predefined solutions (defined outside delivery project). This project is ongoing and results are not available yet. These case studies show that the companies had different motivations for the way the physical assets were structured and that suppliers/producers of assets deal with multiple sub-goals that usually lead to making trade-offs.

Table 1. VCMs in a target setting of modularisation (x: considered)

Life cycle stages and VCMs	Company A	Company B	Company C
Product development			
Research and development investment	x	x	x
More capacity for new product development	x	x	x
Design by reuse (parts, sub-systems, modules, interfaces, other)	x	x	x
New technologies	x	x	x
Encapsulated technologies			
Marketing and sales			
Product cost	x		
Brand impact			
Reactivity to market changes	x		
Product fit to customer needs, standards and local legislation	x	x	x

(*continued*)

Table 1. (*continued*)

Life cycle stages and VCMs	Company A	Company B	Company C
Support for sales and distribution tendering or product variant definition	x	x	x
Production			
Improving controllability of production through transparency and predictability	x	x	
Component availability and number of supply sources			
Late point differentiation	x		
Relocating production to more favourable areas or environments			
Reduction in component quality issues		x	
Decreased ramp-up time and costs with distributed module testing and learning		x	
Distributing or decentralising assembly			
Reduction in product quality issues			x
Transport			
Decreased cost of transportation			
Commissioning			
Reduction in component quality issues		x	
Reduction in product quality issue		x	x
Effective commissioning by learning		x	
Operation			
Ability to reconfigure product during use		x	
Maintenance, decommissioning and replacement			
Reduced down time and maintenance costs by replaceable modules	x		
Increased end-of-life value by reusing modules	x		

4 Discussion

The purpose of the paper was to study (RQ1) the main VCMs of modularisation in the manufacturing industry, and (RQ2) how case study companies related VCMs to essential life cycle stages of their engineering assets in a target setting of a modularisation project. VCMs presented in Sect. 2 and the same VCMs shown in the first column in Table 1 are the answers to RQ1. The VCMs of modularisation describe possible objectives, phenomena and problems that can arise from modularisation. These VCMs derived from the studied literature consider the whole engineering asset life cycle. The answers to RQ2 show that case study companies consider VCMs differently. All companies in the case studies had a strong focus on VCMs related to product development and marketing and sales, but not all VCMs were relevant.

Further differences could be found when considering life cycle stages from production to end-of-life. The case studies show that different objectives may affect modularisation and the resulting asset structure. For example, modularisation can be done in a way that it a) supports asset suppliers by improving asset quality, b) enables reconfiguring of assets, c) facilitates maintenance, and d) increases possibilities to reuse some elements of the assets when the asset as a whole reaches the end of the life cycle. However, one must remember that modularisation, like product development in general, is about making trade-offs and typically all the desired properties for different stages of the life cycle cannot be assigned to a single asset or even a family of assets.

Our hypothesis for future research is that the recognized VCMs support understanding, communicating and estimating potential business impacts and rationale of modularisation. The VCMs derived from the literature review present guiding principles whose importance needs to be defined by the suppliers of engineering assets to find a reasonable structure for asset development. Similar work has been done by Fixson (2006) in product architecture costing, but the engineering dimensions in that approach only highlighted function-component allocation schemes and interface characteristics; from a modularisation point of view, partitioning logic, modules, architecture and configuration rules and constraints should be considered to get a more comprehensive perspective. Our goal is to connect the VCMs to these key elements of a module system (partitioning logic, set of modules, interfaces, architecture and configuration rules and constraints) and develop more systematic tools for estimating the impacts of modularisation, such as comparing different module concepts. For future research topics, VCMs need to be connected to quantitative values to increase their usefulness for management.

References

Amadi-Echendu, J., Willett, R., Brown, K., Hope, T., Lee, J., Mathew, J., Vyas, N., Yang, B.: What is engineering asset management? In: Joe, A.-E., Brown, K., Willett, R., Mathew, J. (eds.) Definitions, Concepts and Scope of Engineering Asset Management, pp. 3–16. Springer, London (2010)

Andreasen, M.M.: 45 Years with design methodology. J. Eng. Des. **22**(5), 293–332 (2011)

Baldwin, C.Y., Clark, K.B.: Design Rules: The Power of Modularity. The MIT Press, Cambridge (2000)

Erixon, G.: Modular Function Deployment - A Method for Product Modularisation. The Royal Institute of Technology, Stockholm (1998)

Fixson, S.K.: A roadmap for product architecture costing. In: Simpson, T.W., Siddique, Z., Jiao, J.R. (eds.) Product Platform and Product Family Design, pp. 305–334. Springer, New York (2006)

Haider, A.: Information systems based engineering asset management evaluation: operational interpretations. University of South Australia (2007)

Harlou, U.: Developing product families based on architectures - contribution to a theory of product families. Technical University of Denmark (2006)

Haug, A., Hvam, L., Mortensen, N.H.: Definition and evaluation of product configurator development strategies. Comput. Ind. **63**(5), 471–481 (2012)

Hopp, W.J., Xu, X.: Product line selection and pricing with modularity in design. Manuf. Serv. Oper. Manage. **7**(3), 172–187 (2005)

Juuti, T.: Design Management of Products with Variability and Commonality. Tampere University of Technology, Tampere (2008)

Lau Antonio, K.W., Yam, R.C.M., Tang, E.: The impacts of product modularity on competitive capabilities and performance: an empirical study. Int. J. Prod. Econ. **105**(1), 1–20 (2007)

Lee, H.L., Tang, C.S.: Modelling the costs and benefits of delayed product differentiation. Manage. Sci. **43**(1), 40–53 (1997)

Lehtonen, T.: Designing Modular Product Architecture in the New Product Development. Tampere University of Technology, Tampere (2007)

Pakkanen, J., Juuti, T., Lehtonen, T.: Brownfield process: a method for modular product family development aiming for product configuration. Des. Stud. **45**, 210–241 (2016)

Pulkkinen, A.: Product Configuration in Projecting Company: The Meeting of Configurable Product Families and Sales-Delivery Process. Tampere University of Technology, Tampere (2007)

Sanchez, R.: modular architectures in the marketing process on JSTOR. J. Mark. **63**, 92–111 (1999)

Soininen, T.: An Approach to Knowledge Representation and Reasoning for Product Configuration Tasks. Helsinki University of Technology, Espoo (2000)

Tiihonen, J., Lehtonen, T., Soininen, T., Pulkkinen, A., Sulonen, R., Riitahuhta, A.: Modelling configurable product families. In: Proceedings of the 12th International Conference on Engineering Design, ICED99, vol. 2. Munich, Germany (1999)

Umeda, Y., Nonomura, A., Tomiyama, T.: Study on life-cycle design for the post mass production paradigm. AI EDAM **14**(2), 149–161 (2000)

Method for Evaluating the Value of Technology in the Manufacturing Industry

Janne Mämmelä[1(✉)], Tero Juuti[1], Pasi Julkunen[2], Pauli Lemmetty[2], and Jarkko Pakkanen[1]

[1] Laboratory of Mechanical Engineering and Industrial Systems, Tampere University of Technology, Tampere, Finland
{janne.mammela, tero.juuti, jarkko.pakkanen}@tut.fi

[2] Rock Drills & Technologies, Sandvik Mining and Construction Oy, Tampere, Finland
{pasi.julkunen, pauli.lemmetty}@sandvik.com

Abstract. The aim of this research is to develop and test a new method for evaluating the value and costs of technology in the manufacturing industry based on Design Science. The overall purpose is to enable the evaluation of product lifetime value creation already in the concept phase. This would improve the manufacturer's capability to design for the total cost of ownership and thus, improve planning in the engineering asset management context. A constructive research strategy is used, and the Design Research Method with an industry case study is applied in developing and evaluating the method. This paper contributes to the literature by presenting a new method for evaluating technology based on Design Science. Product characteristics and properties are modeled, and the technology characteristics are evaluated against them. The proposed method opens the product, which is traditionally seen as a black box in techniques for valuing technologies, to the technology evaluation process. The target application for the proposed method is the existing product and business environment where sufficient design knowledge is available. This method supports decision making in technology-related questions at the managerial level. This paper recognizes and uses value creation mechanisms to evaluate the value and costs of technology.

1 Introduction

Selecting a technology in the manufacturing industry can be complex because technology does not have intrinsic value. The value of a technology is realized only after it is commercialized. Developing a new technology is an uncertain process, and choices can have far-reaching effects. The overall purpose is to enable the evaluation of product lifetime value creation already in the concept phase. This would improve the manufacturer's capability to design for the total cost of ownership and thus, improve planning in the engineering asset management context (Amadi-Echendu et al. 2010). The focus of this paper is to develop and test a method that supports early phase decision making for technologies by evaluating the monetary effects of technology in the context of the manufacturing industry. The method is directed to take into account

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 855–863, 2020.
https://doi.org/10.1007/978-3-030-48021-9_94

the value of the technology in the whole product life cycle. Motivation for this research came from industry where new technologies are seen as an important source of competitiveness. The target application for the proposed method is the existing product and business environment where sufficient design knowledge is available. Results of the valuation are given in monetary terms.

Traditionally, methods for valuing technology originate in financial management and are based on capital budgeting techniques. Understanding and estimating new technologies is challenging with pure monetary-based analysis and backward-looking financial reviews. Technology affects the product, and therefore, understanding the value creation mechanisms through the product is seen as beneficial. Financial estimations of traditional methods for valuing technology are mostly based on scarce forecasts where the link between the real effect of the technology and the product is invisible. Therefore, this paper contributes to current knowledge by presenting a method based on Design Science where the focus is shifted from financial data to the product.

We developed this method for technology evaluation purposes and are interested in how to gather and organize the knowledge needed in the context of manufacturing industry. Our previous study (Mämmelä et al. 2018) showed the information needed in the technology valuation context, and in this paper, we construct and develop a method based on the information. Therefore, this paper includes the following research question (RQ): How to create and structure the information to evaluate the value and costs of technology in the manufacturing industry? A constructive research strategy is used in this paper. The method is developed and tested using Design Research Method (Blessing and Chakrabarti 2009) with case study research (Yin 2014) in a real industry environment. The method is developed using a type 3 research project (Blessing and Chakrabarti 2009, p. 18), which includes four main phases: research clarification, descriptive study 1, prescriptive study and descriptive study 2. The first two phases are based on a literature review, and the last two are focused on developing and testing the method which is the main focus of this paper.

Because the value of technology is situational, company-specific knowledge is the center of the analysis. The proposed method is a five-step method, and the information is gathered mainly in workshops with company and technology experts. During the process, the main targets and limitations for technology exploitation are defined. The product characteristics and properties are modeled, and the technology characteristics are evaluated against them. The technology business impact is evaluated based on modeled value creation mechanisms and available financial data. In this paper, the value is defined as follows (Mämmelä et al. 2018): "the term "value" refers to both monetary and intangible values. Valuable properties are determined by the business owner and can be converted to monetary estimations based on the best available knowledge. Costs are used to represent monetary costs using the same logic as value."

2 Literature Review and Theoretical Foundation

This research is mainly based on ideas of Design Science. According to Hubka and Eder (1996, p. 73), "The term Design Science is to be understood as a system of logically related knowledge, which should contain and organize the complete knowledge about and for design." The Theory of Technical Systems describes the purpose and nature of technical systems, and this theory is a major part of Design Science (Hubka and Eder 1988). The theory of dispositions is used to catch and foresee the effects of technology exploitation decisions (Olesen 1992). Olesen (1992) defined a disposition as "that part of a decision taken within one functional area which affects the type, content, efficiency or progress of activities within other functional areas" (p. 53). Product and business modeling is done and structured according to Property-Driven Development (PDD) approach, which originates in explaining the design phenomenon and improving the design (Weber and Deubel 2003). PDD is described in more detail in Subsect. 3.3.

Technology valuation can be done for many purposes and from many perspectives. The three basic valuation approaches are the cost, market and income approaches (Parr and Smith 2005). The cost approach is based on the idea that the value of technology can be compared to the replacement cost of the technology. The market approach seeks the value of the technology based on others' consensus. The income approach looks at the future earning power of the technology. All other valuation methods are based on these approaches. The income approach is the most recommended valuation method for technology valuation, and this paper develops a method based on the income approach (Jang and Lee 2013; Park and Park 2004). One general challenge in technology valuation is that the value is framed in the eye of the beholder (Boer 1999). As the real benefits of the technology are realized only after commercialization, only a few studies have combined technology and business (Park and Park 2004; Schuh et al. 2012). There are also approaches based on wider perspective of technologies and innovation, for example Fox (2013), but the general challenge of understanding the real effect of the technology remains.

The knowledge needed to evaluate the value and costs of technology in manufacturing starts from defining the technical system intention and the business intention. This determination is done in the first phase of the proposed method. To understand the wanted properties of the product, it is essential to know the product life cycle phases, which are the target of phase 2. The product structure defines the characteristics of the product. The product structure is modeled in phase 3 using Design Reasoning Pattern tool. The main idea is to find similarities between product and technology characteristics and understand the value creation mechanisms between them, phase 4. Combining the financial data in recognized technology effects, it is possible to evaluate the monetary benefit of the technology; this evaluation is done in phase 5 (Mämmelä et al. 2018).

3 Method for Evaluating the Technology

The proposed method is targeted to evaluating the value and costs of the technology in the existing business and product environment. Therefore, we can assume that the product and the business are accepted by the market at a certain level, and information related to the technology exploitation environment is available by company. The results of the evaluation support decision making in technology-related questions. In this paper, the method is understood as described by Newell (1983) by using four statements: 1. A specific way to proceed, 2. A rational way to proceed, 3. involves subgoals and subplans and 4. The occurrence is observable. The proposed method has five main phases, discussed in Subsects. 3.1, 3.2, 3.3, 3.4, 3.5. The method is created to gather the information from the targets for exploitation possibilities. The first two phases focus on target setting, phases 3 and 4 on modeling and recognizing the value creation mechanisms, and the final phase is for monetary estimation of the technology effects.

3.1 Preliminary Targets and Limitations of the Evaluation

To understand the scope of plans for exploiting the technology, the first phase defines the preliminary targets and limitations of the evaluation. Questions to answer include the following:

- Is the main purpose to improve the performance of the product or reduce the operating expenses; that is, what is the business intention?
- Which products are the target of analysis, and what is the technical system intention?
- What is the technology, and what are the characteristics of the technology?
- Which organizations are being analyzed?
- What are the development areas, and what information is static?

Information is gathered in workshops with management in relevant areas. After phase 1, tentative information about the technology and the scope of the evaluation is known. This information supports execution of the next phase and participant selection.

3.2 Targets from Business Environment

The second phase of the method is to set targets for exploiting the technology based on the business environment. The technology influences the technical system, but the effects are evaluated in the business system. Relations between the business environment and the technical system should be set as the center of analysis. Juuti *et al.* (2007) presented Company Strategic Landscape (CSL) framework which defines the elements related to the product development operations and the production of the company. The CSL framework is used to collect the product development and technology targets from the business environment.

The CSL framework includes five main areas related to a company's business environment: strategy, process, product, value chain and organization. To get comprehensive knowledge, the CSL framework is reviewed in a workshop with the management of all relevant departments defined in phase 1. The leader of the technology evaluation process should act as a facilitator in the workshop. The main outcome of this phase is a list of desired properties; see the explanation in Subsect. 3.3, for the technical system from the life cycle phases. The recognized properties act as the target values for the value creation mechanisms in phase 3.

3.3 Modeling Value Creation Mechanisms

The originality of this paper is to show and explain the value creation mechanisms of technology in the context of the manufacturing industry. Therefore, this phase is one of the most critical steps of the proposed method. The value and costs of technology are evaluated according to recognized value creation mechanisms. Modeling is executed based on the same principles that Property-Driven Development presents (Weber and Deubel 2003). The concept is based on the distinction between product characteristics and properties. Characteristics are design parameters over which the designer has direct influence. Properties are understood as a behavioral aspect, such as the power or the price of a technical system. Properties are the results of characteristics and can be designed.

A Design Reasoning Pattern (DRP) is focused on modeling relations between product structuring and the value chain defined in the CSL framework. A DRP is a chart where the value creation mechanisms between the product characteristics and the business goals are recognized and presented. The CSL framework acts as the source of input in this phase, and product structuring is selected as the starting point of the DRP. Value capture and cost properties are set as the goals for the DRP. The DRP also needs input from the company's product development department. The most experienced designers should be involved in this phase. Modeling uses pull control principles to value capture and cost properties toward product structuring. In the product structuring elements are the parts and characteristics where modeling ends. The technology's characteristics should be observed and used to guide modeling the DRP for efficient results.

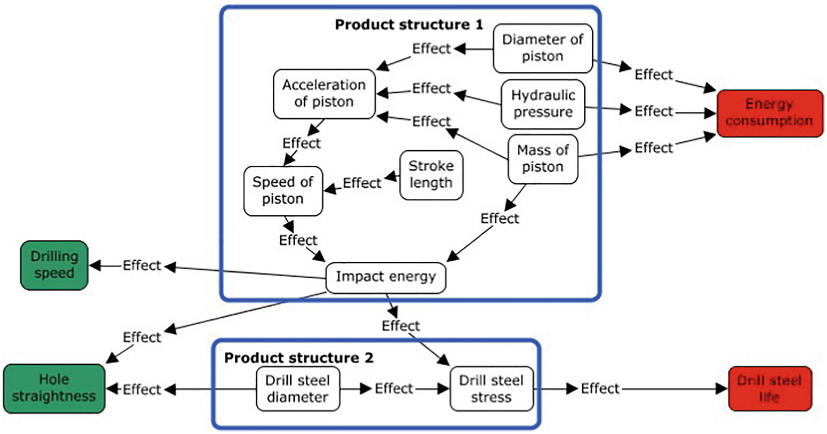


Fig. 1. Example of a simplified DRP chart

Figure 1 represents an example of a simplified DRP chart from the case study. On the left (green boxes) are value creation and the right cost (red boxes) properties recognized in phase 2. The big blue boxes in the middle are product structures. Inside the product structure is the characteristics of the technical system. Arrows describe the relations between the characteristics and properties. Relations between product characteristics and properties can be complex, and all characteristics are not equal. Generally, a few main characteristics have to be understood and controlled for an efficient design solution. There are also links between product structures.

3.4 Evaluation of Exploitation of the Technology

Evaluation of the technical potential of the technology in a defined scope and business environment is done in phase 4. Basic information is structured in previous phases, and we have recognized value creation mechanisms in the DRP. Comparing the characteristics of the technology and product from the DRP reveals the potential of the technology. The effects of the detected similarity can be evaluated based on the DRP chart. For example, the case presented in this paper was focused on metal coating with additive manufacturing. The characteristics of the additive manufacturing (AM) coating technology are related to the surfaces of the parts: the hardness, accuracy or surface quality. Therefore, we looked at this characteristic in phase 1 to model the correct product characteristics in the DRP. Based on our approach, it is not essential to know exactly the technical solution. Instead, we focus on the relevant characteristics that support understanding the real potential of the future possibilities of the technology. The evaluation criteria are the value creation and cost properties defined in phase 2. The evaluation also supports understanding the changes needed in the product and seeing the scope of the new design task.

Exploitation of the technology is evaluated in two steps. In the first step, an experienced designer who has the best knowledge of the product and its dispositions evaluates the technology. The second step is to evaluate the potential of the technology with a technology expert if one is available. Combining the views of these two groups gives a comprehensive understanding of the technical potential of the technology in the defined scope. The recognized potential of the technology is evaluated in monetary terms in the next phase.

3.5 Communicating the Value of the Technology

The fifth and final step is to estimate the business effects of exploiting the technology. The focus is to evaluate the monetary effects of recognized value creation mechanisms from the previous phase. The business impact is analyzed in a workshop with the same group for the CSL framework. The valuation of the technology should be done in a holistic view through the company. The real picture of the effects of the technology will not be shown if only one area is evaluated. Therefore, representatives from all relevant departments are requested. The proposed method allows using a wide perspective in the valuation of technology. At the same time, it is possible to evaluate changes in the company's business and the customers' or suppliers' businesses. Intangible values, such as safety, can be evaluated through the proposed method if the disposition between technology and safety is recognized during the process.

Evaluation is done in the order of the phases of the product life cycle from design to disposal. The product life cycle phases can be found in the CSL framework. Estimations of monetary effects are allocated to specific life cycle phases. A more detailed analysis can be done, for example, using partition of time, quality and amount of sell according to company policies. In addition, different technical solutions can be compared based on the same principles. Estimates of monetary sums are done in rough magnitude of order, EUR 1 000, EUR 10 000 and so on. The idea is to find the maximum potential yearly effect related to the current situation and business. Evaluation of the business impact also needs information about production numbers and other related financial data.

4 Case: Metal Coating with AM in Rock Drills

In this research, the method was tested in the real industry environment. The case study was performed in a global original equipment manufacturer (OEM) in the mining business. The technology to be explored was metal AM printing and especially coating techniques. Steps of the proposed method are used to show the case and the results of specific phases.

Phase 1: The main target was to improve the performance of the product instead of cost savings. Therefore, the focus was the properties defined by the user of the technical system. The technology was the AM coating, and we examined the characteristics related to the surfaces of the parts.

Phase 2: The CSL framework was modeled in a workshop according to the method description. Based on selections done in the previous phase, the main focus was to recognize the properties defined by the user of the rock drill: the drilling speed, hole straightness, energy consumption and drill steel life. See Fig. 1.

Phase 3: The idea of the DRP is to link the characteristics of the product to the properties defined in phase 2. In this case, the DRP was extensive because of the selected technology. Product characteristics have to be modeled until the individual parts and surfaces appear to understand the effects of the surfaces of the product. This approach and the selected technology limited the evaluation of significant structural changes in the product. The analysis focused on the current function principle and structure. Designers involved in creating the DRP were interested and active participants. During the process, the common understanding of the product and design was improved.

Phase 4: In the evaluation of the exploitation of the technology phase, the designers and technology experts recognized two potential areas for the coating. More detailed research revealed that the recognized value could not be captured only by using AM coating techniques. Other parts and the structure of the product prevent the practical benefit of the planned technology. The main outcome of the case was that the explored technology does not have common characteristics with the product. Positive value creation mechanisms were not recognized during the evaluation process. To benefit from the AM technology, the product should have new structures or function principles that can be supported by AM characteristics.

Phase 5: Monetary estimation of the effects of the surface coating were done although the technical potential of the technology was unfavorable. The planned design solution was evaluated with managers of all related areas of the product life cycle. The new solution was minor and caused changes only in the production process in the product life cycle. In the case of the AM coating technology, one step has to be added in the manufacturing process, and it will increase the cost of the product.

5 Discussion and Conclusion

In this paper, a method for evaluating a new technology was developed and tested in the industry environment. The novelty of this research is to explain and show value creation mechanisms for technology. The research question was focused on gathering and structuring the information needed to evaluate the technology. The proposed method is a five-step method. The information gathering starts with the targets and continues toward technology exploitation possibilities and monetary estimation of effects. The main source of information is company personnel, and information is gathered in workshops. Selected theories and the research method provide relevant answers to and explanations for the phenomenon of the value creation of the technology. As the case study shows, the method is dependent on available design knowledge and the structure of the product. A totally new product concept can be evaluated if knowledge needed about the function principle is known. This method is generally related to the product, technology or business environment. The future research agenda is to perform more case study research related to verifying and developing the method for additional applications for practical purposes.

References

Amadi-Echendu, J.E., Willett, R., Brown, K., Hope, T., Lee, J., Mathew, J., Vyas, N., Yang, B.-S.: What is engineering asset management? In: Amadi-Echendu, J.E., Brown, K., Willett, R., Mathew, J. (eds.) Definitions, Concepts and Scope of Engineering Asset Management, vol. 1, pp. 3–16. Springer, London (2010)

Blessing, L., Chakrabarti, A.: DRM, A Design Research Methodology. Springer, London (2009)

Boer, F.P.: The Valuation of Technology: Business and Financial Issues in R&D. Wiley, New York (1999)

Fox, S.: The innovation big picture: including effectiveness dependencies, efficiency dependencies, and potential negative effects within the framing of new technologies. Technol. Soc. **35**, 308–314 (2013)

Hubka, V., Eder, W.E.: Theory of Technical Systems: A Total Concept Theory for Engineering Design. Springer, Berlin (1988)

Hubka, V., Eder, W.E.: Design Science: Introduction to the Needs, Scope and Organization of Engineering Design Knowledge. Springer, New York (1996)

Jang, W.-J., Lee, C.: Technology Valuation Model for the Defense R&D With Income Approach. Int. J. Innov. Technol. Manage. **10**(04), 1350017 (2013)

Juuti, T., Lehtonen, T., Riitahuhta, A.: Managing re-use for agile new product development. In: ICED07: 16th International Conference of Engineering Design, 11 August (2007)

Mämmelä, J., Juuti, T., Korhonen, T., Julkunen, P., Lehtonen, T., Pakkanen, J., Vanhatalo, M.: Evaluating the value and costs of technology in the manufacturing industry. NordDesign 2018, 14–17 August 2018, Linköping (2018)

Newell, A.: The heuristic of George Polya and its relation to artificial intelligence. In: Methods of Heuristics, pp. 195–243 (1983)

Olesen, J.: Concurrent Development in Manufacturing. PhD, Technical University of Denmark (1992)

Park, Y., Park, G.: A new method for technology valuation in monetary value: procedure and application. Technovation **24**(5), 387–394 (2004)

Parr, R.L., Smith, G.V.: Intellectual Property: Valuation, Exploitation, and Infringement Damages. Wiley, New York (2005)

Schuh, G., Schubert, J., Wellensiek, M.: Model for the valuation of a technology established in a manufacturing system. Procedia CIRP **3**(1), 602–607 (2012)

Weber, C., Deubel, T.: New theory-based concepts for PDM and PLM. In: ICED 03 (2003)

Yin, R.K.: Case Study Research: Design and Methods, 5th edn. Sage, Los Angeles (2014)

Co-value Creation Within the Business Model for Smart Grids: Case of Russian Autonomous Energy Complex

Irina Volkova(✉), Yegor Burda, and Elizaveta Gavrikova

National Research University Higher School of Economics, Moscow, Russia
{iovolkova, yburda, egavrikova}@hse.ru

Abstract. The trend on electricity grids digitalization is gradually leading to the shift of business value towards more sustainable and efficient electricity services. Sustainability and efficiency are challenged by the increasing demand for electricity which is followed by a dramatic transformation of energy systems. While smart grids seem to be crucial in this process, there is a discrepancy in understanding the costs and benefits for the multiple actors involved. In addition, there are benefits of smart grids that cannot be measured directly in terms of money, such as higher energy system reliability or commitment to carbon reduction. Despite the rise of interest to the managerial aspects of smart grids implementation and development, many aspects remain out of the scope. This paper contributes to the research of smart grids by providing a conceptualized business model that would allow for value co-creation, delivery and capture. A Russian energy sector perspective is primarily considered throughout the paper and the results are supported by evidence from interviews with of industrial experts.

1 Introduction

Liberalization of the power market, characterized by unbundling power generation and distribution activities (Friedrichsen 2015) created a fruitful basis for the development of a competitive market for electric energy. At the same time, recent developments in smart grid technologies widened these opportunities by decreasing the barriers for new entrants and through widening the spectrum of activities within the power sector (Shomali and Pinkse 2016). For the last few years we are able to observe a rapid growth of various projects on smart grids all over the world (Colak et al. 2016) and co-occurring increase in academic research on the topic (Cardenas et al. 2014). Smart grids can be perceived as an instrument for achieving a set of energy goals: political, societal, environmental, economic, etc. At the same time, there is still no common understanding on how smart grids have to be implemented, while roles and responsibilities have to be clearly defined and assigned to major actors within the paradigm.

Most of the papers are devoted either to the technical details of smart grid projects – like communication systems (Mahmood et al. 2015; Reddy et al. 2014), cloud computing and data protection (Bera et al. 2015; Eissa 2015; Xia and Wang 2012) and questions of optimization (Vardakas et al. 2016; Zhou et al. 2014) or to smart grids'

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 864–872, 2020.
https://doi.org/10.1007/978-3-030-48021-9_95

general description (Farhangi 2010) and various case studies (Brunekreeft et al. 2015; Cambini et al. 2016; Losch and Schneider 2016) as well as success stories (Crispim et al. 2014). At the same time, there are very few publications focused on the managerial aspects of smart grid projects (Cowan 2013) and most of these studies are descriptive ones.

Nevertheless, there are studies devoted to a very important issue – the essence of managerial implication of smart grids – their business models (Niesten and Alkemade 2016). However, the question of value delivery and capture remains out of the scope of scholars, whereas there is a concrete evidence that lot of firms involved in the system failed to generate revenues (Shomali and Pinkse 2016).

This calls for the identification of corresponding business models as well as incentives and values that are important for the implementation of smart grids in energy systems. A growing body of research is devoted to the dynamics and aspects of energy systems transformation, while most of the literature is focused on exogenous pressures and technological advancements as stimuli for change. The aim of the paper is to provide a schematic business model for smart grid, basing on the case of the Autonomous Energy Complex being implemented in Russia (smart grid analogue), with the specific attention being paid to the mechanisms of value co-creation, delivery and its capture for different actors involved.

2 Concept of the Autonomous Energy Complex

In accordance with the "Concept of development of the electric power market on the basis of new technologies" (EnergyNet Working Group 2017), which was created for the execution of the "Internet of Energy" paradigm in Russia, the necessity for new opportunities for the development of Russian power markets is determined by two major factors. First, the development of new power and information technologies (industry cannot neglect the innovative path). Second, the stagnation of the Russian power industry (ageing assets and scarceness of resources), which determines the necessity to increase its efficiency (including the implementation of various "smart" technologies and methods).

To fully understand the basics of AEC, it must be noted that the Concept considers the electric power market as an "*organized system for the relationships between actors for purchasing, selling and exchanging of energy as well as any other goods and services, relevant for the fulfillment of demand for electric energy*". This definition determines the market much wider and "*foresees the competitiveness for various technologies of power supply together with the ability for implementing an open platform for new market entrants*". That is, the market is considered as an innovative ecosystem[1] with its own actors and their interrelationships. The Concept is aimed at creating an opportunity for a new actor to participate in the market (*which would be universal both from technical function within the system and its market role. It will*

[1] This paper operates the term "innovative ecosystem" proposed in Gawer, A., 2009. Platforms, Markets and Innovation. Edward Elgar Publishing Limited. 396 p.

incorporate the ability to produce, consume, store and distribute energy, manage its production schedule and load, provide other services within the system, etc.). This actor is called an "Autonomous Energy Complex".

As we can see from the provided description, AEC is considered as a micro-smart grid, which can both function autonomously from the United Power Grid of Russia or collaborate with it via a Smart Connection, which determines the set of standard technical and legal conditions. The latter opportunity is crucial for AEC, as the benefits of smart grids are fully utilized only in the case of coordination of decentralized users and the global network (Friedrichsen 2015). The implementation of AEC will create a major change in the list of stakeholders of the power system (EnergyNet Working Group 2017) as shown in Table 1.

Table 1. Comparison of basic actors' list within traditional power system and AEC

Traditional power system	AEC
Generation – Distribution – Consumers	Members
System Operator	Independent AEC Operator
Regulators (Ministry of Energy, Antitrust Committee, etc.)	–
Financial Payments Operator	Service Providers: Trading System; Infrastructure Developers; Digital Data Operators; Energy-Management Service Providers; …

The list includes traditional actors of the Russian United Power Grid

AEC diminishes the distribution of roles within the system via establishing a concept of a universal member of the system, which can combine all 3 roles (generation, distribution and consumption). Thus, AEC is aimed at large commercial consumers, i.e. industrial plants, that hold generation equipment (in order to assure an emergency back-up) and own electricity grids (in order to distribute electricity for internal use). Such consumers may provide local households with power either by selling them internal generation surplus or through providing grid connection and selling surplus power energy from the United Power Grid of Russia. AEC is supposed to allow such consumers to sell power energy surplus not only to local households, but also to other consumers as well as liberate them from payments to the System Operator and Financial Payments Operator (as they have their own competence to control the grid and settle accounts).

At the same time, the System Operator remains as a part of the energy system. Being a government institution, it balances the discrimination concerns and coordination needs, while AEC's independence implies efficiency gains from coordination of investments. The financial payments operator is replaced by an electronic trading system (created upon the specific requirements of a particular AEC). The data generated by the smart devices within the system (meters, storage, etc.) is collected by a digital data operator, which aggregates and processes it for the purpose of the whole network via various Big-data solutions (for forecast purposes). The absence of the state support for creating and supporting infrastructure is compensated by developers who, being independent actors, are able to provide efficient solutions on a competitive basis (initial costs may be depreciated during the life cycle of the assets). The AEC model also creates numerous opportunities for establishing complementary services like fiscal metering or smart data processing.

The actions within the United Power Grid are controlled by the Ministry of Energy and the Antitrust Committee, while AEC functions on purely market conditions. AEC is considered being a united power facility and a single subject of legal relations which is fully responsible for any actions executed via the Smart Connection. Membership in the AEC is purely voluntary and participation conditions are fixed in contracts and constitutional documents (in case if several legal entities merge and make up a single one in any form of legal ownership). Any business activities and contracts (concerning energy supply and related services) executed within the AEC aren't subject to additional regulation apart from that declared by the norms of Civil Law (general requirements regarding electrotechnical and ecological safety). A member that breaks the contract and leaves AEC has to connect to the United Power Grid over again, because AEC has the right to stop energy delivery. AEC is able to participate both in the wholesale and retail electricity markets without any additional obligations concerning internal technical characteristics (devices and equipment), while it has to conform to the technical and market rules of interaction with the United Power Grid. At the same time, there are limitations regarding the range of objects that could be included in the AEC, in example, regulated electrical grids or power generation facilities that pertain to the United Power Grid and couldn't be excluded due to their societal importance. The conditions of electricity and electric power supply, reliability and quality parameters as well as terms of payment have to be determined by contractual relations and technical settings of the Smart Connection (through the "smart contract" technique).

Thus, the AEC has a reasonable amount of freedom when it comes to entering the electricity market and establishing relationships with its members and consumers, however, AEC is also fully liable for contractual agreements (including technical aspects). Internal balancing of peak demands (within the AEC) with a regulated Smart Connection and clear lines of accountability and responsibility of all the parties has a significant potential to improve the quality of the market. It also leads to increasing reliability of energy supply and electricity quality both for individual participants and the whole energy system.

Evidently, the AEC model favors the promotion of both business (open and competitive market) and societal (safe and reliable energy supply) goals. This fact unleashes the full potential of smart grid, when members act not to fulfill the established requirements of the government – which is not effective in the end (Boer et al. 2014), but rather act basing on the criteria of maximizing value.

Given that the list of stakeholders is changing, a new business model that would be able to outline major value streams (co-creation, delivery and capture) is required. The next section of the paper will provide a conceptualized business model for an autonomous energy complex, which will take into account its major differences from the existing power systems.

3 AEC's Business Model

The methodology of the research is presented by a holistic case study based on a review of academic literature and a set of interviews with experts from the Russian power industry. The Agency for Strategic Initiatives (ASI) and Ministry of Energy initiated a special industrial strategic session on the development of smart grid in Russia that was held in May 18–20, 2017. Field data was gathered during this event. This strategic session involved participants from all the major institutes of the industry: federal executive authorities, private sector, academia, etc. The main goal of the session was to understand the perspectives of the power sector for the nearest 10–20 years. The session included a number of profile tracks of discussion (within working groups): management systems, energy-market, energy-science, start-ups, cooperation, etc. During the session of the Energy Market Working Group AEC was considered as a possible prototype for the energy market in general. The question of the future for AEC and its business model in particular was raised.

There were three sequential sessions during which we were able to interview 28 experts from this working group. The Delphi Method and brainstorming technique were implemented in order to generate ideas, develop creative solutions and come to mutual agreement. During the first session general questions were raised, including the specificity of the Russian power energy sector, its current state and possibility of adapting the European experience. The second and third session were focused of the formulation and validation of key requirements for the proper AEC functioning. The results of the three sessions allowed us to create a conceptualized business model for AEC (Table 2). Smart grids' business models are usually studied via the framework of Teece as in some research papers (Niesten and Alkemade 2016; Shomali and Pinkse 2016). This framework distinguishes 3 components: Value Creation, Value Delivery and Value Capture (Teece 2010) and we applied it to the case of AEC.

AEC creates a number of value streams for all the basic actors of the system. Members get an opportunity to choose their role within the system (pure consumer or mixed type), achieve higher level of availability of electricity (transparency of the system delivers higher level of sustainability) as well as a much more observable system and ability to predict own consumption or/and generation. Moreover, the system is able to provide various additional services, which increase the overall value created within the AEC. AEC operator, in line with its goals, receives higher level of observability and predictability of the whole system, which gives an opportunity to create a more efficient (flat) load curve. Service providers get an ability to fulfill their major goal – they receive a platform where they can provide specific services for all of the actors in the system. All the relationships are governed and processed via specific ICT system/platform, which creates an opportunity for balancing the interests of various actors of the system. Trading and additional services are executed within this system as well.

At the same time, the system requires a number of specific resources and capabilities vital for successful participation. Members should have necessary skills for acting in the market (trading skills, understanding or basic market mechanisms within AEC) and corresponding equipment (smart meters, specific software). Operator of AEC

Table 2. AEC's business model

		Members	AEC operator	Service providers
Value creation	Value proposition	Choice of role; Availability; Observability and predictability; Additional Services	System' observability and predictability	New market (new revenue streams)
	Customer relationship	Customized solution via ICT platform	Trading system (ICT based)	Customized solutions via ICT platform
Value delivery	Resources & capabilities	Skills for acting in the market; Specialized equipment	Trading system; Specific skills	Specific knowledge, know-how
	Value network composition	Market relationships	Trading system	Services
Value capture	Revenue streams	Payments for generation; Savings from controlled consumption	Profit from trades; Payments for providing data	Payments for services
	Cost structure	CAPEX (smart equipment); OPEX (equipment maintenance)	CAPEX (hardware and software for the system); OPEX (system support and maintenance)	CAPEX (equipment and R&D, necessary for providing service); OPEX (support and maintenance)

The table does not include the regulator due to the nature of the AEC concept, which was explained in the paragraph 2 of the paper

should have the trading system itself (it does not matter, whether this activity is performed within the AEC or outsourced) and corresponding skills as well. Service providers must be able to survive the competition (they should have valuable knowledge and know-how). The value network composition is presented by relationships of basic members, trading system and services.

Members get an ability to capture value through payments for generation (in case if they sell their surplus to the system[2]) and/or they can save certain amount of money through the control on their own consumption (for example, by shifting it to off-peak periods). AEC operator is able to receive profit from trades (as a fixed percent-based fee – by analogy with banks, who charges a certain amount for all the transactions held within the system) and payments from those members of the system, who are ready to

[2] Activity, performed by so-called "active consumers"/"prosumers".

buy aggregated[3] data accumulated within the system (for example, for future proposition of various optimization mechanisms and/or software). The main revenue stream for service providers is the payment for these services.

At the same time, such flexible and market-like system requires certain investments (some of which are upfront and some even stranded). Members have to invest in equipment and incur expenses for its future maintenance. AEC operator should invest (upfront) in the equipment and software necessary for establishing the trading system and maintain its stability and reliability in the future. Cost structure for service providers is determined purely by the service they provide and may vary from high CAPEX/low OPEX to the vice versa situation, basing on the existing economies of scale of a particular service.

In terms of risks and value for the whole energy system the proposed AEC business model allows for additional value creation through the introduction of market mechanisms, while risks are mitigated due to the implementation of smart grid technologies and the employment of “smart contracts”.

4 Conclusions

To the authors’ best knowledge, this is the first time such a concept is being discussed regarding the Russian energy sector. As a conclusion, we would like to outline some limitations and suggest directions for future research. As far as our findings are explorative in nature and mainly describe the situation in the energy sector of Russia, their transferability to other countries’ situations is limited. Therefore, future research has to be devoted to conducting additional case studies and cross-country comparisons – it is worth investigating to how similar concepts are executed in various countries in terms of the main components of the business model. The most interesting would be to determine whether there is any difference in value co-creation, delivery and capture. And if there is, then what it is due to – specificity of the power industry of a particular country or other factors.

Changes in generation, transmission and distribution as well as consumption of power energy are driven by the global transition towards more sustainable and efficient energy systems. However, additional costs related to developing smart infrastructures may weaken incentives for the implementation of micro grids like AEC. An important issue concerning the adoption of new business models and practices is related to the resistance from incumbent market actors, especially network operators, financial intermediaries, generation companies and T&D utilities, as their sales depend heavily on the existing integrated energy system and large fuel producers. Network operators will lose some of their authority, while financial intermediaries will be excluded from the whole process as well as retailers.

In addition, the influence of institutional changes (i.e. higher demand for EV’s, renewables penetration or high-capacity energy storage systems development) has to be

[3] No specific and precise data should be sold, because it may arise the question of securing (and even freedom).

taken into account. Thus, national level regulatory frameworks and energy policies concerning public energy infrastructures are crucial for the process of modernization of the sector and establishing incentives for all stakeholders. At the same time, new market driven business goals have to be balanced with societal goals, so that to guarantee affordable and secure energy supply. The possibility of such balancing was shown on the example of AEC which provides incentives for businesses (there is an open and competitive market) as well as for end-users (by ensuring a safe and reliable energy supply) and contributes to mitigating risks and increasing value of the whole energy system.

The proposed conceptualized business model of smart grid by the example of the Autonomous Energy Complex in Russian power industry supports the thesis about the necessity of establishing clear methods for value co-creation, delivery and capture in order to provide a concrete evidence for participants that they will benefit from the membership. These methods are different for all the actors and should be balanced in order to achieve certain technical and economical synergies and maintain stability of the whole system. Additional analysis is required in order to establish those synergies. Business model provided in Table 2 is an attempt of such balancing, as all the components are carefully tailored in order not to damage anyone of the participants (neither at the initial stage, nor in the future). The research and its results may be used for developing the roadmap of innovative development of Russian power industry.

References

Bera, S., Misra, S., Rodrigues, J.: Cloud computing applications for smart grid: a survey. IEEE Trans. Parallel Distrib. Syst. **26**(5), 1477–1494 (2015)

Boer, D., Salles-Filho, S., Bin, A.: R&D and innovation management in the Brazilian electricity sector: the regulatory constraint. Int. J. Technol. Manage. Innov. **9**(1), 44–56 (2014)

Brunekreeft, G., Luhmann, T., Menz, T., Muller, S., Recknagel, P.: Regulatory Pathways for Smart Grid Development in China. Springer, 190 pp. (2015)

Cambini, C., Meletiou, A., Bompard, E., Masera, M.: Market and regulatory factors influencing smart-grid investment in Europe: evidence from pilot projects and implications for reform. Util. Policy **40**, 36–47 (2016)

Cardenas, J., Gemoets, L., Jose, H., Rosas, J., Sarfi, R.: A literature survey on smart grid distribution: an analytical approach. J. Clean. Prod. **65**, 202–216 (2014)

Colak, I., Sagiroglu, S., Fulli, G., Yesilbudak, M., Covrig, C.: A survey on the critical issues in smart grid technologies. Renew. Sustain. Energy Rev. **54**, 396–405 (2016)

Cowan, K.: A new roadmapping technique for creatively managing the emerging smart grid. Creativity Innov. Manag. **22**(1), 67–83 (2013)

Crispim, J., Braz, J., Castro, R., Esteves, J.: Smart grids in the EU with smart regulation: experiences from the UK, Italy and Portugal. Util. Policy **31**, 85–93 (2014)

Eissa, M.: Protection techniques with renewable resources and smart grids – a survey. Renew. Sustain. Energy Rev. **52**, 1645–1667 (2015)

Farhangi, H.: The path of the smart grid. IEEE Power Energ. Mag. **8**(1), 18–28 (2010)

Friedrichsen, N.: Governing smart grids: the case for an independent system operator. Eur. J. Law Econ. **39**, 553–572 (2015)

Losch, A., Schneider, C.: Transforming power/knowledge apparatuses: the smart grid in the German energy transition. Innov.: Eur. J. Soc. Sci. Res. **29**(3), 262–284 (2016)
Mahmood, A., Javaid, N., Razzaq, S.: A review of wireless communications for smart grid. Renew. Sustain. Energy Rev. **41**, 248–260 (2015)
Niesten, E., Alkemade, F.: How is value created and captured in smart grids? A review of the literature and an analysis of pilot projects. Renew. Sustain. Energy Rev. **53**, 629–638 (2016)
Reddy, K., Kumar, M., Mallick, T., Sharon, H., Lokeswaran, S.: A review of integration, control, communication and metering (ICCM) of renewable energy based smart grid. Renew. Sustain. Energy Rev. **38**, 180–192 (2014)
Shomali, A., Pinkse, J.: The consequences of smart grids for the business model of electricity firms. J. Clean. Prod. **112**, 3830–3841 (2016)
Teece, D.: Business models, business strategy and innovation. Long Range Plan. **43**, 172–194 (2010)
EnergyNet Working Group: The concept of development of the electric power market on the basis of new technologies (2017). https://drive.google.com/file/d/0BwqlN9yVQ-hCTmI4QXE0UzMtWWc/view
Vardakas, J., Zorba, N., Verikoukis, C.: Power demand control scenarios for smart grid applications with finite number of appliances. Appl. Energy **162**, 83–98 (2016)
Xia, J., Wang, Y.: Secure key distribution for the smart grid. IEEE Trans. Smart Grid **3**(3), 1437–1443 (2012)
Zhou, K., Yang, S., Chen, Z., Ding, S.: Optimal load distribution model of microgrid in the smart grid environment. Renew. Sustain. Energy Rev. **35**, 304–310 (2014)

The Impact of Digitalization on Product-Service System Development in the Manufacturing Industry

Jyri Hanski, Helena Kortelainen, and Teuvo Uusitalo(✉)

VTT Technical Research Centre of Finland Ltd., Tampere, Finland
{jyri.hanski,helena.kortelainen,teuvo.uusitalo}@vtt.fi

Abstract. Digitalization is a trend that is changing society and that will also have significant effects on the asset management and manufacturing industry. Simultaneously with the digitalization trend, formerly product-centric companies have been increasingly adopting service components in their products and basing their competitive strategies on services. The purpose of this study is to increase the understanding of how manufacturing companies see the effects of digitalization in their business, and how digitalization might affect their product and service provision including asset related services. The paper utilises explorative case study approach. The research data was collected by semi-structured interviews with experts representing different roles in manufacturing ecosystems. The paper builds a classification for digitalization-enabled PSS for manufacturing industry.

1 Introduction

Digitalization is one of the major global trends that is expected to transform the manufacturing industry by enhancing the collection and analysis of information in a crucial way. One of the commonly used definitions of digitalization is using "digital technologies to change a business model and provide new revenue and value-producing opportunities" (Gartner 2016). The digitalization process has been on-going in the manufacturing industry for decades, for instance in forms of condition-based maintenance and remote control. However, the recent developments in technology and the reduced cost of sensors and IT systems enable a more advanced form of digitalization called the Internet of Things (IoT). According to Gubbi et al. (2013) IoT can be defined as interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified frame-work, developing a common operating picture for enabling innovative applications. This requires seamless large-scale sensing, data analytics and information representation using cutting edge ubiquitous sensing and cloud computing.

Simultaneously with the digitalization trend, formerly product-centric companies have been increasingly adopting service components in their products and basing their competitive strategies on services (Lerch and Gotsch 2015; Baines and Lightfoot 2013). The service research has moved from dyadic supplier customer relationships to service ecosystems and value co-creation through service platforms (Lusch and

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 873–880, 2020.
https://doi.org/10.1007/978-3-030-48021-9_96

Nambisan 2015). This development has been commonly known as servitization, or expanding the offering to product-service systems (PSS). PSS can be defined as "a marketable set of products and services capable of jointly fulfilling a user's needs" (Goedkoop et al. 1999).

Recently de Senzi Zancul et al. (2016) have studied the application of IoT technologies in PSS. They argue that one major opportunity for IoT adoption is the continuing trend towards servitization. Digitalization offers possibilities to extend service offering by enabling the service provider to monitor and to gather data from its products during the usage phase. Rymazewska et al. (2017) present case studies and a framework for developing digitalization-enabled PSS. They conclude that IoT-based solutions can play a significant role in the development of the PSS of the future. IoT-based servitization strategy enable companies to create value propositions based on reliable data on product usage and performance.

Companies are at varying levels of development in relation to implementing the digitalization-enabled PSSs. The level of adapting digital solutions is dependent, for instance, on the size, industrial sector, and location of the company. Advanced PSSs can be classified into four levels, based on their capabilities: monitoring, control, optimisation, and autonomy (Porter and Heppelmann 2014). Tukker (2004) divides PSS into product-oriented, use-oriented, and result-oriented services where product-oriented services represent a low level and result-oriented a high level of servitization.

The goal of this study is to increase the understanding of how manufacturing companies see the role of digitalization and PSS in their present business. In this paper, we first introduce the methodology of the study, and then present the details about the studied companies and the results of the case studies. Finally, we present conclusions that can be drawn from the results.

2 Research Design

The methodological approach utilised in this study is qualitative case study research (e.g. Yin 2003). The nature of the study is exploratory and its goal is to generate theory from case study evidence (e.g. Eisenhardt 1989). Selective (or purposeful) sampling was used to choose the case companies for the research (Morse 1991; Coyne 1997). A typical manufacturing value network includes actors such as the lead producer and its suppliers and customers (e.g. Miltenburg 2005). For this study, we focus on lead producers who provide the PSS and their subcontractors and IT infrastructure providers who are the crucial suppliers in the value network. The case firms were selected based on their reputation as advanced actors in digitalization. The selection was made in an expert group that conducted a study on status of industrial internet in Finland (Ailisto et al. 2015).

Representatives from ten companies operating in the Finnish manufacturing industry were interviewed on their views and insights regarding the effects of the digitalization on their industrial branch and their company's competitive position Five of the companies were manufacturing companies providing solutions with emphasis on systems and services. Two companies were IT infrastructure and service providers and two companies were subcontractors for manufacturing industry. One interviewee was from a development agency that supports regional companies. The interviewees were responsible for new digital services or product development in their companies or organisations. Semi-structured interviews were conducted either face to face, by phone, or using IP-based communication software. The interviewees were asked to express their views on the industrial internet, digitalization, and the related product-service systems, both in their own company and in the relevant industrial sector. The interviewees were responsible for new digital services or product development in their companies or organisations. Semi-structured interviews were conducted either face to face, by phone, or using IP-based communication software. The interviewees were asked to express their views on the industrial internet, digitalization, and the related product-service systems, both in their own company and in the relevant industrial sector.

The authors analysed the results as a team, thus increasing the creative potential of the study and the confidence in findings (Eisenhardt 1989). Detailed write-ups were made for each case. The interview data was coded. The cases and actors were analysed individually to identify similarities within them. Additionally, the patterns across actor groups were analysed.

3 How Is the Manufacturing Industry Affected by Digitalization?

Lead producers consider that digitalization enables better integration of companies and business networks, which has great potential for increasing the productivity of the manufacturing industry. Overall, companies in the manufacturing industry are following the development and developing digital applications at varying levels; some companies are integrating sensors into their products and manufacturing plants, whereas others are already exploring the opportunities of big data and other forms of advanced analytics. There are some frontrunners regarding the intelligent products, but also a lot of those who have neither noticed nor reacted to the trend. Large companies are more advanced than SMEs in applying digital technologies. Some bigger manufacturing companies are not yet very advanced in digital services, because their machine fleet is so versatile; newer products are technologically advanced but a large part of their revenue comes from previous product generations. Most companies are still uncertain about what kind of impact digitalization will have on their business.

The dominance of hardware is diminishing as the new business environment emphasises services and software. Frontrunner lead producers aim for new ecosystems and business related to the intelligent products through partnerships. Information produced by other actors in the business ecosystem, and also by third parties, is utilised. This forms the basis for more transparent as-a-service business models, which are expected to be more popular in the near future. Advanced companies consider the digitalization of existing services that could be offered to new customer segments. Many lead producer companies already have a strong tradition in utilising digital applications. Currently, the digitalization affects mostly assets and asset-related services, and the main applications concentrate on how to get better predictions from the data, how to automate analytics, and how to increase the productivity of maintenance activities. More advanced companies develop innovation platforms and application programming interfaces (APIs), organise innovation contests, and discuss with start-ups and SMEs to provide novel content for their platform and third-party services. Lead producers develop and offer autonomous and intelligent machines, fleet-level solutions (optimising and prioritisation), and PSSs. Some advanced companies have a standard portfolio of advanced digital services in place. However, digital applications currently represent only a small part of the whole offering of the companies.

IT infrastructure provider companies see digitalization as a big opportunity that is affecting all the industrial sectors. Intelligent products are a natural growth area for their current competencies, as cloud services and servers are central in solution delivery. ICT solutions will be increasingly cloud based, and electronic commerce is going to change radically. IT infrastructure providers see two perspectives of the digitalization: (1) having lots of data in their own networks offers business opportunities through analysis, and (2) network management and remote control services can be offered to customers. Internally, the intelligent products have a major role; there is a large amount of data in the companies' networks that can be analysed. The goal of the analysis is to improve their processes and prioritise investments better. Digital services have been utilised already for a long time in network management. For instance, the management of customers' networks, remote control, and network development are digital services that are already offered. Companies can sell their digitalization competencies and strive for pilot cases to test potential new services, which also include joint offerings with third-party companies.

Subcontractors that deliver mainly components and individual subsystems to system providers find that digitalization is not yet actualised at subcontractor level. Subcontractors think that lead producers are the main utilisers of intelligent products, and see the applications as an increased number of wireless sensors. The interviewed companies do not have digital applications, apart from sensors in the manufacturing process, in use, but they offer products in which it could be utilised in the form of condition monitoring. Additionally, there are product development projects in place that also concern the supporting IT solutions. The interviewed subcontractors raised up the use of solutions that do not require as much initial investment and that are "good enough" for their use, such as social media, cloud services, and microblogs. These internet-based solutions are in use to spread information and hold discussions with the lead producer's development department.

4 Classification of Digitalisation-Enabled PSS in Manufacturing Industry

Based on the interviews and classifications of Porter and Heppelmann (2014) and Tukker (2004), we classify digitalization-enabled PSSs into four categories, reflecting the complexity and potential benefits of a service delivery:

1. digital services related to products and manufacturing processes,
2. optimising the performance of a fleet of machines, components, and manufacturing equipment,

3. advanced asset management and productivity services, and
4. integrated business ecosystems.

The first category deals with single product-related services and aims to increase the effectiveness of maintenance and other product-related services through operation and business-related data. From the manufacturing perspective, the goal of these services is to improve the reliability and quality of the product or the manufacturing process. In the second category, the focus is on the fleet of products, services, and manufacturing equipment. These solutions aim to improve the reliability, resource efficiency, quality, and performance of PSS delivery and manufacturing equipment, through analysis of a large amount of data and understanding the best practices in the industry.

Advanced asset management and productivity services (third category) consider intelligent, resilient, and self-adapting manufacturing equipment and the PSS. Digital services are integrated into customers' systems, optimising the performance and life-cycle costs of the PSS. For instance, remote monitoring of products, and information on asset location, condition, and use, are necessary on this level. From the manufacturing perspective, the third category contains automated decision-making, based on predictive and condition-based maintenance. The integrated business eco-system level (fourth category) includes all the features on the previous levels on an ecosystem level. It focuses on creating the service experience together with an extended business ecosystem consisting of customers, subcontractors, service providers, consumers, and so on. In this category, the manufacturing ecosystems are transparent and enable the different actors in the network to plan their functions optimally. It includes the choice of a manufacturing method that is optimal for the PSS.

5 Discussion and Conclusions

In this paper, we presented views of different actors in the manufacturing eco-system, lead producers, IT infrastructure companies and subcontractors, on digitalization. Additionally, we presented a framework of different levels of a digitalization-enabled PSS. The study increased the understanding of how Finnish manufacturing companies see the role of digitalization and digital services and how different actors in the manufacturing industry position themselves in the framework. Manufacturing industry companies have been adopting, or have considered adopting, digital services as a part of their offering, or to enhance their manufacturing processes. Based on the interview results, lead producers are the most advanced in providing a digital PSS to their customers.

Solution and IT infrastructure providers, in particular, see that digitalization brings new types of PSS and business models. The importance of service components increases and the nature of service delivery may change, especially in maintenance. These changes require new information management and information systems-related competencies in companies. With the increased data intensity of the PSS, questions about the ownership, availability, and security of data become more important.

In our study, the subcontractors (all SME companies) expected that the lead producers would drive the change. On the other hand, lead producers were anticipating new entrants to the market and changing roles in the supply chain. IT providers expecting changes in the logistics chains and major changes in the industry. This suggests that manufacturing ecosystems are in transformation. We still lack understanding on the transformation towards digitalization enabled PSS. Especially, how to involve SMEs to the change and new value sharing models.

The importance of IT and service components increases when moving towards more advanced PSSs (see Fig. 1). The increased importance of IT is a potential source of disruption for the manufacturing industry. It is uncertain which actors in the business ecosystem will take control of the increased value in IT, and which will be reduced to a lesser role. IT infrastructure providers or platform providers may take over the management of platforms and take a major share of the value of manufacturing. However, it is also possible that their role will be reduced to providing connectivity between the actors in the network. Similarly, solution providers may control the whole ecosystem or may be "just" hardware product providers. Additionally, subcontractors may take a larger role in the supply chain, and may start to compete with solution providers, retain a similar position as currently, or be forced out of the market.

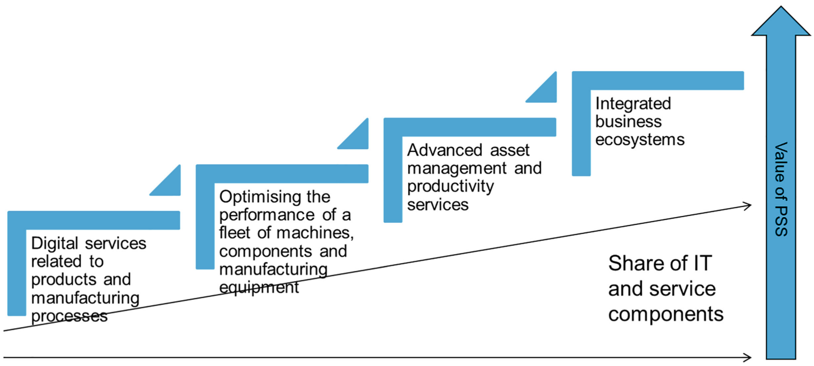


Fig. 1. Advanced forms of the PSS increase the importance and value of IT and service components

The most advanced solution providers are already active or have plans to be involved in providing PSSs to their customers. All the actors in the manufacturing industry have been adopting or have considered adopting digital services as a part of their offering, or to enhance their manufacturing processes, but there are significant differences in the level of digitalization between actors. Subcontractors are adopting digital services in their service portfolio more slowly than solution providers or IT infrastructure providers. All the actor groups see a transition towards deeper partnerships and ecosystems in the future. This is enabled by connecting all the machines to the internet, if there is a benefit seen in the connection. The connection of machines and applications is enabled by platforms, which are of rapidly growing importance (e.g., Porter and Heppelmann 2014; Yoo et al. 2012). Solution providers seem to be the most likely leaders in developing both advanced asset management and productivity services, and integrated business ecosystems. However, the other actors, or the functions they provide, are also needed in the complex integrated business ecosystems of the future.

Companies provide a PSS to their customers based on the data created in their business ecosystem. Advanced manufacturing companies already offer knowledge as a service component, as a part of their PSS. In the future, an increasing number of companies may move from providing data or information as a service, to providing comprehensive knowledge-based services as a part of their PSS.

The research explores the state of digitalization in the manufacturing sector. As such it contributes to the scare body of literature that deals with digitalization enabled PSS. Our research also raises up some important topics like data ownership and changing service delivery beyond traditional PSS. These findings should also be reflected in the future PSS research. This paper gives managers guidelines on how to develop a more advanced digitalization-enabled PSS in their companies. Additionally, it enables the comparison of their digitalization-enabled PSS to some of the more advanced companies. The case study research has been criticized of its lack of rigor and basis for generalization, and requirement of lots of resources and time to conduct (Yin 2003). The case studies presented in this paper rely on qualitative data, which can be regarded as a limiting factor. In this study, the number of case companies and interviews in the three actor groups is limited which affects the generalisability of results. The effects of the limited amount of case companies and interviews on reliability of results are diminished by careful sampling of companies and interviewees.

Acknowledgements. This work has been supported by the Strategic Research Council at the Academy of Finland, project Digital Disruption of Industry (grant number 292889).

References

Ailisto, H., Mäntylä, M., Seppälä, T., Collin, J., Halén, M., Juhanko, J., Jurvansuu, M., Koivisto, R., Kortelainen, H., Simons, M., Tuominen, A. Uusitalo, T.: Finland—the silicon valley of industrial internet. Publications of the Government's analysis, assessment and research activities 10/2015 (2015)

Baines, T., Lightfoot, H.W.: Servitization of the manufacturing firm. Int. J. Oper. Prod. Manage. **34**(1), 2–35 (2013)

Coyne, I.T.: Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? J. Adv. Nurs. **26**(3), 623–630 (1997)

Eisenhardt, K.M.: Building theories from case study research. Acad. Manag. Rev. **14**(4), 532–550 (1989)

Gartner: IT Glossary (2016). http://www.gartner.com/it-glossary/digitalization/. Accessed 21 April 2016

Goedkoop, M.J., van Halen, C.J.G., te Riele, H.R.M., Rommens, P.J.M.: Product service systems, ecological and economic basics. Ministry of Housing, Spatial Planning and the Environment Communications Directorate, the Netherlands (1999)

Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener. Comput. Syst. **29**, 1645–1660 (2013)

Lerch, C., Gotsch, M.: Digitalized product-service systems in manufacturing firms: a case study analysis. Res.-Technol. Manage. **58**(5), 45–52 (2015)

Lusch, R., Nambisan, S.: Service innovation: a service-dominant logic perspective. MIS Quart. **39**(1), 155–175 (2015)

Miltenburg, J.: Manufacturing Strategy: How to Formulate and Implement a Winning Plan, 2nd edn. Taylor & Francis, Abingdon (2005)

Morse, J.M.: Strategies for sampling. In: Morse, J.M. (ed.) Qualitative Nursing Research: A Contemporary Dialogue, pp. 127–145. Sage, Newbury Park (1991)

Porter, M.E., Heppelmann, J.E.: How smart, connected products are transforming competition. Harv. Bus. Rev. **92**, 64–88 (2014)

Rymazewska, A., Helo, P., Gunasekaran, A.: IoT powered servitization of manufacturing – an exploratory case study. Int. J. Prod. Econ. **192**, 92–105 (2017)

de Senzi Zancul, E., Takey, S.M., Barquet, A.P.B., Kuwabara, L.H., Miguel, P.A.C., Rozenfeld, H.: Business process support for IoT based product-service systems (PSS). Bus. Process Manage. J. **22**, 305–323 (2016)

Tukker, A.: Eight type of product-service system: eight ways to sustainability? Experiences from SusProNet. Bus. Strategy Environ. **13**, 246–260 (2004)

Yin, R.: Case Study Research: Design and Methods, 3rd edn. Sage Publications, London (2003)

Yoo, Y., Boland, R.J., Lyytinen, K., Majchrzak, A.: Organizing for innovation in the digitized world. Organ. Sci. **23**(5), 1398–1408 (2012)

Creating Value from Fleet Life-Cycle Data in Business Ecosystem

Sini-Kaisu Kinnunen[1(✉)], Salla Marttonen-Arola[2], and Timo Kärri[1]

[1] Lappeenranta University of Technology, Lappeenranta, Finland
{sini-kaisu.kinnunen, timo.karri}@lut.fi
[2] University of Sunderland, Sunderland, UK
salla.marttonen-arola@sunderland.ac.uk

Abstract. To manage successfully a fleet of assets requires data collection from a fleet that can be distributed globally and to several companies. Thus, data collection is often conducted by multiple actors in business ecosystem, which makes it difficult to get access to all the data concerning a fleet. There is a huge potential to benefit from fleet data due to increasingly gathered data, Internet of Things technologies, and data analysis tools. It is important to demonstrate the value that can be achieved by systematically utilizing fleet data as a support of fleet-level decision making. In this paper, a conceptual model is proposed to illustrate the costs and benefits of fleet life-cycle data utilization in business ecosystem. The model has been developed based on the prior literature and research conducted in collaboration with industry. An example ecosystem is proposed, formed by an equipment manufacturer, its customer company, and an information service provider. The model demonstrates the costs and benefits for each actor in the ecosystem and works as a managerial tool to develop the collaboration, fleet data utilization, service development, and data-based value creation in the ecosystem. The results deepen the scientific discussion about value of information and emphasize the importance of measuring the benefits that need to exceed the costs of data refining in order to create value from data. Further research focuses on the actual modelling based on the structure presented in this paper.

1 Introduction

Data from industrial assets are increasingly gathered with advanced technologies, and Internet of Things (IoT) applications and services are developed based on the data. Technologies for data collection are developed but the focus is shifting to the utilization of the data in decision-making. The massive amounts of data need to be refined and utilized to create value rather than only be stored in data warehouses and systems. In asset management, the technologies have enabled analyzing and predicting the state of assets and managing large groups of assets, i.e. fleets. Assets of fleet can be similar units, share features or certain purposes why they are considered and managed as a fleet. The benefit of managing asset fleets is improving the management of assets so that it enables to gain cost savings and maximize profit during the life cycle of an asset. Cost savings and increased revenues can be related e.g. to improved maintenance

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 881–888, 2020.
https://doi.org/10.1007/978-3-030-48021-9_97

planning (Feng et al. 2017), improved resource and capacity utilization (Archetti et al. 2017), increased asset availability (Gavranis and Kozanidis 2015), and learned best practices (Galletti et al. 2010). However, the challenge in fleet-level analysis is that the fleet data is often fragmented and none of the actors in ecosystem has access to the full life-cycle data of assets. Often manufacturer has certain technical data of their products, but the process and event data is stored in customer's systems. In addition, other actors, such as service providers, might be involved. Thus, the fleet-level consideration requires the reviewing of eco-system, i.e. who has been involved in producing, storing and refining the data. It is meaningful to look into the process in order to develop the value creation from data.

Literature lacks models and frameworks that present how fleet data can be turned into value in business ecosystem. To develop data utilization and to create value for business, research is needed to increase the understanding how the costs and benefits of data utilization are realized and how value is created. In fleet data utilization, the key is to understand how the actors of ecosystem around a fleet are involved in data refining process and what are the costs and benefits for each actor. To respond this need, this paper aims to clarify the generation of costs and benefits in the eco-system around a fleet. The research question is: *How can the costs and benefits of utilizing fleet data in decision-making be categorized for each actor in ecosystem?*

The research question is answered by developing a model that illustrates the generation of costs and benefits of fleet data in the ecosystem around a fleet. This paper presents the structure of the model and works as a basis for further research in which the analytical modelling and testing is conducted. The model is developed based on the prior research and research work done in research program dealing with the topic of the service solutions for fleet management (DIMECC S4Fleet). The themes have been studied in close collaboration with companies involved in the program. Different types of working methods, such as seminars and workshops, have been used and empirical materials, such as interviews and secondary data (e.g. discussions, company brochures), have had an effect on the research process. Theoretical background for the model is discussed in Sect. 2 in which the essential literature is reviewed and theoretical frame is presented. Section 3 presents the structure of the model and Sect. 4 concludes results and proposes ideas for further research.

2 Theoretical Frame of Benefits and Costs in Fleet-Level Decision-Making

The benefits of fleet management are recognized in literature and the basic idea is to improve the management of assets so that cost savings and maximized profits can be achieved during the life cycle of an asset. Fleets are widely discussed in marine, military, logistics, and aviation industries, where fleets consist of ships, aircrafts, trucks, or other vehicles (Archetti et al. 2017; Feng et al. 2017; Tran and Haasis 2015). Machineries and equipment are also considered as fleets in asset management and maintenance contexts (Al-Dahidi et al. 2016). Fleet management and decisions concerning fleets are often related to optimization problems, including capacity and routing problems, but also to proactive and real-time decisions e.g. in maintenance planning and performance monitoring, as well as to strategic decisions, including e.g. investment management (Al-Dahidi et al. 2016; Archetti et al. 2017; Feng et al. 2017; Richardson et al. 2013; Tran and Haasis 2015). Fleet management and decisions at fleet level pursue the economies of scale, minimizing unit costs, and maximizing profits during the lifetime of assets (Tran and Haasis 2015; Archetti et al. 2017). Savings can be e.g. savings in maintenance operations, spare parts, quality costs, and decreased downtime (Al-Dahidi et al. 2016; Feng et al. 2017; Gavranis and Kozanidis 2015; Yongquan et al. 2016). Other benefits can be e.g. increased revenues in the form of new service sales and new product development (Kortelainen et al. 2016). The connection between fleet-level decision-making, supportive analyses and models, achievable benefits and generated costs is described in Fig. 1. Literature has presented pieces of this puzzle (see e.g. Al-Dahidi et al. 2016; Berghout and Tan 2013; Feng et al. 2017; Gavranis and Kozanidis 2015; Miragliotta et al. 2009; Richardson et al. 2013) but the pieces have not been previously put together, as is presented in Fig. 1.

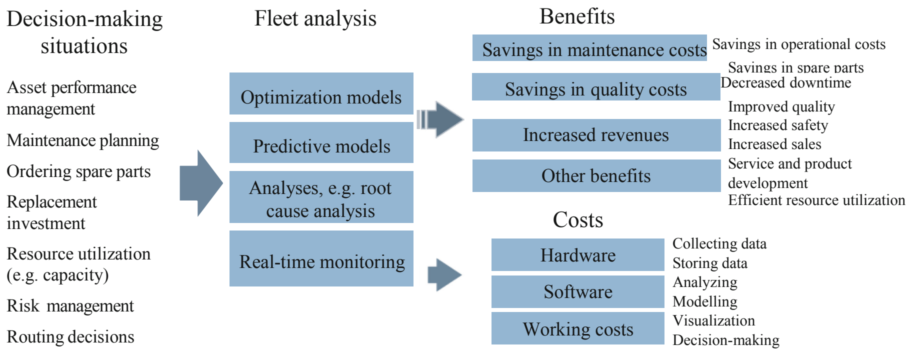


Fig. 1. Theoretical frame of benefits and costs in fleet-level decision-making

Although the benefits of fleet level analysis are seen, the comprehensive understanding of the value of fleet information is limited. Value of information is challenging to determine which can be seen as relatively slight attempts to evaluate the value of data or information. Instead, discussion focus on describing single applications and models to solve certain decision-making needs (see e.g. Archetti et al. 2017). However, there is literature that describe the principles to the value of information in general (Evans and Price 2012; Moody and Walsh 2002). For example, it can be said that the value of information increases with accuracy and use and it is valuable only when utilized in decision-making (Moody and Walsh 2002). When discussing the value of fleet data we need to bring other aspects to discussion, such the roles of actors in ecosystem, due to the nature of fragmented fleet data (Kinnunen et al. 2017). The challenge is that the value is often contextual (Evans and Price 2012) and the value of fleet data vary depending from which perspective it is evaluated. In the case of fleet information, data is generated, processed and utilized by multiple actors in ecosystem around a fleet. Who should actually benefit from the data and to whom the value should be maximized? Should the asset owner get the most value from information concerning the fleet (Archetti et al. 2017) or equipment provider who delivers the assets, and can e.g. utilize the data in product development to gain value in a form of improved products and increased new product or service sales (Kortelainen et al. 2016)? And what is the role of information service provider, who refines the fleet data into analyses and models, in this value formation?

When considering the value of information the costs of data processing need to be considered (see Fig. 1). Costs are generated through the whole process from data collection to storage, maintenance, and utilization of information (Miragliotta et al. 2009; Moody and Walsh 2002). Costs of data utilization can include costs of initial (e.g. hardware), running (e.g. software licenses and maintenance), and other organizational (e.g. personnel working and training) costs (Berghout and Tan 2013). In the case of fleet information it is essential acknowledge that the costs and benefits are realized to multiple actors in ecosystem depending on which phases and to which extend they participate in data refining process (Kinnunen et al. 2017). However, the ecosystems around fleets are distinctive and the roles in fleet data refining process may differ depending the case, and thus the roles need to be defined and modelled in each case separately.

3 Model to Illustrate the Value of Fleet Information in Ecosystem

The conceptual model presented in this section is developed based on prior literature and supported by empirical observations in the research program. Previously presented model (Kinnunen et al. 2016) discusses the value of fleet information and it is extended to ecosystem level in this paper. The model illustrates the value formation of fleet data for the actors of ecosystem. We have developed the model for an example ecosystem consisting of an equipment provider, its customer company, and an information service provider. The structure of the model is presented in Fig. 2 and the components of the model are described in this section.

The ecosystem consisting of three actors represents a simplified example of the fragmented fleet data refining process from data collection to the value for each actor. Typically, the ecosystem includes an equipment provider who manufactures the assets and owns e.g. product development data of the assets. Customer company usually owns the assets and has the most extensive data from the operations phase of assets' life cycle. In addition, service providers are often involved in the fleet life-cycle data generation by producing and mastering e.g. service history data. In this model, we have an information service provider as the third actor. Information service provider provides e.g. the technical solution for fleet level analysis. As a presumption in the model, the fleet is the same for each actor: the assets owned by the customer and provided by the equipment provider. Equipment provider supports customer's maintenance decisions by providing fleet level analysis. Both, customer and equipment provider, utilizes fleet life-cycle data also for their own decision-making, e.g. equipment provider can use fleet data to support product development.

In Fig. 2, the value for each actor is illustrated as the components of costs and benefits followed from data processing and data utilization in fleet decisions. In

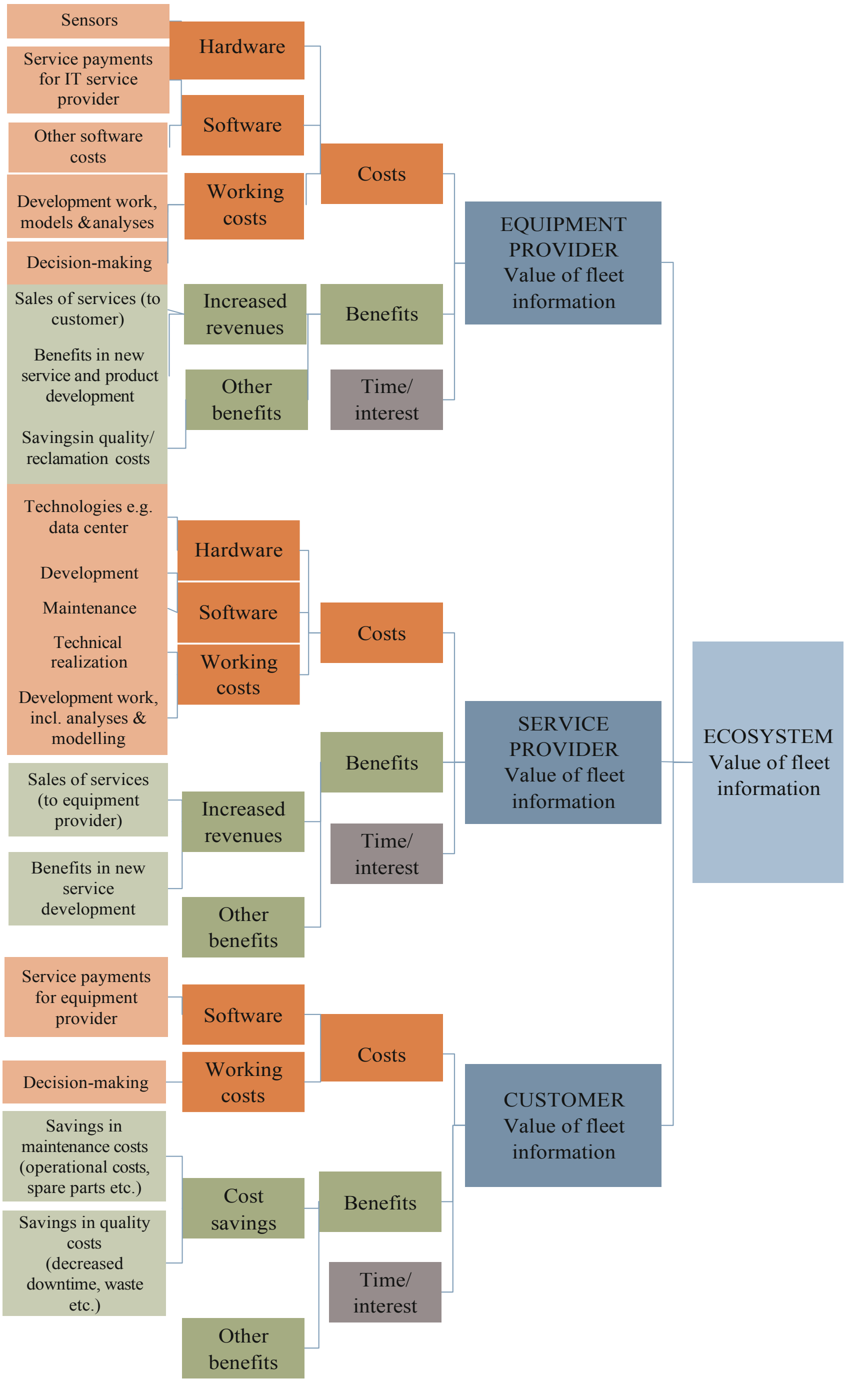


Fig. 2. Structure of the model

general, benefits for each actor are defined as cost savings, increased revenues or other benefits, as was discussed in theoretical frame in Sect. 2. Benefits are monetary benefits that can be achieved when better decisions are made because of fleet-level analyses. Cost savings are gained if e.g. maintenance costs can be reduced with better decisions compared to the situation where the fleet-level analyses are not available. Each actor of ecosystem views the fleet data utilization from their own perspective and thus, the benefits for each actor are different and reflect their decision making needs. In the model, costs are monetary costs that are followed from data collection, data processing and data utilization. In general, the costs include hardware, software and data processing-related work costs, as was presented in theoretical frame in Sect. 2. Because it is assumed that the actors of ecosystem are responsible for different phases of the data refining process, e.g. equipment provider and customer gather fleet data and information service provider is mainly responsible for data processing, different costs are caused for each actor. Next, the benefits and costs from each actor's perspective are discussed.

First, from customer's perspective the benefits are related to asset management decisions and include, e.g. savings in maintenance and quality costs and other benefits. Costs in maintenance can be reduced, i.e. cost savings can be achieved, when better decisions can be made concerning the fleet of assets thanks to informative fleet analyses. According to empirical findings, significant potential is related to savings in decreased downtime from customer company's point of view. In other words, customer can reduce downtime by predicting the maintenance needs more effectively and gain savings in quality costs. For customer the cost components of processing fleet data include service costs for equipment provider, who is assumed to offer the fleet analysis services, and working costs derived from the utilization of analyses and decision-making. Costs can be non-recurring or recurring costs by nature.

Secondly, from equipment provider's perspective the benefits of fleet analyses are increased revenues that can be gained from increased product and service sales, and other benefits, such as savings in reclamation costs and benefits in product development. Costs from data processing for the equipment provider include (1) hardware costs (e.g. sensors in assets), (2) software costs for information service provider for providing the technical solution of fleet analysis, and (3) working costs including development work used to develop the fleet analyses in collaboration with information service provider, and utilization of analyses e.g. in own product development decisions.

Thirdly, the benefits for information service provider include revenues from service sales for equipment provider, and other benefits, e.g. benefits for service development in future. These other benefits can be evaluated for example as annual monetary benefits. Costs for information service provider are mostly from the technical solution development of fleet analyses, including hardware costs (e.g. data center), software development and maintaining costs, and working costs of technical realization and development.

Finally, the value of fleet information is considered as economic value for the investment in fleet data refining process. In Fig. 2, time or interest suggest that the costs and benefits of fleet data utilization are realized during certain period and when evaluating the value, the time value of money need to be considered. The value for each actor can be viewed e.g. as the net present value of the discounted total benefits and total costs. The value can be evaluated for each actor separately but the value of fleet information can be defined for the whole ecosystem. However, if we consider the value of ecosystem as a whole, it is not obvious which value we should maximize. The model can be adapted to other ecosystems but the fleet and the actors in the ecosystem need to be defined. The roles in fleet data refining process may vary as well as the benefits and costs for each actor.

4 Conclusions

The conceptual model was developed to demonstrate the value of fleet life-cycle data for the actors in fleet ecosystem. The model works as a basis for further research that focus on analytical modelling of the value of fleet life-cycle data in eco-system. Understanding the value of fleet data is important for companies in order that they can develop the fleet data utilization both inside company but also with the stakeholders in their fleet ecosystem. The fleet data are usually fragmented between multiple actors and actors in ecosystem need to develop the fleet life-cycle data management process to generate the value. Because the data has the value only when used, it is important to promote the utilization of data in decision making, instead of just storing the data in data warehouses. The efforts to assess the value of information in literature are scarce and thus more research is needed to evaluate and to model the value of information.

The developed model presents a simplified example of fleet ecosystem and the value of fleet data for each actor in ecosystem. The model can work as a tool to understand the relations of actors in fleet data refining process and how the ecosystem can more efficiently create value from fleet data that is currently fragmented and the full potential of data remains unused. If we consider the value of information in ecosystem we need to remember that targets of actors may differ and the value that should be maximized need to be considered case-specifically. Further research should focus on examining the value of data for business in ecosystems, which can be done e.g. by modelling the logic of valuation and by testing and simulating case ecosystems with real or illustrative data.

References

Al-Dahidi, S., Di Maio, F., Baraldi, P., Zio, E.: Remaining useful life estimation in heterogeneous fleets working under variable operating conditions. Reliab. Eng. Syst. Safe. **156**, 109–124 (2016)

Archetti, C., Bertazzi, L., Laganà, D., Vocaturo, F.: The undirected capacitated general routing problem with profits. Eur. J. Oper. Res. **257**, 822–833 (2017)

Berghout, E., Tan, C.-W.: Understanding the impact of business cases on IT investment decisions: an analysis of municipal e-government projects. Inf. Manage. **50**(7), 489–506 (2013)

Evans, N., Price, J.: Barriers to the effective deployment of information assets: an executive management perspective. Interdiscip. J. Inf. Knowl. Manag. **7**, 177–199 (2012)

Feng, Q., Bi, X., Zhao, X., Chen, Y., Sun, B.: Heuristic hybrid game approach for fleet condition-based maintenance planning. Reliab. Eng. Syst. Safe. **157**, 166–176 (2017)

Galletti, D.W., Lee, J., Kozman, T.: Competitive benchmarking for fleet cost management. Total Qual. Manag. Bus. Excell. **21**(10), 1047–1056 (2010)

Gavranis, A., Kozanidis, G.: An exact solution algorithm for maximizing the fleet availability of a unit of aircraft subject to flight and maintenance requirements. Eur. J. Oper. Res. **242**(2), 631–643 (2015)

Kinnunen, S.-K., Hanski, J., Marttonen-Arola, S., Kärri, T.: A framework for creating value from fleet data at ecosystem level. Manag. Syst. Prod. Eng. **25**(3), 163–167 (2017)

Kinnunen, S.-K., Marttonen-Arola, S., Kärri, T.: Value of fleet information in asset management. In: Proceedings of 6th International Conference on Maintenance Performance Measurement and Management, pp. 76–80, Luleå, Sweden, 28 November 2016

Kortelainen, H., Happonen, A., Kinnunen, S.-K.: Fleet service generation – challenges in corporate asset management. In: Koskinen, K.T., Kortelainen, H., Aaltonen, J., Uusitalo, T., Komonen, K., Mathew, J., Laitinen, J. (Eds.) Proceedings of the 10th World Congress on Engineering Asset Management. Lecture Notes in Mechanical Engineering, pp. 373–380. Springer, Cham (2016)

Miragliotta, G., Perego, A., Tumino, A.: A quantitative model for the introduction of RFId in the fast moving consumer goods supply chain: are there any profits? Int. J. Oper. Prod. Manag. **29** (10), 1049–1082 (2009)

Moody, D., Walsh, P.: Measuring the value of information: an asset valuation approach. In: Morgan, B., Nolan, C. (eds.) Guidelines for Implementing Data Resource Management (2002)

Richardson, S., Kefford, A., Hodkiewicz, M.: Optimized asset replacement strategy in the presence of lead time uncertainty. Int. J. Prod. Econ. **141**, 659–667 (2013)

Tran, N.K., Haasis, H.-D.: An empirical study of fleet expansion and growth of ship size in container liner shipping. Int. J. Prod. Econ. **159**, 241–253 (2015)

Yonquan, S., Xi, C., He, R., Yingchao, J., Quanwu, I.: Ordering decision-making methods on spare parts for a new aircraft fleet based on a two-sample prediction. Reliab. Eng. Syst. Safe. **156**, 40–50 (2016)

Context-Awareness in Internet of Things - Enabled Monitoring Services

Ali Al-Shdifat(✉), Christos Emmanouilidis, and Andrew Starr

Cranfield University, Cranfield, UK
{ali.alshdifat, christosem, a.starr}@cranfield.ac.uk

Abstract. Remote monitoring services are required to meet the very high demands on availability and efficiency of industrial systems. The fast evolution of technologies associated with the deeper penetration of Internet of Things in industry creates considerable challenges for such services. These are related to the whole data lifecycle, encompassing data acquisition, real-time data processing, transmission, storage, analysis, and higher added value service provision to users, with adequate data management and governance needed to be in place. The sheer complexity of such activities the need to ground such processing on sound domain knowledge emphasises the need for context information management. The aim of this paper is to survey and analyse recent literature that addresses Internet of Things context information management, mapping how context-aware computing addresses key challenges and supports delivering appropriate monitoring solutions.

Keywords: Internet of Things · Context management · Remote monitoring services

1 Introduction

In recent years, research into context management for Internet of Things (IoT) has received increased attention in academia, aimed to address the increasing complexity challenges of IoT-enabled data value chains (Perera et al. 2015). When considering IoT usage in industrial environments, the term Industrial Internet of Things (IIoT), or simply Industrial Internet, is often employed, and is being considered synonymous to Industrie 4.0 (Jeschke et al. 2017). IoT brings together many functionalities such as identification, sensing, communication, computation, services, and semantic information management (Al-Fuqaha et al. 2015). Although it offers many benefits and solution enablers, substantial effort is required to manage and exploit the data generated by things. Among the key instruments to tackle such complexity is the concept of context information management. The term context in computing, originally employed in computational linguistics and later adopted in web-based information management, refers to establishing the background circumstances or specific situation regarding specific data or computing service requests. Context-awareness is the ability of a system to give appropriate information or services to consumers utilising context information (Abowd et al. 1999).

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 889–896, 2020.
https://doi.org/10.1007/978-3-030-48021-9_98

With the deeper penetration of IoT technologies in monitoring tasks, the need for context information management increasingly manifests itself as a requirement for industrial applications. Context gathering, modelling, reasoning and dissemination are needed for the efficient handling of vast amounts of data, produced by numerous devices, and their efficient integration in enterprise systems. This is further fueled by the accelerating shift to service-based business models, wherein service level agreements must be ascertained, supported by adequate monitoring systems.

This paper offers a survey and analysis of recent literature that addresses context management in IoT. It maps how context-aware computing techniques have contributed to delivering solutions and identifies key challenges that IoT-enabled monitoring services need to address. The rest of this paper is organised as follows: Sect. 2 provides a review of context-awareness, including a critical analysis of their strengths and weaknesses. Section 3 deals with context information lifecycle management, offering also a mapping view of context-awareness in IoT for remote monitoring services. The analysis results in identifying some key challenges to be addressed by further research in the field, as summarised in conclusion.

2 Context Awareness in IoT

Context-awareness has an increasingly significant role to play in deploying IoT solutions in complex industrial environments. Many different definitions of context are reported in the literature. Abowd et al. (1999) have argued that context is used to give necessary information and services to the consumer, where relevance depends on the consumer's task. According to Dey et al. (2001), some definitions relied on examples and therefore could not be utilised to identify a new context, proposing instead to employ synonyms of context, such as environment and situation. Five W's (Who, What, Where, When, Why) were identified as the basic information that is required to understand context (Abowd and Mynatt 2000). The management of large-scale sensing was recognised as a prime target for context management, as such data need to be gathered, modelled, analysed, fused, and interpreted (Raskino et al. 2005). The data produced by sensors may not supply the useful information that could be utilised to understand the whole situation. Consequently, additional knowledge and context-relevant information may have to be fused with sensor data for successful context identification. In order to address this, middleware solutions have been proposed, addressing various aspects of IoT data management, such as context-awareness, interoperability, device management, platform portability, as well as security and privacy (Perera et al. 2014). Several surveys have been conducted in this area (Bellavista et al. 2012; Kjaer 2007; Molla and Ahamed 2006; Saeed and Waheed 2010). The viewpoint of such surveys is mapped in chronological order from left to right in Fig. 1.

Surveyed worked shows that context management has largely dealt with the challenges of ubiquitous environments, as well as the data heterogeneity and services scalability. It plays a central role in defining what data needs to be collected and how to be processed, as well as in determining what information and services that require being presented to the consumer. Context management issues increasingly progressed from dealing with context acquisition and modelling to context reasoning and dissemination.

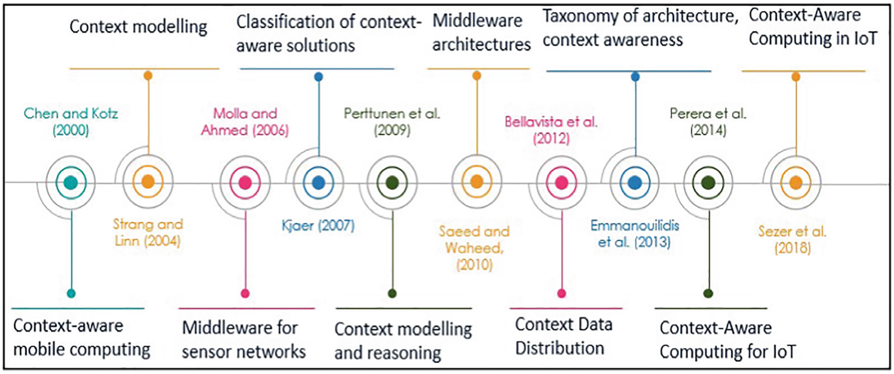


Fig. 1. Summary of surveys on context awareness

Early context information management literature targeted mobile computing and web-based information processing. IoT has expanded the range of applications with substantial needs for context management, and this was reflected in the focus of relevant surveys. Nonetheless, while substantial research efforts have been devoted to context lifecycle management in web-based, mobile, and ubiquitous computing, including IoT-enabled computing, little attention has been given to translate these advances to tangible progress in remote monitoring services. Various types of context have been identified by researchers based on different perspectives (Henricksen 2003; Chong et al. 2007; Guan et al. 2007; Rizou et al. 2010; Li et al. 2015; Valverde-Rebaza et al. 2018). Abowd et al. (1999) distinguished context between primary and secondary, as well between conceptual and operational. Operational context can be further classified as sensed, static, profiled, and derived. Chen and Kotz (2000) distinguish between passive and active context, depending on whether context is directly actionable or not, considering the way it is used in applications. Liu et al. (2011) classify context into user, physical, and networking. An overview of different context categorisation schemes in chronological order is shown in Fig. 2.

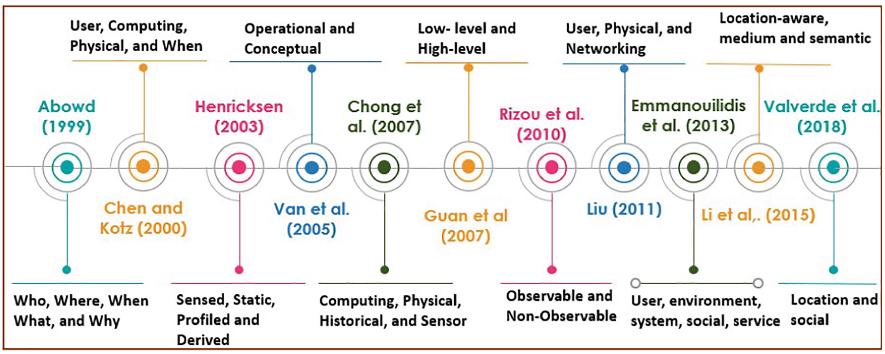


Fig. 2. Different context categorisation schemes

An outline of strengths and weaknesses of typical context classification approaches are depicted in Fig. 3.

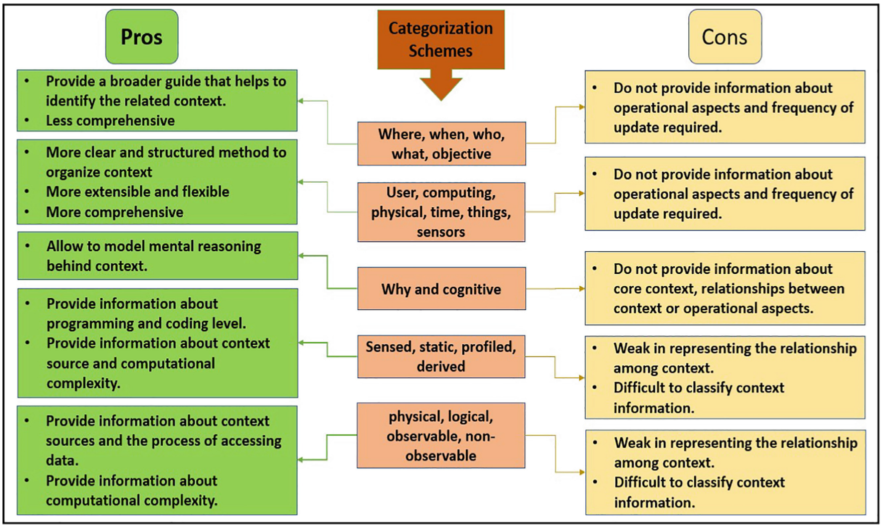


Fig. 3. Comparison of context categorisation schemes adapted from Perera et al. (2014)

Based on the assessment of context classification, it is clear that existing context classification schemes have weaknesses and it is particularly unclear to what extent they meet needs for monitoring services. Therefore, in order to design an appropriate framework to manage context for IoT-enabled monitoring services efficiently, further analysis is needed. In order to do so, an outline of how the lifecycle of context information can be managed is first introduced.

3 Monitoring Services Context Information Life Cycle

Context lifecycle refers to how data is gathered, modelled, processed, and how knowledge is deduced from the obtained data (Sezer et al. 2018). The context lifecycle management generally consists of four steps, namely context acquisition, modelling, reasoning, and dissemination (Perera et al. 2014). However, a more detailed handling and analysis of what these steps actually involve when considering monitoring services are largely missing in the relevant literature. Figure 4 offers an illustration of the different stages of context information management, placing them against the monitoring services functionality.

Context acquisition involves acquiring and bringing together data from physical objects in different ways, based on sensor types, responsibility, acquisition process, frequency, and source (Aguilar et al. 2018). From a remote monitoring perspective, it would be of contextual relevance to understanding what type of measurement data need

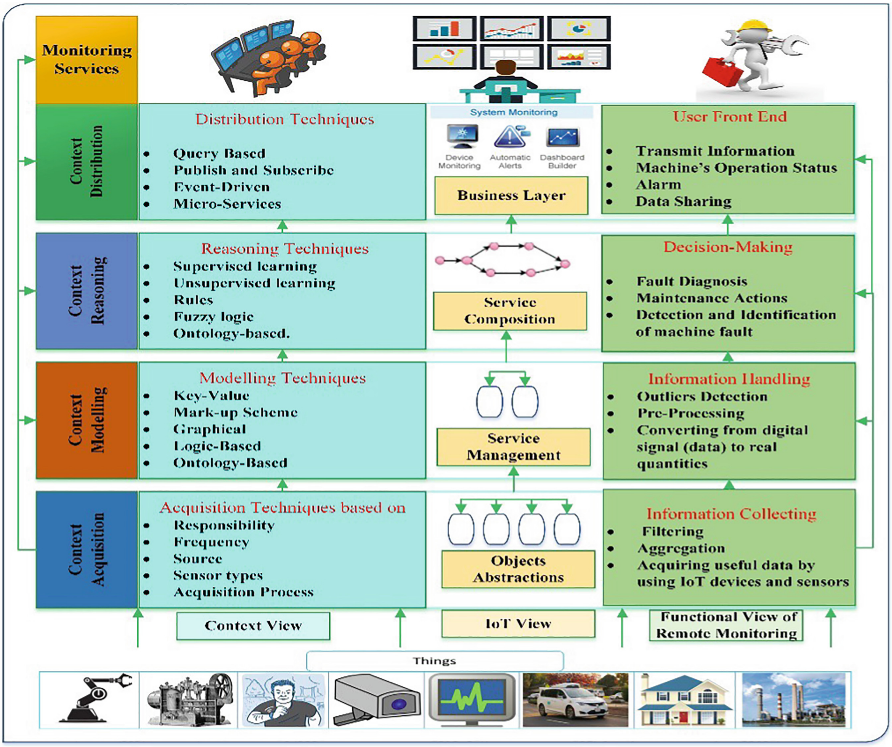


Fig. 4. Context management lifecycle in IoT for remote monitoring services

to be collected (e.g. temperature, vibration, and pressure) by using IoT devices and sensors, making also sure that these are indeed acquired. **At context modelling level,** is further generally referred to as representation and formalization of the context, through certain modelling approaches (Cabrera et al. 2017). Context modelling techniques have been surveyed by (Chen and Kotz 2000; Strang and Linnhoff-Popien 2004), and include Ontology-Based, Key-Value, Logic-Based, Markup Scheme, and Graphical ones. From an IoT perspective, service management enables to work with heterogeneous objects, and it also concerns the operations to manage and orchestrate the services exposed through it. From a remote monitoring perspective, information handling is the process of converting digital data into real quantities of working conditions of machines to produce meaningful information, while filtering out perceived outlier data.

Context Reasoning can be defined as a process that contributes significantly to the collection of new knowledge based on the acquired contextually relevant data (Bikakis et al. 2008). Typical context reasoning techniques include Rule-based approaches, Supervised learning, Fuzzy logic, Unsupervised learning, and Ontology-based ones (Bikakis et al. 2008; Perttunen et al. 2009). The importance of this layer of context management lies the ability to provide high-quality intelligent services to meet end-user needs. In remote monitoring ser-vices, this can enable not only fault detection, diagnostics, and prognostics, but also action recommendations consistent with the inferred context of the analysed situation.

Context Dissemination. This is where actionable context is made available to other applications and services, or users. Two methods are typically used for context distribution 1) Query: The user requests the context, such that the context management system answers to that query. 2) Subscription also called publication (Perera et al. 2014). This constitutes the high-end of the IoT-generated data process chain and can fuel added value services, such as visual or other types of analytics, as well as decision support. From the end-user's perspective, this stage is actually the most important, as the initial data are now disseminated in enhanced form and are essentially converted to visual information, insights and action recommendations.

4 Conclusion and Further Research

This paper's aim was to outline and analyse key issues related to context management for IoT – enabled remote monitoring services. This highlighted the need for handling the whole context information management lifecycle, from context acquisition and modelling, through reasoning, all the way to context dissemination and the relevance that each such phase has to monitoring services. IoT has expanded the range of applications and the scale of involved data, creating a clear need for context management, and this was reflected on the recent focus of relevant surveys. Such research is contributing towards filling the gap in relevant literature, which focused on context lifecycle management in web-based, mobile, and ubiquitous computing, including IoT-enabled computing, while paying little attention to translating these advances to tangible progress in remote monitoring services. Consequently, further research is required to develop context-aware approaches and architectures to deliver more efficient IoT-enabled monitoring services, including non-functional issues, such as IoT security, which constitute a critical adoption barrier in current IIoT – enabled systems.

References

Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a better understanding of context and context-awareness. In: Proceedings of 1st International Symposium on Handheld and Ubiquitous Computing, HUC 1999, London, UK, pp. 304–307. Springer (1999)

Abowd, G.D., Mynatt, E.D.: Charting past, present, and future research in ubiquitous computing. ACM Trans. Comput. Hum. Interact. **7**(1), 29–58 (2000)

Aguilar, J., Jerez, M., Rodríguez, T.: CAMeOnto: context awareness meta ontology modeling. Appl. Comput. Inform. **14**(2), 202–213 (2018)

Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M.: A survey on Internet of Things. IEEE Commun. Surv. Tutor. **17**(4), 2347–2376 (2015)

Bellavista, P., Corradi, A., Fanelli, M., Foschini, L.: A survey of context data distribution for mobile ubiquitous systems. ACM Comput. Surv. **44**(4), 1–45 (2012)

Bikakis, A., Patkos, T., Antoniou, G., Plexousakis, D.: A survey of semantics-based approaches for context reasoning in ambient intelligence. Constr. Ambient Intell. **11**, 14–23 (2008)

Cabrera, O., Franch, X., Marco, J.: 3LConOnt: a three-level ontology for context modelling in context-aware computing. Softw. Syst. Model. **18**, 1345–1378 (2019)

Chen, G., Kotz, D.: A survey of context-aware mobile computing research. Dartmouth Computer Science Technical Report, 3755, pp. 1–16 (2000)

Chong, S.K., McCauley, I., Loke, S.W., Krishnaswamy, S.: Context-aware sensors and data muling. In: Context Awareness for Self-managing Systems (Devices, Applications and Networks) Proceeding, pp. 103–117. VDE-Verlag, Berlin (2007)

Dey, A., Abowd, G., Salber, D.: A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. Hum. Comput. Interact. **16**(2), 97–166 (2001)

Emmanouilidis, C., Koutsiamanis, R.A., Tasidou, A.: Mobile guides: taxonomy of architectures, context awareness, technologies and applications. J. Netw. Comput. Appl. **36**(1), 103–125 (2013)

Guan, D., Yuan, W., Lee, S., Lee, Y.K.: Context selection and reasoning in ubiquitous computing. In: Proceedings of the 2007 International Conference on Intelligent Pervasive Computing, IPC 2007, pp. 184–187 (2007)

Henricksen, K.: A framework for context-aware pervasive computing applications, The School of Information Technology and Electrical Engineering, The University of Queensland, pp. 13–20 (2003)

Jeschke, S., Brecher, C., Meisen, T., Özdemir, D., Eschert, T.: Industrial Internet of Things and cyber manufacturing systems. In: Jeschke, S., Brecher, C., Song, H., Rawat, D. (eds.) Industrial Internet of Things Cybermanufacturing Systems. Springer Series in Wireless Technology, pp. 3–19. Springer, Cham (2017)

Kjaer, K.E.: A survey of context-aware middleware. In: Proceedings of 25th Conference on IASTED International Multi-Conference: Software Engineering, pp. 148–155. ACTA Press (2007)

Li, X., Eckert, M., Martinez, J.F., Rubio, G.: Context-aware middleware architectures: survey and challenges. Sens. (Switz.) **15**(8), 20570–20607 (2015)

Liu, W., Li, X., Huang, D.: A survey on context awareness. In: 2011 International Conference on Computer Science and Service System (CSSS), pp. 144–147 (2011)

Molla, M., Ahamed, S.: A survey of middleware for sensor network and challenges. In: Proceedings of 2006 International Conference Workshops on Parallel Processing, ICPPW 2006, pp. 223–228. IEEE Computer Society, Washington, DC (2006)

Perera, C., Liu, C.H., Jayawardena, S., Chen, M.: A survey on Internet of Things from industrial market perspective. IEEE Access **2**, 1660–1679 (2015)

Perera, C., Zaslavsky, A., Christen, P., Georgakopoulos, D.: Context aware computing for the Internet of Things. IEEE Commun. Surv. Tutor. **16**(1), 414–454 (2014)

Perttunen, M., Riekki, J., Lassila, O.: Context representation and reasoning in pervasive computing. Int. J. Multimedia Ubiquit. Eng. **4**(4), 1–28 (2009)

Raskino, M., Fenn, J., Linden, A.: Extracting value from the massively connected world of 2015. Technical report, Gartner Research (2005). https://www.gartner.com/doc/476440/extracting-value-massively-connected-world. Accessed 13 June 2018)

Rizou, S., Haussermann, K., Durr, F., Cipriani, N., Rothermel, K.: A system for distributed context reasoning. In: Sixth International Conference on Autonomic and Autonomous Systems (ICAS), pp. 84–89 (2010)

Saeed, A., Waheed, T.: An extensive survey of context-aware middleware architectures. In: 2010 IEEE International Conference on Electro/Information Technology, EIT 2010 (2010)

Sezer, O.B., Dogdu, E., Ozbayoglu, A.M.: Context aware computing, learning and big data in Internet of Things: a survey. IEEE Internet Things J. **5**(1), 1 (2018)

Strang, T., Linnhoff-Popien, C.: A context modeling survey. In: Workshop on Advanced Context Modelling, Reasoning and Management, UbiComp 2004 - The Sixth International Conference on Ubiquitous Computing, pp. 1–8 (2004)

Valverde-Rebaza, J.C., Roche, M., Poncelet, P., de Andrade Lopes, A.: The role of location and social strength for friendship prediction in location-based social networks. Inf. Process. Manag. **54**(4), 475–489 (2018)

Van Bunningen, A.H., Feng, L., Apers, P.M.G.: Context for ubiquitous data management. Proceedings of the International Workshop on Ubiquitous Data Management, Washington, DC, USA, pp. 17–24, 4 April 2005

Part X: Sustainable Assets and Processes

Asset Management for the Energy Transition

Ype Wijnia(✉)

AssetResolutions B.V., Zwolle, The Netherlands
ype.wijnia@assetresolutions.nl

Abstract. Climate change and the associated rise of the sea level is one of the major concerns for urban deltas, which house about half the population of the world. Part of the mitigation pack is decarbonizing the energy production, by means of renewable energy sources. These are geographically distributed by nature, putting a higher demand on the distribution grid. Because of regulatory pressures, most network operators connect them by means of an incremental strategy. However, decarbonizing the energy system requires electrification. This means the capacity of the grid has to increase with a factor 10 in 30 years. An incremental strategy is unlikely to result in such a growth, on the contrary is more likely to waste resources. In this paper we demonstrate the waste of resources by incremental strategies under strong growth scenarios with a case study of wind turbines in the Netherlands. Visionary strategies can outperform incremental strategies in these conditions by as much as 30%. We end the paper with an appeal to all asset managers to develop visionary strategic plans to support the coming energy transition.

1 Introduction

According to the IPCC (2014), the climate change will result in a sea level rise, projected to be between 0,5 and 1 m by 2100. Furthermore, the risks from extreme weather events tend to increase progressively with further warming. Both risks tend to impact coastal areas and especially deltas, which house about half the population of the world. The ability to deal with these impacts depends strongly on the wealth of nations, and thus may have a very uneven distribution around the world. This is why it is argued that 4 °C warmer world should be avoided (World Bank by the Potsdam Institute for Climate Impact Research and Climate Analysis 2012).

A major element of this preventative strategy is decarbonizing the energy production, by means of renewable sources like wind and sun. However, integrating these into the existing energy system is not a triviality. The average annual production power density of renewables is orders of magnitude below that of traditional power plants. Wind farms reach about 2 megawatt (MW) per square kilometer of earth surface, photovoltaic (PV) about 8 megawatt per square kilometer, compared to over 1 gigawatt per square kilometer for a coal fired plant (based on load factors of 1000 h per year for PV, 3000 h for wind and 8000 h for Coal) (MacKay 2008). Renewable energy is thus distributed much more in the geographical sense than traditional power plants, putting a higher demand on the distribution grid. But renewables are also distributed in terms of decision making, as in principle each wind turbine (i.e. each MW) or PV plant (down to

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 899–906, 2020.
https://doi.org/10.1007/978-3-030-48021-9_99

about 1 kW) can be a separate decision. This is in strong contrast with power plants that come in hundreds to thousands of megawatt at the time. The capacity need for the distribution grid is therefore highly unpredictable.

Many network operators struggle with these new demands. Distribution Network Operations (DNO's) tend to seek a solution in smart grids so that as much renewables as possible can be allowed in the current grids (Jonathan Fox (SP Energy Networks) 2017). This is due to the regulatory regime in most European countries: DNO's are only funded for necessary investments, to be paid by consumers and not producers of electricity. Necessity can only be proven by existing or planned loads, but due to distributed decision making no long term plan exists for renewables. As a result, network operators have to implement an incremental approach, building on what already exists. However, given that decarbonizing the energy system essentially comes down to electrification, the amount of electricity needs to grow with a factor of 10 in the coming 30 years. This means doubling the capacity roughly every 7 years. An incremental approach is unlikely to result in this growth, on the contrary it is much more likely to waste resources following inefficient development paths.

In this paper, we will demonstrate this waste of resources by incremental strategies under strong growth scenarios with a case study of wind turbines in the Netherlands. We will end this paper with a call for more asset management in the energy infrastructure planning.

2 Case Introduction[1]

The situation of the case study is substation Zeewolde, located in the Flevopolder, a part of the Netherlands claimed over the sea in the period between 1950 and 1968. The area is several meters below sea level. The map below shows the Zeewolde area with wind turbines, a power station and power lines marked on the map. The map is a screen shot of the ENIPEDIA power plant database of the Delft University of Technology (Fig. 1).

The land use is mostly agricultural, with a few population centers. The dotted line on the top left is a 380 kiloVolt (kV) line, the dotted line on bottom right is a 150 kV line. To indicate the scale of the map, the distance between those lines is 9,8 km. The Zeewolde 150 kV substation is marked by the red circle. The red dots represent wind turbines. Detailed information on the wind farms can be found in an online database (The Windpower 2012). The table below lists the windfarms in this area (Table 1).

[1] This case is an update from Wijnia (2013), accessible via http://www.assetresolutions.nl/userfiles/Pubs/Delta_Conference_2013paper.pdf.

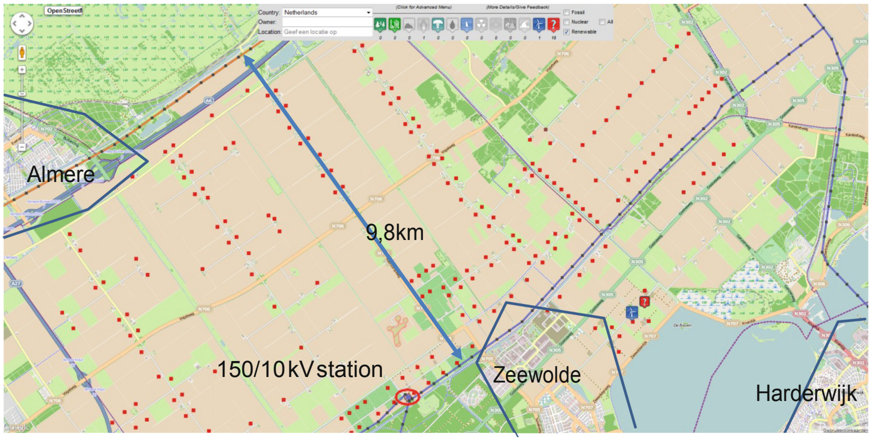


Fig. 1. Map of the Flevoland area showing the locations of wind turbines (ENEPEDIA 2012)

Table 1. List of windfarms in the "Zeewolde Area 4" (The Windpower 2012). The voltage guess is based upon the power. Up to 13 MW can be supplied by a single 10 kV at 800 mm^2 cable.

ID	Name	Power [kW]	Number of turbines	Average power per turbine [kW]	Voltage [kV] (guess)
24	Bloesemlaan	18.150	20	908	20
58	Dodaarsweg	18.780	23	817	20
67	Eemmeerdijk	18.000	18	1.000	20
74	Eolienne	10.000	10	1.000	10
82	Futenweg	10.500	6	1.750	10
92	Gruttoweg	19.710	23	857	20
151	Lepelaarweg	11.400	12	950	10
187	Ooievaarsweg	15.280	16	955	20
204	Pijlstaartweg	36.000	24	1.500	20
207	RachelCarson	18.000	18	1.000	20
208	Reigerweg	12.570	16	786	10
221	Schollevaarweg	20.400	22	928	20
239	Sterappellaan	10.950	12	913	10
289	Wulpweg	14.600	17	859	20
295	Zeewolde	125.560	126	997	20
	Total	**359.900**	**363**	**991**	
	Total 10 kV	**55.420**			
	Total 20 kV	**304.480**			

These wind turbines are connected to the Zeewolde substation. Apparently it was easier to connect them to Zeewolde, even though there are other 150 kV stations in the neighbourhood as shown in Table 2 and Fig. 2.

Table 2. The connected load and production to the stations of Fig. 2 (source: Quality and Capacity document 2011–2016 (LIANDER 2011))

Substation	Voltage [kV]	Capacity 2011 [MVA]	Peak load basis 2012 [MW]	Wind 10 kV [MW]	Wind 20 kV [MW]	Total
Almere	150	132	57,6	0	0	0
De Vaart	150	132	56,8	0	0	0
Zeewolde	150	380	28,2	54,9	301	355,9
Zuiderveld	150	132	59,4	56,2	0	56,2

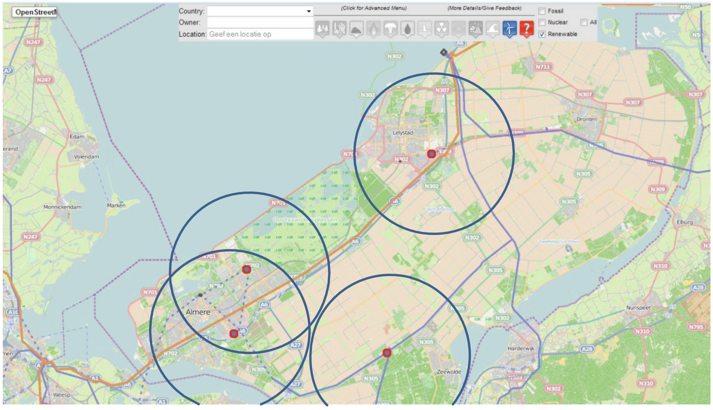


Fig. 2. Substations near the wind turbines

This list reasonably complies with the estimate based on the wind farms. That these numbers match should not be surprising. The wind farm operators would have to pay for the cable connecting their farm to the substation. As long as the only voltage was 10 kV, it makes sense to match the wind farm power output with the cable capacity. Most 10 kV farms therefore have a power in the 10 MW range, whereas 20 kV farms are in the 20 MW range.

3 Alternative Technical Solutions

The wind farms were first connected to the grid using the existing transformers. However, in hindsight this may not have been the best solution. Implementing the 20 kV solution earlier might have been more cost effective from a total societal perspective. And what if the voltage would have been even higher, up to 30 kV? Given the large amount of wind power, that higher Voltage might have been a better solution right from the start.

To answer this question, a simple model has been developed to evaluate the options economically. The average length of the cables to connect a wind farm to substation Zeewolde has been set at 10 km. At the average power per wind turbine is set at 1 MW and the distance between the turbines is about 500 m. It is assumed wind parks could be reshuffled to reach the cable capacity (Table 3).

Table 3. Base data

Item	Operating voltage		
	10 kV	20 kV	30 kV
Cable capacity (630 Al)	10 MW	20 MW	30 MW
Max number of turbines	10	20	30
Base cable length	10 km	10 km	10 km
Additional length	500 m	500 m	500 m
Total length at full capacity	15 km	20 km	25 km
Cable cost	100 k€	110 k€	120 k€
Energy loss at full load per km (2200 h, 40€ per MWh, 20 years equivalent use)	45 kW 135 MWh/yr 6.750 €/yr 135 k€ NPV[a]	45 kW 135 MWh/yr 6.750 €/yr 135 k€ NPV	45 kW 135 MWh/yr 6.750 €/yr 135 k€ NPV
Connection cost per turbine (compact substation)	60 k€	65 k€	70 k€
Connection cost 150 kV substation	100 k€	120 k€	140 k€
Transformer cost 150 kV (150 MW)	2.400 k€	2.400 k€	2.400 k€
Transformer cost per MW	16 k€	16 k€	16 k€

[a]NPV = Net Present Value

These costs are high level estimates, based upon the tariff list of LIANDER (ACM 2012). The tariff list applies to 10 kV connections. The costs for 20 kV and 30 kV are higher, but not exceptionally. Manufacturer TKF supplies their Twenpower cable for all voltages in the table above, with the only difference between cables in isolation thickness (TKF 2012)[2]. The cost difference therefore is estimated at +10% for 20 kV and +20% for 30 kV. The same reasoning holds for the compact substation. The largest part of the costs are the building and its installation, the costs for the switchgear inside is in the range between 10 and 20 thousand Euros. Switchgear for higher voltages is more expensive, but again not exceptionally. ABB supplies their Ring Main Unit SafeRing up to 36 kV (ABB AS Power Products Division 2015). Extra costs of 5 k€ and 10 k€ are used for 20 and 30 kV respectively (Table 4).

[2] The specs used are 50420 (10 kV), 54218 (20 kV), 56195 (30 kV).

Table 4. Cost calculation

Cost item	Operating voltage		
	10 kV	20 kV	30 kV
Feeder (base length plus sub connection)	1.600 k€	2.320 k€	3.140 k€
Transformer cost per cable	160 k€	320 k€	480 k€
Connections	600 k€	1.300 k€	2.100 k€
Total per farm	2.360 k€	3.940 k€	5.720 k€
CAPEX Per turbine (a)	**236 k€**	**197 k€**	**191 k€**
Energy loss connection NPV	1.350 k€	1.350 k€	1.350 k€
Energy loss intra farm NPV (full load at 1/3 of length)	225 k€	450 k€	675 k€
Total OPEX NPV per farm	1.575 k€	1.800 k€	2.025 k€
OPEX NPV per Turbine (b)	**158 k€**	**90 k€**	**8 k€**
Total costs per turbine (a + b)	**393 k€**	**287 k€**	**259 k€**
Total cost excluding transformer	**377 k€**		

It is clear that from the perspective of the needed distribution grid the higher voltages are preferred. The 20 kV solution performs about 20% better than the 10 kV solution and about 10% worse than the 30 kV solution. This holds even if the cost of the transformer is excluded for the 10 kV solution, given that the first amount can be connected without installing a transformer. The difference is also large enough for the conclusion to hold if the estimates for cable and equipment costs at higher voltages are more than 20% off.

4 Comparing Strategies

If the average costs per wind turbine are considered, it is clear the 30 kV solution performs best. However, when the first wind farm requested a connection, it may not have been clear that the number of turbines would grow that high. Building a 150/30 kV substation for wind turbines with a total power of just 10 MW is not a wise decision, especially if it is possible to connect about 55 MW to the existing station. Therefore the decision to be made was on the path to be followed, and not the technology as such. In Table 5, several strategies are drafted with their total costs of ownership. In the incremental strategy the first farms are connected to the existing transformer. Once the existing capacity is used, the substation is expanded by means of the cheapest investment, i.e. another 10 kV transformer. In the enhanced incremental strategies, a new technology is chosen when the limit of the existing installation is reached. In the visionary strategies, that new technology is implemented from the start.

As is clear from the this table, the "visionary plus" strategy performs best, whereas the "incremental strategy" performs worst. The actual situation is the result of the "enhanced incremental" strategy. It is perfectly understandable why this strategy was

Table 5. Total cost of ownership of several implementation strategies

NR	Strategy	Description	Total costs of ownership
1	Incremental	10 kV, first 55 MW on existing installation	55 * 377 k€ (existing) + 308 * 393 k€ (new) = 142 M€
2	Enhanced incremental	20 kV, first 55 MW on existing 10 kV installation (this is the actual situation)	55 * 377 k€ (existing) + 308 * 287 k€ (new) = 109 M€
3	Visionary	20 kV all new	363 * 287 k€ = 104 M€
4	Enhanced incremental plus	30 kV, first 55 MW on existing 10 kV installation	55 * 377 k€ (existing) + 308 * 259 k€ (new) = 100 M€
5	Visionary plus	30 kV all new	363 * 259 k€ = 94 M€

chosen. The first 55 MW required least investments from the DNO. And when it became clear the growth would be significant, the change was made to a technology already commonly used in the Netherlands, i.e. 20 kV. But in hindsight, this was not the best choice to be made. Switching to 30 kV at the decision point would have been better.

5 Conclusion

Sustainable energy production is required to decarbonize the energy system. Because of the distributed nature, connecting renewable sources to the grid is not a triviality. If a (almost regulatory enforced) incremental strategy is used, it is highly likely that resources will be wasted on inefficient grid designs. In this paper it was demonstrated that the difference between incremental and visionary strategies in a high growth scenario can be in the magnitude of 20–30%. This is significant by all means. Given that decarbonizing the energy system is a high growth scenario for electricity, asset management for the energy transition should commit itself to developing visionary strategic plans which allow for radical changes by means of new technologies in the asset base.

References

ABB AS Power Products Division: Product catalogue SafeRing 36 and SafePlus 36. 1VDD006114 GB (2015)

ACM: Tarievenbesluit Liander (2012)

ENEPEDIA (2012). http://enipedia.tudelft.nl/maps/PowerPlants.html

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland (2014)

Jonathan Fox (SP Energy Networks): DSO Transition. 4th Annual Asset Management Forum. TBM Group, Amsterdam (2017)

LIANDER: Kwaliteits en Capaciteitsdocument Elektriciteit (2011)

MacKay, D.J.C.: Sustainable Energy - Without the Hot Air. UIT Cambridge, Cambridge (2008)

The Windpower: Wind energy data for Netherlands (2012). http://www.thewindpower.net/country_windfarms_en_10_netherlands.php

TKF: Medium voltage power cables Catalog (2012)

Wijnia, Y.C.: Grid planning for renewable energy in urban deltas. In: Brown, K. (ed.) International Conference on Strategic Infrastructure Asset Management for Deltas 2013, Sydney, Australia (2013)

World Bank by the Potsdam Institute for Climate Impact Research and Climate Analysis: Turn Down the heat: Why a 4 °C Warmer World must be avoided, Washington D.C. (2012)

Robust Wind Farm Layout Optimization Under Weibull Distribution by Monte Carlo Simulation

Longyan Wang[1] and Andy C. C. Tan[2(✉)]

[1] Research Center of Fluid Machinery Engineering and Technology, Jiangsu University, Zhenjiang, Jiangsu, China
1000004851@ujs.edu.cn

[2] LKC Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Bandar Sungai Long, Kajang, Selangor, Malaysia
tanacc@utar.edu.my

Abstract. The wake shadowing problem is the main cause to the wind power exploitation losses for a wind farm. In order to boost the cost-competitiveness of wind energy compared to other renewable energy sources, it is imperative to reduce the wake losses for wind power exploitation in the wind farm. Among different approaches, the wind farm layout optimization is one of the most effective tools, which however, is very dependent on the condition of local wind resources for the optimization results. This paper aims to study the optimization of wind farm layout considering the uncertainty of wind conditions due to its unpredictability and randomness. The effect of uncertain wind condition on the wind farm optimization results is investigated by conducting the robust optimization of wind farm layout with the description of wind speed variation by Weibull distribution, while the evaluation of wind farm layout under the Weibull distribution is achieved by the Monte Carlo simulation. The robustness of the wind farm layout optimization results is evaluated by the wind farm expected total power output and the power variation for the sampled wind speed and wind direction obtained by Monte Carlo simulation. Through the robust optimization, it is found that under the Weibull distribution for a square-shape wind farm, the optimal wind farm layout is very sensitive to the incoming wind directions. It shows a very distinctive distribution of wind turbine placement even when only the probability of wind directions changes while the wind directions are fixed.

1 Introduction

The wake shadowing problem for developing a wind farm project has attracted great attention for the engineer and academia due to its adverse effect of reducing the power production and increasing the fatigue load on exploiting the wind energy [1]. In order to mitigate the effect of wake interaction for wind turbines, the optimization of wind farm layout plays a key role as one of the most effective tools to solve the problem [2, 3]. However, the wind farm layout optimization results are extremely vulnerable to the wind condition employed for the optimal design, which is hard to be accurately predicted in the real world [4]. Hence, the robust optimization results of wind farm layout

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 907–914, 2020.
https://doi.org/10.1007/978-3-030-48021-9_100

considering uncertain wind conditions are vital for wind farm developers to obtain the reliable optimal design outcomes.

The study of wind farm layout optimization problem began in 1994 by Mosetti et al. [5] to study a wind farm with the ideal square shape. Different types of wind conditions including the constant wind speed and wind direction, constant wind speed and variable wind direction, and the variable wind speed and wind direction, are investigated. The wind farm power output has an evident increase compared to the random wind farm layouts after optimization for all wind conditions. In order to provide a more realistic assessment of wind properties of a given region, more complex wind condition (speed-direction joint distribution of wind) is employed for the wind farm layout optimization study in reference [6]. The wind farm layout optimization under the similar joint distribution of wind speed and wind direction is also investigated in reference [7]. Nevertheless, the study of wind uncertainties has rarely been incorporated into the wind farm layout optimization in literature. Among those few study, the uncertainty of wind direction is evaluated by Gaumond et al. [8], and then applied for correcting the commonly used analytical wake models for wind farm optimization. In reference [9], the optimization model is accommodated to incorporate the uncertain parameters in wind characteristics and applied for the wind farm layout optimization.

Up to now, the robust optimization of wind farm layout considering the uncertainty of wind conditions (both wind speed and wind direction) has not been studied in literature, which will be the topic of interest for this paper. When studying the uncertainty of wind condition, the wind speed is described by Weibull distribution and its uncertainty study is carried out based on the Monte Carlo method, while two dominant wind directions with different probability index values are studied to conduct the wind direction uncertainty study for simplicity. The remainder of this paper is organized in the manner that Sect. 2 describes the modelling and methodology applied for the wind farm layout optimization study, Sect. 3 discusses the optimization results, and Sect. 4 concludes the research of this paper.

2 Modelling and Methodology

2.1 Wind Farm and Wind Turbine Models

The wind farm model studied in this paper is a square shape with 2 km length indicated in Fig. 1. This model has been extensively studied for the wind farm layout optimization as a benchmark model for the test of new optimization algorithms or new wind farm optimization methods in literature.

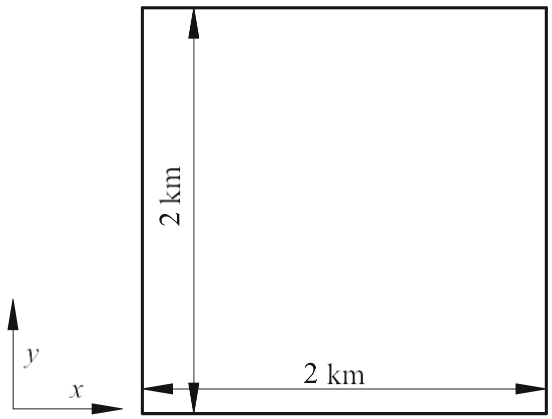


Fig. 1. Ideal wind farm with square shape with two-dimensional Cartesian coordinate system

The properties of wind turbine model used for the wind farm optimization study are extracted from literature given the consistency of study for the same wind farm. They are indicated in Table 1. In this paper, constant 38 wind turbines are studied for the wind farm optimization work.

Table 1. Properties of wind turbine model

Parameters	Values
Rated power	630 kW
Rotor diameter	40 m
Default hub height	60 m
Cut-in wind speed	2.3 m/s
Rated wind speed	12.8 m/s
Cut-out wind speed	18 m/s
Power coefficient	0.59

2.2 Weibull Distribution Model of Wind Speed

The wind scenario features can be typically represented with two variables: wind speed and wind direction. For the wind condition studied in this paper, the wind speed is described by the Weibull distribution which has been widely used to characterize the wind scenario all across the world. Its Probability Density Function (PDF) is represented by:

$$p(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1}\exp\left(-\left(\frac{v}{c}\right)^{k}\right) \tag{1}$$

where $p(v)$ represents the probability density of wind speed v, while c and k denote the scale parameter and the shape parameter, respectively. In this paper, the constant scale parameter equals 9 and the constant shape parameter equals 2 for the Weibull distribution. Two predominant wind directions from 0° and 90° hit the wind turbines, and their probabilities are set to be 20% and 80% to study their impact on the optimization results.

2.3 Objective Function Representation

The objects of the wind farm layout optimization are the total wind farm power output (P_{total}) and the variance of power output (Var) under Weibull distribution. Single objective genetic algorithm is adopted as the optimization algorithm and hence, the two optimization objects are incorporated into one objective function (f) connected with the weight parameter (α) as:

$$\text{Objective function: } f = \alpha \times P_{total} + (1 - \alpha) \times Var$$

$$s.t.: \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \geq d_{min} \tag{2}$$

where, x_i and y_i are the wind turbine Cartesian coordinates while d_{min} is the minimal allowed distance between any two wind turbines. As can be seen, the optimization objective function is the variance of the power output when α equals 0 and it is the total power output when α equals 1. For the optimization, weight values from 0 to 1 with the increment of 0.2 are adopted to investigate the trade-off between the two objects under the Weibull distribution wind condition.

2.4 Monte Carlo Method for Objective Function Evaluation

In order to calculate the wind turbine power output under the continuous Weibull distribution, the classic way of discretizing the wind speed is dividing the wind speed into certain number of bins with the constant interval, and the power output is calculated using the average wind speed of different bins [10]. However, such method has distinctive disadvantages including complex discretization steps and equations which lead to a large computational burden for the optimization, and its dependency of accuracy on the selection of wind speed interval value. Hence, more advanced Monte Carlo simulation method is introduced in this paper for the calculation of the wind farm power output under Weibull distribution wind condition. The steps are described in detail as bellow:

Firstly, the random variables ($x_i, i = 1, 2, \ldots, S$) that are uniformly distributed between 0 and 1 are generated using the random variable generator in the Matlab software. In this paper, the sample number is set to be 10000.

According to the Cumulated Density Function (CDF) of Weibull distribution, i.e., $F(v; k, c) = 1 - e^{-\left(\frac{v}{c}\right)^k}$, transform the samples of [0, 1] uniform variables (x_i) into samples of random variables (X_i) that follows the CFD by the inverse transformation $X_i = F^{-1}(x_i)$.

Based on the transformed random variables which are the discrete wind speed data for the study in this paper, the j-th wind turbine power and power variance can be calculated based on the equations of wind turbine power curve and the discrete data variance, that are:

$$P_i(j) = \begin{cases} 0 & \textit{if } v_i < 2.3\,\text{m/s } \textit{or } v_i > 18\,\text{m/s} \\ 0.3v_i^3 & \textit{if } 2.3\,\text{m/s} \leq v_i < 12.8\,\text{m/s} \\ 630 & \textit{if } 12.8\,\text{m/s} \leq v_i \leq 18\,\text{m/s} \end{cases}$$

and $Var(j) = \frac{1}{S-1}\sum_{i=1}^{S}\left(P_i(j) - \frac{\sum_{i=1}^{S}P_i(j)}{S}\right)^2$. The total wind farm power output and total power variance is the accumulation of all individual wind turbine power output and power variance as: $P_{total} = \sum_{j=1}^{N}\sum_{i=1}^{S}P_i(j)$ and $Var = \sum_{j=1}^{N}Var(j)$.

3 Results and Discussion

Two different wind conditions are employed for the wind farm layout optimization study in this paper. Firstly, the results of 80% probability of 0° wind direction and 20% probability of 90° wind direction are presented. Figure 2 shows the trade-off of power output variance and total power output of the wind farm and the wind farm efficiency (calculated as the ratio of real wind farm power output deducting power losses to the theoretic wind farm power output neglecting power losses) for different α values. As the α value increases from 0 to 1 (indicating the weight of power output optimization objective goes up from zero to one hundred percent, or the weight of power output variance goes down from one hundred percent to zero), the variance and total power output generally increase (see Fig. 2 (a)). At particular points (α equals 0.6, 0.8 and 1), the trade-offs almost maintain constant. When the weight value is small, the variance of power output exhibits higher weights than the total power output, and hence less variance of power output along with less power output are obtained after optimization

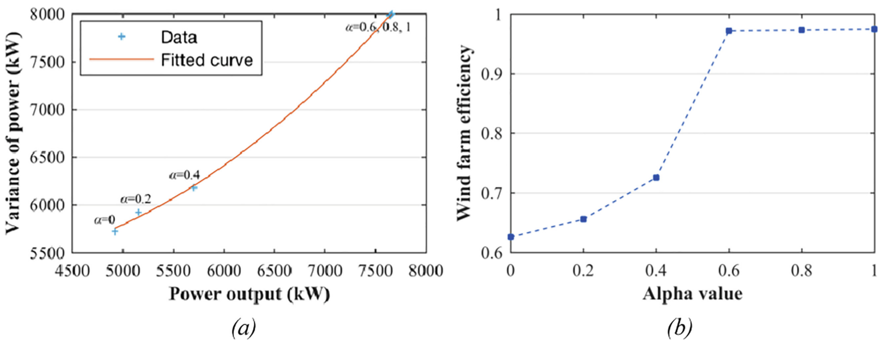


Fig. 2. Wind farm layout optimization results under the dominant wind direction of 0°: (a) variance of power and total power output comparison and (b) wind farm efficiency for different weight values of α

than that when the weight value is large. Less power output implies more wake losses, which can be seen in Fig. 2 (b). As the weight value increases, the wind farm efficiency also increases in general.

The optimal wind farm layouts after optimization with different weight values are presented in Fig. 3, and three particular values of α with 0, 0.4 and 1 are selected. When α equals 0, the objective function is the variance of power output. It has the largest wake losses, which can also be reflected from the optimal layout that the wind turbines are in line with the dominant wind direction and spreading all over the wind farm area. When α equals 1, the objective function is the total wind farm power output and it has the least wake losses. Hence, the wind turbines are staggered inside wind farm to avoid the wake interaction as much as possible. When α is in between, the objective function aims to achieve the trade-off between variance and total power output. Hence, the wind turbine positions are characterized with the properties of both aligned and staggered distributions.

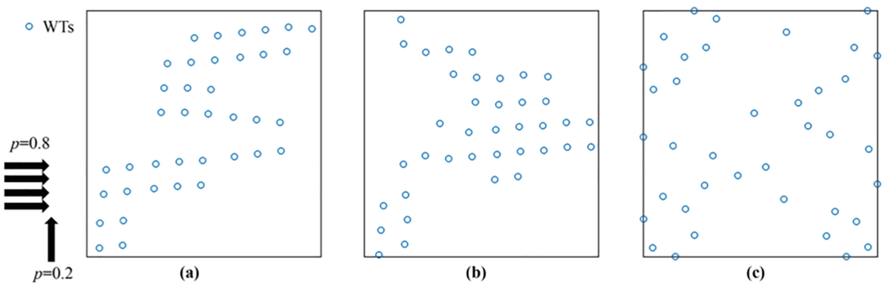


Fig. 3. Optimal wind farm layouts under the dominant wind direction of 0° for different weight values of the objective function: (a) $\alpha = 0$, (b) $\alpha = 0.4$ and (c) $\alpha = 1$

Next, the results of 80% probability of 90° wind direction and 20% probability of 0° wind direction are presented. As can be seen in Fig. 4 (a), similar to the above wind condition, both the variance of power and total power output increase as the weight value increases. At the points of $\alpha = 0$ and $\alpha = 0.2$, the trade-offs are close to each other. Almost unchanged trade-offs are obtained for the weight values of 0.6, 0.8 and 1. Close trade-off of variance and power output indicates approximately the same wake losses, which can be seen in Fig. 4 (b). In general, the wind farm efficiency increases when bigger weight value is employed. Then the optimal wind farm layouts with the weight values of 0, 0.4 and 1 are presented in Fig. 5. When the objective function is equivalent to the variance of the power output, the distribution of wind turbines is well organized in line with the dominant wind direction. And they are concentrated inside one portion of the wind farm area with the safety distance apart from each other. When the objective function is equivalent to the total wind farm power output, the wind turbines are staggered across the wind farm to achieve the least wake interaction.

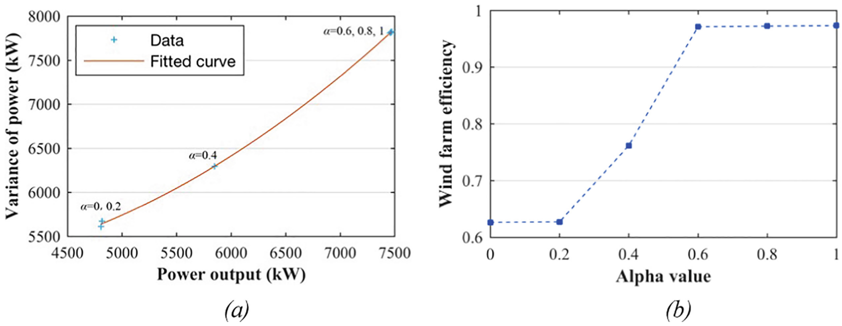


Fig. 4. Wind farm layout optimization results under the dominant wind direction of 90°: (a) variance of power and total power output comparison for different weight values, (b) wind farm efficiency for different weight values

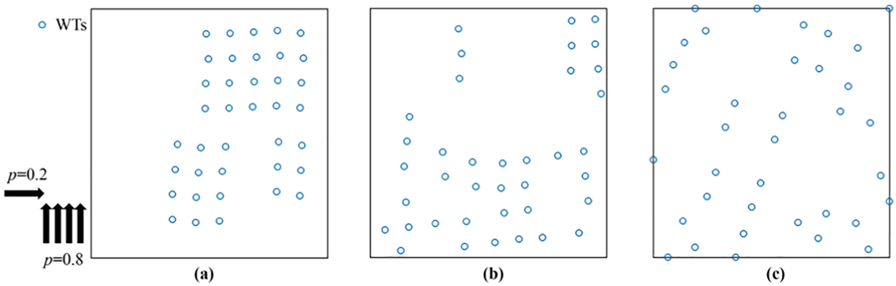


Fig. 5. Optimal wind farm layouts under the dominant wind direction of 90° for different weight values of the objective function: (a) α = 0, (b) α = 0.4 and (c) α = 1

4 Conclusion

The robust optimization of wind farm layout considering the uncertainty of wind conditions is conducted in this paper. Uncertainty of wind speed (represented by Weibull distribution) is investigated by the Monte Carlo simulation with randomly generated wind speed data. Uncertainty of wind direction is investigated for two constant dominant wind directions which have variable probability of occurrence. Trade-offs between the variance of power and total power output are obtained for different weight values of the objective function. It is found that as the weight value increases from 0 to 1 (implying the optimization objective changes from variance of power to total power output), both the variance and the total power output increase generally, along with the increase of wind farm efficiency. When the objective function is equivalent to the variance of the power output, the distribution of wind turbines is well organized in line with the dominant wind direction. When the objective function is equivalent to the total wind farm power output, the wind turbines are staggered across the wind farm to achieve the least wake interaction. By comparing the results under two different wind direction conditions, it shows a very distinctive distribution of wind turbine placements even though only the probability of wind directions changes while the wind directions are fixed.

References

1. Christiansen, M.B., Hasager, C.B.: Wake effects of large offshore wind farms identified from satellite SAR. Remote Sens. Environ. **98**(2–3), 251–268 (2005)
2. Kuo, J.Y.J.. Wong, I.A., Romero, D.A., Beck, J.C., Amon, C.H.: Wind farm layout optimization in complex terrains using computational fluid dynamics (2015)
3. Rašuo, B., Bengin, A.: Optimization of wind farm layout. FME Trans. **38**, 107–114 (2010)
4. Wang, L., Cholette, M.E., Zhou, Y., Yuan, J., Tan, A.C.C., Gu, Y.: Effectiveness of optimized control strategy and different hub height turbines on a real wind farm optimization. Renew. Energy **126**, 819–829 (2018)
5. Mosetti, G., Poloni, C., Diviacco, D.: Optimization of wind turbine positioning in large wind farms by means of a Genetic algorithm. J. Wind Eng. Ind. Aerodyn. **51**(51), 105–116 (1994)
6. Rahbari, O., Vafaeipour, M., Fazelpour, F., Feidt, M., Rosen, M.A.: Towards realistic designs of wind farm layouts: Application of a novel placement selector approach. Energy Convers. Manag. **81**, 242–254 (2014)
7. Feng, J., Shen, W.Z.: Modelling wind for wind farm layout optimization using joint distribution of wind speed and wind direction. Energies **8**(4), 3075–3092 (2015)
8. Gaumond, M., Réthoré, P.-E., Ott, A.P.S., Bechmann, A., Hansen, K.S.: Evaluation of the wind direction uncertainty and its impact on wake modeling at the Horns Rev offshore wind farm. Wind Energy **17**, 1169–1178 (2014)
9. MirHassani, S.A., Yarahmadi, A.: Wind farm layout optimization under uncertainty. Renew. Energy **107**, 288–297 (2017)
10. Kusiak, A., Song, Z.: Design of wind farm layout for maximum wind energy capture. Renew. Energy **35**(3), 685–694 (2010)

Science and Technology Parks as Innovation Intermediaries for Green Innovation

Roberto Rivas Hermann[1], Elisa Thomas[2](✉), and Mario Pansera[3]

[1] Nord University Business School, Nord University, Bodø, Norway
roberto.r.hermann@nord.no

[2] Center for Innovation Research, University of Stavanger Business School, Stavanger, Norway
elisa.thomas@uis.no

[3] Departament d'Empresa, Universitat Autònoma de Barcelona, Bellaterra, Catalunya, Spain
mario.pansera@uab.cat

Abstract. This paper discusses how science and technology parks (STPs) act as intermediaries for projects regarding green innovation. The empirical evidence is gathered through a case study of the City of Knowledge in Panama. For the recent Panama channel's expansion, local authorities faced the need to improve the water resource management to secure enough fresh water for the canal's operation. We inductively analysed data from 24 interviews, documents and participant observer. Preliminary results show the intermediation of STPs in green innovation processes in three phases: a first intermediation process is the STP as a hub for knowledge generation, including training for entrepreneurship. A second stage of the park as an innovation intermediary regards to an arena for knowledge and technology transfer, including collaboration with universities. A third phase implies financing and brokerage of green innovation between local and global actors. Our results challenge the existing literature about STPs with a narrow focus on economic spillover effects, or as hubs for attracting and developing cutting-edge technological innovations.

1 Introduction

Literature about science and technology parks (hereafter just STP) has grown exponentially in recent years, becoming a central topic within innovation management, industrial policy and science and technology studies (Hobbs et al. 2017). Research about STP in sustainability issues, however, remains fragmented. Some scholars focus on pollution control and end-of-pipe technologies, especially in southeast Asia (Wu et al. 2006). Other researchers seek to apply cleaner production principles to the operation of science parks (Chen et al. 2015) or the inclusion of sustainability goals in strategic planning (Ribeiro et al. 2016). However, there is a lack of studies about STP' role in supporting the generation of green innovation, contrasting to the vast literature examining support to start-ups and knowledge spillovers (Todo et al. 2011), or research on regional growth and science parks (Zhou 2005; Zhu and Tann 2005). As actors that join several other organizations, STP act as intermediaries (Löfsten and Lindelöf 2002)

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 915–922, 2020.
https://doi.org/10.1007/978-3-030-48021-9_101

providing different sets of services towards helping collaborative innovation projects (Thomas et al. 2017). However, more theorizing is necessary to address the gap of knowledge on the role of science parks in green innovation as intermediaries for projects in their hosting regions. Therefore, in this article we address the question: *how are science and technology parks acting as intermediaries for the generation of green innovation?*

To help answering this question, we research one case study, City of Knowledge, a STP located in the proximity of the Panama Canal (Dettenhofer and Hampl 2009). Panama authorities have recently faced the challenging situation of having to balance the construction of the channel's extension for commercial purposes with the needs of the country's citizens, and the social and environmental problems this activity causes (Floris 2012; Carse and Lewis 2017). The science park, along with other actors in the Panamanian innovation system, is a key actor in achieving this goal.

Our paper contributes to two streams of literature. First, we aim to fill the gap of knowledge about the role of STP in fostering and supporting the generation of green innovation. Second, we add to innovation intermediaries' literature by theorizing about intermediaries' roles in inter-organizational collaboration for green innovation. We structure the paper as follows: Sect. 2 presents a literature review focused on the science parks as innovation intermediaries. Section 3 presents the materials and methods. Section 4 presents findings from the case study and the discussion of results. Due to limitations of space for the paper, we opted to present them together. Finally, Sect. 5 presents preliminary conclusions taking into account that this study still seeks further discussion.

2 Literature Review

2.1 Science and Technology Parks as Places in Which Innovation Is Created

Science and technology parks (STP) are organizations with the main aim of promoting the culture of innovation and competitiveness for firms, universities and R&D institutions. This can be delivered for the associated organizations via facilities and high value-added services (IASP 2018). STP can also include incubators and accelerators to facilitate the creation and growth of new technology-based companies. Ratinho and Henriques (2010) argue that the two most important characteristics for the success of STP are university links and suitability of management. Löfsten and Lindelöf (2002) also state that "the assessing of academic knowledge and expertise by businesses located on-site is a key principle of Science Parks". By creating supportive spaces for knowledge and technology-based firms, STP may also facilitate technology transfer and help companies' growth (Guadix et al., 2016). This way, STP may become important drivers of regional development (Löfsten and Lindelöf 2002).

2.2 Science Parks as Intermediaries of Inter-organizational Relationships and Governance of Networks

The fact that actors geographically located close to each other interact more easily helps to explain the role of STP in the support for the generation of innovation through inter-organizational collaboration (Villani et al. 2017). As seen, one of the main characteristics of STP is the connection of companies located in the facilities with universities and research centres for knowledge sharing aiming at the creation of innovation. Networks provided by science parks benefit also new firms based on technology (Löfsten and Lindelöf 2002). However, Ratinho and Henriques (2010) found that the mere proximity of STP with universities is not enough for their success. Besides internal partners, Guadix et al. (2016) say that STP also aim to cooperate and promote cooperation with other actors in the public and private sectors, and these external collaborations have a positive effect on the outcomes of companies. Therefore STP act as innovation intermediaries by facilitating the identification of external knowledge sources and by making external knowledge accessible (Agogué et al. 2017).

As such, STP are considered innovation intermediaries or brokers, which are broadly defined as organisations that provide services and support role for collaboration between two or more parties during different stages of the innovation process (Howells 2006). Intermediaries are central to creating and maintaining a successful innovation ecosystem (De Silva et al. 2018). Park (2016) summarizes the roles of innovation intermediaries in facilitation, mediation, or coordination. According to Villani et al. (2017, p. 87), intermediaries can purposefully influence proximity among actors through specific direct and indirect activities, taking into consideration that “proximities are strongly related to context-specific characteristics, such as the complexity of the knowledge being exchanged and the type of actor involved in the technology transfer.” Besides promoting proximity among companies, intermediaries provide a broad set of services to innovative firms and collaborative projects, such as coordinating knowledge between actors through to commercializing new technologies (De Silva et al. 2018).

Regarding green innovation, “intermediaries can be critical to the exploration of new opportunities and the development of new ways to address shared issues, such as sustainability and environmental issues” (Agogué et al. 2017, p. 20). According to Ekins (2010), eco-innovation have complex political, institutional and cultural, in addition to technological and economic, dimensions. Coordination capability, which is the involvement of individuals and other firm resources across a company in regard to creating value for customers and other stakeholders, is positively related to green innovation (Huang and Li 2017). Green innovation or eco-innovation, in this paper, is defined as innovation that results in a reduction of environmental impact (OECD 2009).

3 Method

This research follows a qualitative approach through the case study technique.

3.1 The Case Study: Watershed Management for the Panama Canal Extension

The case study is the science and technology park City of Knowledge (COK) involved in the Panama Canal extension. COK is at the same time a "Science, business and technology park" at the former USA military base of Fort Clayton, besides the Panama Canal; and, a "Knowledge Management Network and Node" (COK 2017). Its origins date back to 1977 when Panamanian president Omar Torrijos and the American president Jimmy Carter signed an agreement by which the U.S.A. government compromised to give back to the Panama government the full sovereignty over the Panama Canal Zone, including the land, water and existing infrastructure by 1999. The main purpose of COK is to encourage innovation, and to foster the establishment of research centres, and knowledge transfer organizations.

In 2007, the Panamanian government started a 6 billion USD project to expand the Panama Canal. The engineering activities required widening the canal's entrances, deepening the canal, and building the infrastructure of the new locks (Spengler et al. 2014). The Panama Canal Authority (PCA) was the manager of the project. The success of such a megaproject in the long term depended on a secure supply of fresh water (Newbery 2017). In parallel, Panama authorities are in the challenging situation of having to balance the commercial activity of the channel with the needs of its own citizens, and the social and environmental problems this activity necessary causes (Floris 2012; Carse and Lewis 2017). The science park, along with other actors in the Panamanian innovation system are key actors in achieving this goal.

3.2 Data Collection and Analysis

A total of 24 interviews were carried during three phases of data collection (2010–2016). The interviews targeted key informants at the higher hierarchy of organizations dealing with one way or another involved in transition processes linked to water resource management. Thus, advisors (6), consultants (2), directors (1), managers (8), professors (1), rector (1), officers (1) and researchers (3). These interviews were carried in the units of analysis: City of Knowledge and PCA, supportive interviews were also conducted. In total 5 interviews involved staff from PCA, NGOs (4), companies (5), research centres (1), universities (5), cooperation agencies (1) and government agencies (3). The semi-structured interview guides started with a broad discussion about previous or existing collaboration initiatives between the science park and the canal authority with focus on water resource management. When specific projects where highlighted more detailed questions were addressed in regards to the type of innovation developed along the project, and specific aspects of the innovation - resources, key actors and stages. Interview transcripts and field notes were coded using the software QSR NVivo. The analysis of data was performed according to the content analysis technique where the authors used previous literature to analyse empirical findings to generate implications to practitioners and to the literature.

4 Findings

Due to limitations of space, we opted to present findings from the case study organized according to the literature regarding the roles of innovation intermediaries.

4.1 Knowledge and Technology Transfer

Firms in the park related to green innovation and connected to the case of the canal's extension include consulting agencies, maritime R&D service companies, and clean tech companies selling technology. Environmental related activities of these organization range from environmental studies (like consulting companies which can prepare environmental impact assessments), capacity building (training on specific topics like ISO norms), sales of "clean technology" (renewable energies or waste management), or technology related to the maritime sector.

The case portrayed the value of STP to develop educational programs consultants, which become reliable for particular insights and capacity building. This finding is not a fortuity, in the literature, consultants' importance in the transformation of knowledge and innovation is often stressed (Swan et al. 2003). Among non-for-profit organizations, there are research centres and academic institutions. From an environmental point of view, research centres' work relates to policy, innovation, knowledge management and environmental studies. Academia comprises universities, NGOs and government bodies. Similarly, innovation promotion government agencies are also located in City of Knowledge. These agencies fund prizes for start-ups and finance settlement of human capital (national or foreigner researchers who propose research agendas). This examples correspond to Guadix et al. (2016) regarding the aim of STP to cooperate and promote cooperation with external actors in the public and private sectors.

4.2 Knowledge Generation

Different from knowledge and technology transfers, interactions between City of Knowledge and ACP's environmental strategies aim at generating knowledge through the provision of education, training, and consultancy. Educational and training activities evidence interactions between industry, government and universities promoted by the park. In a first example, one private firm established at City of knowledge, created agreements with a major national university in order to launch a master program in Environmental management information systems. The government was involved as it provided funding for the program.

Another example of knowledge generation highly involving City of Knowledge is in the fields of medicine and biotechnology. The park has been attracting international firms which are gradually creating a cluster in these domains. Foreign Direct Investments may fund R&D in this case. Local universities train human capital doing research in these two areas. The knowledge generation, in the case of the cluster, happens around tacit and explicit types of knowledge through formal and informal means (Thomas 2018). In this case, the STP facilitates the identification of external knowledge sources and makes external knowledge accessible to firms, acting as an intermediary (Agogué et al. 2017).

Besides educational activities, consulting offers opportunities for interaction between organizations in City of Knowledge and ACP environmental division. For example, CATHALAC[1] has developed close links with ACP. From the one side, it has exploited its integrated water resource management experience by training ACP hired educators working on environmental education in the watershed: “We are about to start a project to integrate climate change adaptation to sustainable development plans” (interview CATHALAC). Similarly, other City of Knowledge based organizations have provided consulting to ACP’s PCW protection programs (i.e. Ramsar, and the NGO Panama Verde).

4.3 Financing Innovation

One interviewee, the head of the entrepreneurial innovation division from the National Innovation Agency, stated that it is not difficult to get funding in Panama; “the problem is not the money. The problem is to have good and competitive projects presented [to calls managed by national or international cooperation agencies]”. In Panama, government agencies provide direct funding for innovation. At lesser extent, other organizations including the City of Knowledge indirectly provide funding to innovation by administering external funds that come from international cooperation agencies. The access to financial resources as a fundamental resource for a collaborative R&D project is regarded as one of the roles of innovation intermediaries (Thomas et al. 2017). At the City of Knowledge, a private equity group called ECOS S.A. finances projects dealing with renewable energy, community tourism, biofuels and others. Yet, its operations are regional (e.g. Latin–America, with most projects located in Colombia and Brazil).

5 Conclusions

This paper aimed to analyse science and technology parks acting as intermediaries for the generation of green innovation. We presented the case study of City of Knowledge, a park located near the Panama Canal and involved in the project for the watershed management on the constructions for the canal’s extension.

From our data, we can see that City of Knowledge Foundation’s interest to promote the science park as a hub to develop knowledge on natural resources management, such knowledge emphasizes new modes of institutional innovation. Hence, the case illustrates a particular type of “green innovation” focused on new institutional modes of natural resource management. This has consequences as ICT and bio-tech organizations seem to quantitatively dominate over environmental services providers. Thus, it gives an impression that no specialization can be expected in City of Knowledge with regard to green technology or services. Even though, our analysis showed the intermediation of the park in green innovation processes. These activities can be organized in three ways: a first intermediation process is the STP as a hub for knowledge

[1] Water Center for the Humid Tropics of Latin America and the Caribbean (*Centro del Agua del Trópico Húmedo para América Latina y el Caribe*).

generation, including training for entrepreneurship. A second stage of the park as an innovation intermediary regards to an arena for knowledge and technology transfer, including collaboration with universities. A third phase implies financing and brokerage of green innovation between local and global actors. Our results add to the existing literature about STPs which present mostly a narrow focus on economic spillover effects, or as hubs for attracting and developing cutting-edge technological innovations.

Nevertheless, the proximity of the STP to the canal has hitherto not yielded with the creation of a “green cluster”, which could be a precedent to better promote green innovations. Our findings suggest that interactions of the science park with the Panama Canal Authority and other actors in the region are not institutionalized but take place through adhoc projects. Therefore, there is opportunity for the STP to establish itself clearly as an intermediary and service provider for collaborative projects between industry, universities and governments. Although this paper generates insight for further discussion on the role of science parks as innovation intermediaries for green innovation, we acknowledge some limitations to its results. As a single case study, it does not allow the validation of results to a wider population of STP. Also, the Panama Canal extension present some unique attributes of its case that may not be found in other projects regarding green innovation.

References

Agogué, M., Berthet, E., Fredberg, T., Le Masson, P., Segrestin, B., Stoetzel, M., Ystrom, A.: Explicating the role of innovation intermediaries in the “unknown”: a contingency approach. J. Strategy Manag. **10**(1), 19–39 (2017)

Carse, A., Lewis, J.A.: Toward a political ecology of infrastructure standards: or, how to think about ships, waterways, sediment, and communities together. Environ Plan A **49**, 9–28 (2017)

Chen, W.-Y., Chen, H.-W., Chang, C.-N., et al.: Particles and metallic elements near a high-tech industrial park: analysis of size distributions. AEROSOL AIR Qual. Res. **15**, 1787–1798 (2015)

CoK: The Foundation- Nonprofit organization that manages the City of Knowledge (2017). https://apps.ciudaddelsaber.org/portal/en/foundation. Accessed 19 Aug 2017

De Silva, M., Howells, J., Meyer, M.: Innovation intermediaries and collaboration: knowledge–based practices and internal value creation. Res. Policy **47**(1), 70–87 (2018)

Dettenhofer, M., Hampl, N.: Development of a biomedical innovation economy-Panama. J. Technol. Manag. Innov. **4**, 21–32 (2009)

Ekins, P.: Eco-innovation for environmental sustainability: concepts, progress and policies. Int. Econ. Econ. Policy **7**(2–3), 267–290 (2010)

Floris, V.: Water and environmental management in the expansion of the Panama Canal. World Environ. Water Resour. Congr. **2012**, 2084–2093 (2012)

Guadix, J., Carrillo-Castrillo, J., Onieva, L., Navascues, J.: Success variables in science and technology parks. J. Bus. Res. **69**(11), 4870–4875 (2016)

Hobbs, K.G., Link, A.N., Scott, J.T.: Science and technology parks: an annotated and analytical literature review. J. Technol. Transf. **42**, 957–976 (2017). https://doi.org/10.1007/s10961-016-9522-3

Howells, J.: Intermediation and the role of intermediaries in innovation. Res. Policy **35**(5), 715–728 (2006). https://doi.org/10.1016/j.respol.2006.03.005

Huang, J.-W., Li, Y.-H.: Green innovation and performance: the view of organizational capability and social reciprocity. J. Bus. Ethics **145**(2), 309–324 (2017)

IASP - International Association of Science Parks and Areas of Innovation, I. The role of STPs and areas of innovation. https://www.iasp.ws/Our-industry/The-role-of-STPs-and-areas-of-innovation. Accessed Apr 2018

Löfsten, H., Lindelöf, P.: Science Parks and the growth of new technology-based firms—academic-industry links, innovation and markets. Res. Policy **31**(6), 859–876 (2002)

Newbery, M.: The critical role of water management and reliability in the Panama Canal expansion. World Environ. Water Resour. Congr. **2017**, 589–599 (2017)

OECD - Organisation for Economic Co-operation Development. Eco-innovation in industry: enabling green growth. OECD Publishing, Paris (2009)

Park, J.H.: Brokerage activities in regional innovation networks: the case of Daegu Technopark in Korea. Int. J. Urban Sci. **20**(2), 260–284 (2016)

Ratinho, T., Henriques, E.: The role of science parks and business incubators in converging countries: evidence from Portugal. Technovation **30**(4), 278–290 (2010)

Ribeiro, L.M., Botelho, S.S.C., Duarte Filho, N.L.: Axis of Sustainability: a proposal for an eighth axis for the tool Strategigram | Eixo sustentável: Uma proposta de um oitavo eixo para a ferramenta estrategigrama. Revista Espacios 37 35 (9) (2016)

Spengler, R., Casanova, F., Imedio, J., Pérez, R.: Panama Canal expansion project - description third set of locks project. In: 37th IABSE Symposium Engineering for Progress, Nature and People. IABSE, Madrid, pp. 783–790 (2014)

Swan, J., Harry, S., Maxine, R.: Linking knowledge, networking and innovation processes: a conceptual model. In: Shavinina, L.V. (ed.) The International Handbook on Innovation, pp. 680–694. Pergamon, Amsterdam (2003)

Thomas, E., Marques Vieira, L., Balestrin, A.: Mind the gap: lessons from the UK to Brazil about the roles of TTOs throughout collaborative R&D projects. BAR - Brazilian Adm. Rev. **14**(4) (2017). https://doi.org/10.1590/1807-7692bar2017170048

Thomas, E.: From closed to open innovation in emerging economies: evidence from the chemical industry in Brazil. Technol. Innov. Manag. Rev. **8**(3) (2018)

Todo, Y., Zhang, W., Zhou, L.-A.: Intra-industry knowledge spillovers from foreign direct investment in research and development: evidence from China's 'Silicon Valley'. Rev. Dev. Econ. **15**(3), 569–585 (2011)

Villani, E., Rasmussen, E., Grimaldi, R.: How intermediary organizations facilitate university–industry technology transfer: a proximity approach. Technol. Forecasting Social Change **114**, 86–102 (2017)

Household Energy Consumption Prediction Using Evolutionary Ensemble Neural Network

Songpu Ai(✉), Antorweep Chakravorty, and Chunming Rong

University of Stavanger, 4036 Stavanger, Norway
{songpu.ai, antorweep.chakravorty, chunming.rong}@uis.no

Abstract. Low-voltage local electricity intelligent management is an essential portion of smart grid research. Thereinto, a precise prediction of domestic energy consumption is a pivot in establishing household/neighbourhood energy management system to achieve local smart solutions including consumption auto-balancing, micro generation & storage system, neighbourhood energy sharing, etc. Recent years, a large amount of literature has considered the use of artificial neural networks (ANNs) on electric load forecasting. Various ANN structures and configurations are employed and discussed. However, the solutions are generally developed as case by case studies. The advised network for each specific problem is commonly selected through empirical or enumerative approaches. In this article, an evolutionary ensemble approach is proposed to pool ANNs with various structures and configurations to forecast domestic energy consumption efficiently. The approach utilizes an evolutionary method to select and reproduce better performed network individuals in a network pool to optimize prediction quality. Forecast results demonstrate that the approach achieves a more accurate energy consumption prediction comparing with ANNs with commonly used configurations.

1 Introduction

As novel electrical characters for instance electric vehicle, micro-generator, household battery array, etc., are integrated into family daily life progressively. A structural transformation is undergoing among domestic electrical environment (Kotilainen and Saari 2018). In the perspective of power system, domestic users are transforming from a single consumer to a combined role gathering electricity generation, consumption and storage. This change in role makes the accurate prediction of domestic electricity consumption a necessity.

In future energy systems, either smart grid or energy internet involves low-voltage local intelligent energy management system (EMS) as an essential portion to realize intelligent and precise domestic load allocation and management (Javaid et al. 2013; Huang et al. 2017). Thereinto, a precise prediction of domestic energy consumption is a pivot in establishing household/neighbourhood EMS to decrease the potential peaks by shifting the operating time of controllable energy intensive non-critical loads from a potential energy demand peak to an off-peak period, in order to reduce the risk of damage caused by energy overload. Through further intelligent integrating with local generation, storage at household/neighbourhood level, as well as interacting with

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 923–931, 2020.
https://doi.org/10.1007/978-3-030-48021-9_102

higher level power grid, the household/neighbourhood EMS is able to achieve advanced local smart solutions including consumption auto-balancing, micro generation & storage system, neighbourhood energy sharing, etc.

Recent years, artificial neural network (ANN) as a promising approach on machine learning and data mining, has attracted the attention of research and industry communities on many fields (Kabir et al. 2018), which include energy informatics. Various ANN topologies and training methods are employed and discussed, such as convolutional neural networks, recurrent neural networks in topology and backpropagation, genetic algorithm for weight updating.

ANN is particularly appealing because of its ability to model the unspecified nonlinear relationship between variables. For instance, the relationship between hourly household energy consumption and the hour of a day, the day of a week, temperature, etc. is discussed in Songpu et al. (2015). However, the network structure and activation function are generally developed as case by case studies. The advised network structure and activation functions of the neurons for each specific problem is commonly chosen by empirical or enumerative approaches (Arifovic and Gençay 2001). At the same time, the initialization of the network and neurons is also an affair which may impact the achievement of efficient analysis. Normally the initial values of weights and bias in neurons are randomly selected, which may cause uncertainties on training time, local optimize, etc.

To overcome the realistic dilemmas discussed above, an evolutionary ensemble neural network (EENN) method is proposed in the article as the main contribution of this study. EENN is a novel type of ensemble learning which pools ANNs with various structures and configurations to obtain analysis result. The network pool is trained through an evolutionary approach. A more accurate hourly household energy consumption prediction is achieved using the proposed method.

The remainder of the article is organized as follows. Section 2 presents the EENN we proposed. Afterwards, the studied case including the description of utilized data and data processing methods is introduced in Sect. 3. In Sect. 4, the performance of the proposed method is illustrated and discussed. Finally, Sect. 5 provides the conclusion of the work.

2 Evolutionary Ensemble Neural Network

In brief, EENN utilizes ensemble method to attain analysis output and to use evolutionary algorithm to update the population of networks in the pool. Random configured ANNs are firstly initialized in a network pool. Then, along with the iterations of training, the individuals in the pool are selected to drop or reproduce (fitness process). Ultimately, the final analysis result is given by the survived networks (which are the most suitable ones for the problem) collaboratively.

2.1 Ensemble Learning

Ensemble learning is the art of integrating different machine learning models to achieve better performance (Zhou 2012). It is developed basing on the outcomes of multiple learning models. A commonly utilized ensemble approach for ANN is ensemble averaging. The results of a set of ANN are combined and averaged to attain the output. Meanwhile, special architectures are also utilized to fit diverse needs, which are called hybrid ensemble learning. In the work of Ai et al. (2018), a two-layer hybrid stacking ensemble learning is developed, in which multiple machine learning algorithms are combined through averaging. Time series data is forecasted in Wichard (2016) using hybrid ensemble learning with weighted mean of several component algorithms.

In this article, an innovative hybrid ensemble learning model is proposed, which can be seen as an integration of evolutionary algorithm and ensemble averaging. Various networks in a network pool are trained separately. The selection is done evolutionarily and the analysis results of networks in the pool are combined and averaged to get a joint output as showed in Eq. 1.

$$OUT = \frac{1}{N}\sum\nolimits_{i=1}^{N} O_i \tag{1}$$

Where N denotes the total amount of networks in the pool; O_i denotes the output of the i'th network.

2.2 Evolutionary Algorithm

Evolutionary is a cluster of machine learning methods which simulates the generic population optimization process happening in nature. Evolutionary methods are of the advantages like robust to dynamic changes, with capability for self-optimization, etc. (Fogel 1997). The methods are generally simple in structure and easy to be realized. Considering that the objective we have here is to achieve a better performed model on numerous potential configurations, it is interesting to utilize evolutionary method to discover a set of networks with good enough network structures and proper parameters configured at the same time.

Specifically, in this study a set of potential network structures with varying activation functions and weights are initialized in the network pool. The network pool here can be analogous to the pool living environment of fishes. Various types of fishes (networks) live in the same pool environment utilizing data as food. Networks fitting well to the environment can survive and reproduce in the pool and become the mainstream.

We denote a neural network in the pool as NN^i, where $i \in [1, N]$. N is the total amount of networks in the pool. Then, each network in the pool is trained *init* iterations before sent be selected. In selection, all networks are evaluated by an amount of evaluation data samples and the ones with the worst performance are dropped. Ones with better performance are reproduced and filled back to the pool to keep the population. The percentage of networks to drop is denoted as P_{drop}. After each selection, the evaluation data set is integrated into the training set to train all networks T iterations before the next round of selection. Through multiple rounds of training and selecting, a test set is sent to the pool and the all networks in the pool give their forecast result collaboratively using Eq. 1. The entire process of the method is shown as a work-flow diagram in Fig. 1.

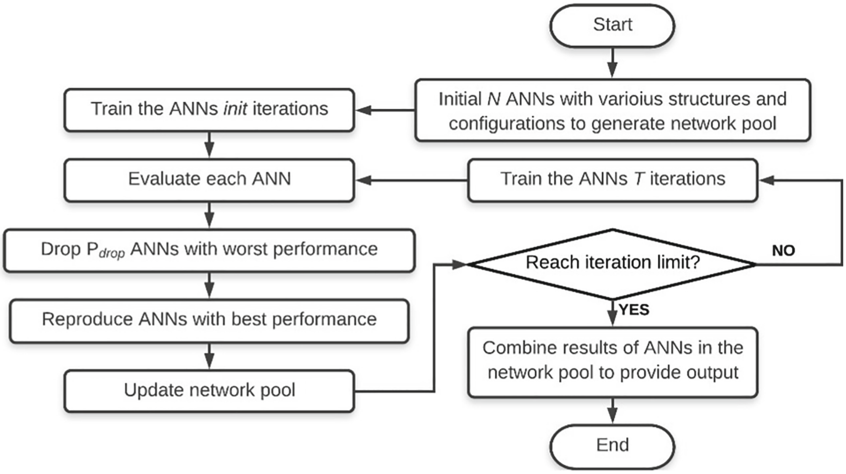


Fig. 1. Work-flow of EENN.

Comparing with genetic algorithms which are adopted in ANN optimization like Arifovic and Gençay (2001), Mahajan and Kaur (2014), etc. evolutionary algorithm in our study is utilized on the selection of better performed network structures and configurations among various network models, rather than simply selecting edges between neurons or adjusting weights of one network.

The word "pool" in this study also has the similar implying sense as the "pool" in the pooling layer in convolutional neural network, which combines the outputs of neurons at one layer as the output of the next layer. In the case study, we adopt the ensemble averaging to obtain the final output, which is analogous with the average pooling in convolutional networks.

In implementation, the required capability of computation decreases along with the progress of the program running, since the duplications of good networks are unnecessary to be computed more than once.

3 Case Study

In this study, a practical case is investigated to forecast hourly energy consumption using EENN. The sample situation is a typical household in western Norway. The study is programmed in Python 3.0 (Python 2018). TensorFlow (2018) and Scikit-learn (2018) libraries are utilized to implement ANN configuration, initialization, training and evaluation.

3.1 Data Description

The power usage of an anonymous western Norwegian domestic household energy consumption data is recorded approximately 4 months. The real-time power read is monitored every 10 s locally and then the data is sent to cloud and stored there. The raw data includes two features: timestamp and power read in Watt. For an intuitive understanding, the power usage of the household among 24 h is presented in Fig. 2.

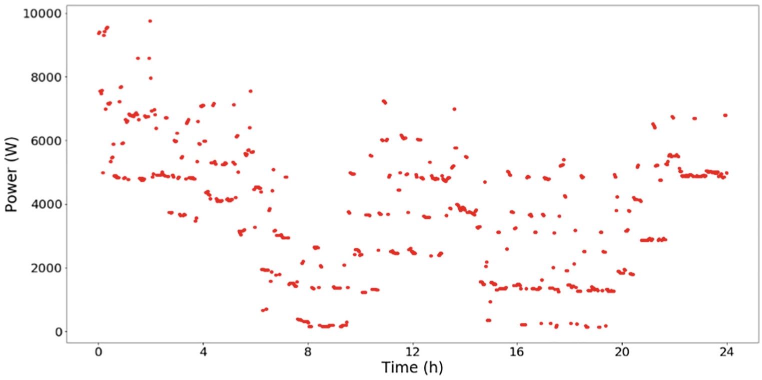


Fig. 2. Power usage data of the household among 24 h.

3.2 Data Processing

Data processing is a necessary step before feeding data into the proposed model. Any errors or incompletions in the raw data is needed to be addressed in a suitable way to prepare proper inputs for the prediction models. Through analysing the available raw dataset, it is observed that the interval between rows are not static. Although 99.0% of them are in range $[9.5, 10.5]$ (s), still a certain amount pair of rows hold longer or shorter intervals. In addition, there are dozens of row-pairs with interval less than 0.5 s. Through enumeration, all pairs with less than 5 s interval hold the same power reads among each pair. Hence it is reasonable to have an assumption that, for a pair of consequent rows, if the interval between them is less than 5 s then the latter row is a redundant data row. In addition, we assume that if the interval between rows is greater than or equal to 15 s, it is considered that one or multiple missing rows exist between these two rows, while the intervals which are greater than or equal to 5 s and less than 15 s are recognized as normal, as expressed in Eq. 2. Due to there are both certain amount of redundant and missing data in the raw dataset, it is supposed that a possible origin of data errors is from the communication process.

$$\begin{cases} 0 \le interval < 5, & Redundant\ Data \\ 5 \le interval < 15, & Correct\ Data \\ interval \ge 15, & Data\ Missed \end{cases} \tag{2}$$

The missed m rows in row-pairs with interval greater than or equal to 15 s are refilled by the m-equidistant values of power reads of the start- and end-row. As observation, the two reads of row-pairs with missing data are relatively stable generally. The differences are normally less than 10 W. Very few of them have larger differences. Therefore, the refilled data rows probably do not have greatly impact on energy consumption in hour level.

The hourly energy consumption is calculated through the integral with respect to time of the piecewise function of power usage. The input components of the ANNs are chosen based on the input combination of our early study on household hourly demand prediction (Songpu et al. 2015), with certain enumerative tuning for the data utilized in this article. The input components consist of eight energy components: energy consumption of five latest hours, energy consumption of the same hour of one, two and seven days ago (with 24-/48-/168-h lag respectively), as well as three components of time: the hour of the day, the day of week, and workday/holiday. All components are normalized to avoid biases.

The purpose of implementing this case study in the article is to evaluate the performance of the EENN contrasting with commonly used ANNs. This article does not discuss whether the input combination is the optimal one for energy consumption prediction. Yet it is important to provide the same data for EENN and ANNs with commonly used structures to compare the performance of the predictions.

4 Model Performance

The performance of EENN is investigated in this section. The model configuration and evaluation criterion are bestowed firstly. Then the model performance is presented and compared with the results of ANNs adopting commonly used structures and configurations.

4.1 Model Configuration and Evaluation Criterion

The total number of ANNs in the network pool is configured as $N = 10$. For each ANN, the number of hidden layer, L_{NN^i}, can be randomly chosen between one to four layers, i.e. $L_{\mathrm{NN}^i} \in [1, 4]$. The number of neurons in each hidden layer, $n_{L_{\mathrm{NN}^i}}$, is set within the range between ten and a hundred, i.e. $n_{L_{\mathrm{NN}^i}} \in [10, 100]$. Each neuron has two options on activation function, either sigmoid function or hyperbolic tangent (tanh) function. Mean squared error is allocated as the cost function and the learning rate is fixed as 0.005. It is because the difference on learning speed may lead to faulty eliminating networks with better structures and configurations indeed. The weights and bias of each neuron are set up randomly.

All the available data samples are divided into three sets, 70% of data samples are utilized as training set, 15% for evaluation and the remaining 15% are utilized on testing. The training steps are configured as $init = 20$, $T = 10$, $P_{drop} = 10\%$, and the iteration limit is set as 100.

To evaluate the performance of prediction models, mean absolutely percentage error (MAPE) is adopted in this study. It measures the performance of a method by calculating the average absolute of occurred errors, as showed in Eq. 3.

$$MAPE = \frac{1}{M}\sum\nolimits_{j=1}^{M}\left|\frac{OUT_j - TGT_j}{TGT_j}\right| \cdot 100\%. \quad (3)$$

Where M denotes the total number rows in the test set.

4.2 Model Performance

The prediction evaluations of EENN and ANNs with several commonly utilized configurations are shown in Table 1. The ones with better performance are marked as bold. It can be observed that among ANNs with commonly utilized network & neuron configurations, the one with four 10-neuron hidden layers and the one with two 50-neuron hidden layers attain the best performance among sigmoid and tanh activation function networks respectively. If an enumerative approach is used to develop the prediction model, these two networks are promising to be chosen. However, the evaluation results elucidate that our proposed method accomplishes even better performance, which obtains 0.0199 on MAPE. The progress of utilizing the EENN achieves approximately 45%.

Table 1. Prediction results of ANNs with commonly used network/neuron configurations and EENN.

Num. of hidden layer	Activation function	Sigmoid			Tanh			EENN
	Num. of neuron each layer	10	50	100	10	50	100	
1		0.0645	0.339	0.273	0.288	0.479	0.826	**0.0199**
2		0.156	0.194	0.0879	0.171	**0.0353**	1.04	
3		0.144	0.274	0.194	0.481	0.991	0.538	
4		**0.0390**	0.159	0.0537	0.349	0.130	0.481	

The relationships between evaluation and network iteration time of the top two ANNs and the proposed method are presented in Fig. 3. It is observed that at the beginning of the selection stage (since $init = 20$), the performance of the proposed method is not as good as the ANNs. The reason could be that the two ANNs acquire relatively "lucky" structures and weight-sets during their network initialization, which means that their structures and weight-sets are originally more fit the problem than others. However, the progress rate of EENN is higher than the ANNs. The performances of the EENN and the ANNs become similar after 90–100 time of iterations for the evaluation set. Then, EENN achieves a better performance on the test set.

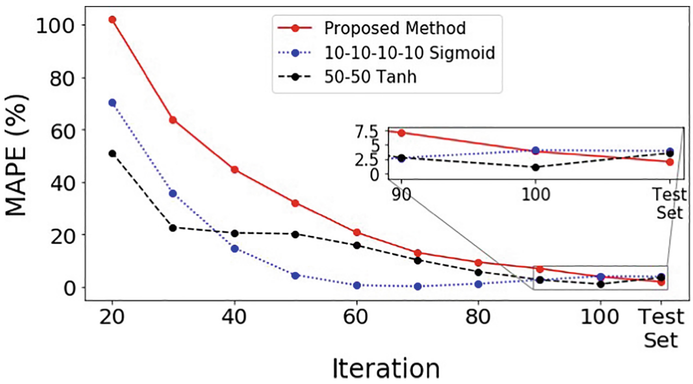


Fig. 3. The relationship between iteration time and MAPE of the proposed method and the top two ANNs.

When looking into the details of the later portion of the iterations, it can be aware of certain overfitting issues on the ANNs. The performance of the "sigmoid" ANN (the blue dot line) gets certain negative growth along with the join of evaluation data items into training set. The evaluation of the "tanh" ANN (the black dash line) grows worse when forecasting the test set. However, EENN improves the performance while the evolutionary process (even at the final time of population update at iteration 100) and becomes better than the ANNs. Meanwhile, the reason that EENN achieves better performance could also get the contribution from the utilization of ensemble averaging which neutralizes errors among the network pool.

5 Conclusion

In this article, a novel type of ensemble learning, EENN, is proposed, which pools ANNs with various structures and configurations. An evolutionary approach is developed to train the network pool. Case study demonstrates that a more accurate household energy consumption hourly prediction is achieved using the proposed method.

References

Ai, S., Chakravorty, A., Rong, C.: Household EV charging demand prediction using machine and ensemble learning. In: 2018 IEEE International Conference on Energy Internet (ICEI), 21–25 May 2018, pp. 163–168 (2018)

Arifovic, J., Gençay, R.: Using genetic algorithms to select architecture of a feedforward artificial neural network. Phys. A Stat. Mech. Appl. **289**, 574–594 (2001)

Fogel, B.: The advantages of evolutionary computation. In: Biocomputing and Emergent Computation: Proceedings of BCEC 1997. World Scientific Press (1997)

Huang, C.J., Jhang, J.Y., Hsiao, J.T., Hu, K.W., Chu, C.C.: A renewables-based load-balancing energy management system for energy internet. In: 2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN), 6–8 May 2017, pp. 217–222 (2017)

Javaid, N., Khan, I., Ullah, M.N., Mahmood, A., Farooq, M.U.: A survey of home energy management systems in future smart grid communications. In: 2013 Eighth International Conference on Broadband and Wireless Computing, Communication and Applications, 28–30 October 2013, pp. 459–464 (2013)

Kabir, H.M.D., Khosravi, A., Hosen, M.A., Nahavandi, S.: Neural network-based uncertainty quantification: a survey of methodologies and applications. IEEE Access **6**, 36218–36234 (2018)

Kotilainen, K., Saari, U.: Policy influence on consumers' evolution into prosumers-empirical findings from an exploratory survey in Europe. Sustainability **10**, 186 (2018)

Mahajan, R., Kaur, G.: Neural networks using genetic algorithms. Int. J. Comput. Appl. **77**, 6–11 (2014)

Python (2018). https://www.python.org

Scikit-Learn (2018). http://scikitlearn.org

Songpu, A., Kolhe, M.L., Jiao, L., Zhang, Q.: Domestic load forecasting using neural network and its use for missing data analysis. In: 2015 9th International Symposium on Advanced Topics in Electrical Engineering (ATEE), 7–9 May 2015, pp. 535–538 (2015)

TensorFlow (2018). http://www.tensorflow.org

Wichard, J.D.: An adaptive forecasting strategy with hybrid ensemble models. In: 2016 International Joint Conference on Neural Networks (IJCNN), 24–29 July 2016, pp. 1495–1498 (2016)

Zhou, Z.H.: Ensemble Methods: Foundations and Algorithms. Taylor & Francis, Boca Raton (2012)

Overview of the US Department of Energy Light Water Reactor Sustainability Program

Bruce P. Hallbert(✉)

Idaho National Laboratory, P.O. Box 1625, Idaho Falls, ID 83415, USA
bruce.hallbert@inl.gov

Abstract. The Department of Energy's Light Water Reactor Sustainability (LWRS Program) is being carried out to enhance the safe, efficient, and economical performance of US nuclear power plants and extend the operating lifetimes of this reliable source of electricity. It has two facets with respect to long-term operations: (1) manage the aging of plant systems, structures, and components (SSCs) so that nuclear power plant lifetimes can be extended and the plants can continue to operate safely, efficiently, and economically; and (2) provide science-based solutions to the industry to implement technology to exceed the performance of the current labor-intensive business model.

The Department of Energy's (DOE's) role is to partner with industry and with the NRC and key industry support groups to conduct the research needed to inform major component refurbishment and replacement strategies, performance enhancements, plant license extensions, and age-related regulatory oversight decisions. When appropriate, research and development (R&D) and demonstration activities are cost-shared with industry or the NRC. Pilot projects and collaborative activities are underway at commercial nuclear facilities and with industry organizations.

1 Background

Nuclear energy is an important contributor to meeting national electricity generation objectives. It provides reliable, base-load capacity at historically high availability rates while supporting national greenhouse gas emission goals. The United States (U.S.) commercial nuclear power industry has demonstrated a substantial history of safe operation and serves as a vital element that ensures the stability of the nation's electricity grids.

Near the year 2030, unless second license renewals are granted, decommissioning of the current fleet of nuclear power plants will begin. Over the next three decades beyond 2030, decommissioning of the existing fleet would result in a loss of nearly 100-GWe of emission-free electrical generating capacity, leaving a shortfall of required emission- free generating capacity. Early (prior to 60 years of operation) shutdowns due to economic factors will increase this shortfall. Hence, the continued safe and economical operation of current plants to and beyond the current license limit of 60 years is an important option for supplying needed electricity and maintaining the existing level of emission-free power generation capability at a fraction of the cost of building new plants.

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 932–936, 2020.
https://doi.org/10.1007/978-3-030-48021-9_103

Decisions on second license renewal and required investments to support long-term operation are being made by plant owners. On January 31, 2018, Florida Power & Light filed to renew its licenses for its Turkey Point nuclear Units 3 and 4 with the NRC that, upon approval, will allow the utility to operate the units until 2052 and 2053, respectively. This follows previous announcements by Dominion Generation Group of second license renewal submittal plans for its Surry Power Station and by Exelon Corporation for two operating reactors at the Peach Bottom Atomic Power Station in Southeastern Pennsylvania.

The LWRS Program has worked with owners/operators to provide the technical basis for second license renewal specifically, as well as to address current and future issues needed to ensure a long-term viable source of nuclear power generation remains available to U.S. electricity markets.

2 Program Description

The Light Water Reactor Sustainability (LWRS) Program is a research and development (R&D) program sponsored by the U.S. Department of Energy (DOE), performed in close collaboration and cooperation with industry. The LWRS Program provides technical foundations for the continued operation of the nation's nuclear power plants, utilizing the unique capabilities of the national laboratory system.

(1) This involves leveraging national laboratory facilities, staff, and expertise to conduct research needed to inform decisions, demonstrate technical solutions, and provide methods needed for the long- term management and operation of nuclear power systems.

The following research and development pathways address the objectives of the LWRS Program:

Materials Research. R&D to develop the scientific basis for understanding and predicting long-term environmental degradation behavior of materials in nuclear power plants. This work will provide data and methods to assess the performance of SSCs essential to safe and sustained nuclear power plant operations. The R&D products will be used to define operational limits and aging mitigation approaches for materials in nuclear power plant SSCs.

Plant Modernization. R&D to address nuclear plant economic viability in current and future energy markets through innovation, efficiency gains, and business model transformation through digital technologies. This includes addressing the long-term aging and modernization or replacement of legacy instrumentation and control technologies by research, development, and testing of new instrumentation and control technologies and advanced condition monitoring technologies for more automated and reliable plant operation. The R&D products will enable modernization of plant systems and processes while building a technology-centric business model platform that supports improved performance at lower cost.

Risk-Informed Systems Analysis. R&D to support decision-making related to economics, reliability, and safety, providing integrated plant systems analysis solutions through collaborative demonstrations to enhance economic competitiveness of the operating fleet. This pathway will demonstrate a risk-assessment method coupled to safety margin quantification that can be used by decision makers as part of their margin recovery strategies; and apply the RISA toolkit that enables more accurate representation (e.g., reduce conservatisms, increase realism) of margins for the long-term benefit of nuclear assets. The R&D products will be used to optimize plant economic performance and safety by incorporating plant impacts, physical aging, and degradation processes.

Reactor Safety Technologies. R&D to improve understanding of beyond- design basis events and reduce uncertainty in severe accident progression, phenomenology, and outcomes using existing analytical codes and information gleaned from severe accidents, in particular the Fukushima Daiichi events. This information has been used to aid in the development of mitigating strategies and improving severe accident management guidelines for the current light water reactor fleet. In addition, methods for enhancing plant resilience to accident initiating events have also been explored.

3 Key Outcomes

Measurable milestones have been developed for each of the pathways; these include both near-term (i.e., 1 to 5 years) and longer- term (i.e., beyond 5 years) milestones. High-level planned accomplishments in the near-term include:

Provide a mechanistic understanding of key materials degradation processes, predictive capabilities, and high-quality data to inform decisions and processes by both industry and regulators including:

- Containment Inspection Guidelines for extended-service conditions
- Predictive models for swelling in light water reactor components, aging of cast austenitic stainless-steel components, cable degradation, and nickel-base alloy stress corrosion cracking susceptibility.
- Model for transition temperature shifts in reactor pressure vessel steels, precipitate phase stability and formation in internal primary water coolant components and reactor pressure vessel steels, and environmentally-assisted fatigue in light water reactor components.
- Methodology and techniques for a system for nondestructive examination of concrete sections, impact assessment of alkali- silica reaction affected concrete, radiation-induced changes and synergistic environmental stressor damage in concrete and cable insulation.
- Development and transfer of weld repair technique for irradiated materials to industry and the evaluation of new replacement alloys.

Technical basis and supporting reports and studies needed to broadly implement digital technologies and modernize plants including:

- Methods and studies on migrating existing analog control rooms to hybrid integrated control room technologies incorporating digital systems, advanced alarm systems, and control room computer-based procedures.
- Cost-benefit studies for deploying technologies that are the subject of R&D in actual nuclear power plants.
- Human performance improvement for nuclear power plant field workers based on application of technologies, such as radio- frequency identification (RFID), for management of tools and materials used in nuclear plant maintenance.
- New analytical capabilities for reducing operational and schedule adherence risks for nuclear refueling outages.
- Advanced online monitoring technologies used for applications of structural health monitoring of nuclear plant passive components, such as structural concrete and process piping.

Integrated Probabilistic Risk Assessments with cost analysis, and multi- physics best estimate plus uncertainty engineering tools to optimize the economic and safety performance of existing nuclear power plants:

- Demonstrate enhanced plant resiliency with industry adoption of accident tolerant fuel, optimal utilization of FLEX equipment, augmented or new passive cooling systems, and improved fuel cycle efficiency.
- Demonstrate improved economic performance of existing nuclear power plants by recovering safety margins by reducing uncertainties and conservatisms of legacy licensing, design and analysis bases through applications of the Risk-Informed Systems Analysis (RISA) toolkit.
- Use the RISA toolkit to quantify predicted market revenue gaps for existing fleet through analyzing electricity market attributes and stress case conditions.

Improved understanding of and reduced uncertainty in severe accident progression, phenomenology, and outcomes, including:

- Gap analysis of accident tolerant components and severe accident analysis.
- Validated Terry turbine models for input into system-level severe accident analysis codes (e.g., MAAP, MELCOR) for evaluating extended core cooling capabilities during off-normal plant conditions.
- Validated a set of high fidelity analysis tools to support industry efforts in the implementation of boiling water reactor (BWR) severe accident water management (SAWM) strategies.

4 Summary

Government cost-sharing and involvement is required to promote the necessary programs that are of crucial long-term, strategic importance.

The LWRS Program, by incorporating collaborative industry stakeholder inputs and shared costs, supports the strategic national interest of maintaining nuclear power as an available resource.

Decisions on second license renewal and required investments to support long-term operation are being made by plant owners. On January 31, 2018, Florida Power & Light filed to renew its licenses for its Turkey Point nuclear Units 3 and 4 with the NRC that, upon approval, will allow the utility to operate the units until 2052 and 2053, respectively. This follows previous announcements by Dominion Generation Group of second license renewal submittal plans for its Surry Power Station and by Exelon Corporation for two operating reactors at the Peach Bottom Atomic Power Station in Southeastern Pennsylvania.

The LWRS Program has worked with owner-operators to provide the technical basis for second license renewal specifically, as well as to address current and future issues needed to ensure a long-term viable source of nuclear power generation remains available to U.S. electricity markets.

The science-based technical results from the LWRS Program provide data, methods, and technologies that are used by owners/operators to make informed decisions and take actions needed to ensure the continued operation of the existing U.S. light water reactor (LWR) fleet. Through the variety of R&D activities carried out together with and used by industry, the LWRS Program reduces some of the key uncertainties and risks that many owners/operators face regarding the long-term performance of vital materials, plant modernization, efficiency improvement, and other issues needed to make the investments required for nuclear power plant operation periods to and beyond 60 years.

Reference

Light Water Reactor Sustainability Program – Integrated Program Plan. https://lwrs.inl.gov. Accessed 5 Feb 2017

Deep Learning for Short-Term Energy Load Forecasting Using Influential Factors

Cristina Heghedus(✉), Antorweep Chakravorty, and Chunming Rong

Department of Electrical Engineering and Computer Science,
University of Stavanger, Stavanger, Norway
{cristina.heghedus, antorweep.chakravorty,
chunming.rong}@uis.no

Abstract. Large scale technological, economical and environmental changes led to increased energy consumption all over the world. Electrical energy became an indispensable daily factor due to the large automation and emergence of smart devices and equipments. Therefore, efficient resource management is a major goal of the energy market. In order to adapt electricity supply to demand and to prevent waste, companies rely on accurate and granular forecasting of future consumption. Electricity demand forecasting is a complex and deep task as it is influenced often by weather conditions and seasonality. Thus, the aim of this work is to offer more accurate short-term forecasting, evaluating data produced by electricity meters, from individual house holds, and considering factors influencing the consumption.

1 Introduction

Energy Load Forecasting (ELF) is a highly treated subject in various and sophisticated levels. Driven by energy supplier's and consumer's needs, researchers and data scientists have been focusing their work on aggregated and non-aggregated electricity demand forecasting. Moreover, forecasting is divided in three main categories which are as follows: long-term, for months, seasons or longer periods of the year, medium-term, for a few days, weeks or up to a few months ahead and short-term, for a few minutes ahead or a few days ahead [1, 3]. This last category is usually more challenging, as forecasting can be also done for seconds, but data is not always available for such short periods.

Another aspect contributing to the complexity of future energy demand estimation is granular forecasting, which means decomposing the overall electricity consumption to the level of appliances or single consumer. Energy Informatics, is meant to address energy related challenges and questions using Information and Communication Technology methods [2]. These methods vary from simple mathematical approaches to modern Artificial Intelligence (AI) algorithms, which capture non-linear relationships and correlation of lagged datasets.

Solid and accurate estimation of future energy consumption holds great advantages for electricity suppliers and consumers as well. Data collected through Electricity Meters combined with high performance AI models offers revolutionary effectiveness

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 937–945, 2020.
https://doi.org/10.1007/978-3-030-48021-9_104

[3]. Capturing characteristics such as trends, seasonality or outliers, highly benefits decision making and optimal planning.

Thus, in this work we propose an AI model, for short-term forecasting based on energy and weather data (as influential factor). The reason of choosing short-term forecasting is that it represents a rather hard task compared to medium and long term forecasting, where it is easier to capture the trend and estimate consumption, based on average daily and monthly usage.

2 Theoretical Framework

In the era of Big Data decision making and process planning within companies are data driven, therefore scientists and engineers are in constant improvement seeking for learning future behavior and customer needs. Predictive Analytics has been highly used in this scope in the energy field, that is, matching energy generation to the consumption. This way, consumers can benefit from the market driven pricing and suppliers have control over the production, managing Earth's resources in the most efficient manners.

2.1 AI Footprint on ELF

Throughout the years AI truly added value to the process of forecasting the energy load with strong and robust techniques. Competitive results have been carried out by Machine Learning (ML) models such as Support Vector Machine (SVM) [4, 5], Back Propagation Neural Network (BPNN) [6], Gradient Boosting Machines [7, 8], Fuzzy Logic [9, 10], Hidden Markov Model and so forth. However, the rapidly increasing volume of generated data in combination with higher computational power led to Deep Learning (DL), a group of more sophisticated ML algorithms.

DL models in fact imply deeper and more complex structure in terms of hidden layers, nodes, interaction of activation functions and information transfer between layers. Deep Neural Networks (DNN) have been used successfully in ELF as these prove a high compatibility with Time Series data. DNN's can vary from Multilayer Artificial Neural Networks to more complex architecture such as Recurrent Neural Networks (RNN). The feasibility of using RNN on electricity data is determined by its chainlike structure which allows learning over time.

Several types of RNN have been applied in the energy field in load and in electricity price forecasting such as Long-Short Term Memory (LSTM) [11, 12], Gated Recurrent Unit (GRU) [12], Non-Linear AutoRegressive with eXogenous inputs Recurrent Network (NARX) [13], Conditional Restricted Boltzmann Machine (CRBM) and Factored CRBM [14]. These, and many more DL models, are suitable for sensor data, having great applicability in the industry of Internet of Thing (IoT), Smart Homes/Cities and supply chain. A major capability of a RNN is recognizing patterns in sequences of data, hitch makes them the most powerful and useful type of neural network.

2.2 Influential Factors

Enormous amount of sensor data is being produced all over the world by different devices. This data is diverse and unprocessed, but it carries extremely valuable information and it contributes to enterprises resource planning. Therefore, the mail goal of data scientists is to extract the most relevant knowledge out of energy consumption data even when this is influenced by other factors. Appliances automation evidently put a major stamp on electricity consumption, especially when it comes to heating devices. Thus, weather conditions undoubtedly influence the consumption pattern of the households.

Energy consumption data combined with weather data highlights two primary characteristics, trend and seasonality. The trend in a time series dataset indicates either an increasing or a decreasing tendency. The seasonality illustrates a pattern change in the usage with respect to different seasons of the year [15]. However, the problem of estimating energy consumption is way more complex due to the global climate change, which induces inconsistency in monthly and seasonal temperatures of consecutive years.

Beside weather imprint, there are many various circumstances influencing the electricity consumption, formed by individual household characteristics. Regarding heating, most relevant aspects are the size of the house, the material constructed with, the number of cohabitants, isolation and geographical position of the house. Moreover, the houses period of occupancy, specific indoor activities such as using the stove, washing machines, TV sets, computers, etc., also effects the consumption pattern [15, 16].

Figure 1 represents the decision making process, based on raw data obtained from Electricity Meters, and it emphasizes the possibilities of using the clean, processed data. Forecasting is made either based on simple reading of the consumption or by including weather and other significant aspects.

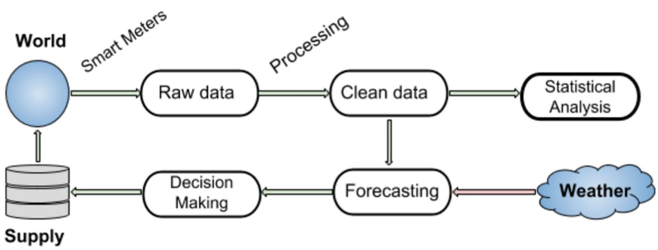


Fig. 1. Process of data analysis and forecasting.

In this work, we follow the steps presented din Fig. 1, which means we get the raw electricity consumption data, apply processing techniques to obtain clean data, then following the vertical arrow, we forecast future consumption. In addition to the clean data we set as input weather parameters as well. This way, our model will have two inputs (past consumption and temperature) and one output (future consumption).

3 Approach

In the first part of this section we will present the data and the necessary steps taken to justify the choice of our model. The second part starts with a short description of neural networks and it continues with a detailed description of our model.

4 Data Characteristics

Time series data for four consecutive months and multiple households was used in this work, provided by the local electricity supplier. The readings are made on interval of 10 s, adding up to over eight thousands of meter points per day. In other words, we have a time series dataset, containing timestamps and read electricity consumption values. In order to obtain a more accurate forecasting, we wish to investigate the relation of the observed electricity consumption values.

Thus, we create 3 more columns of data containing observations with lag 1, lag 2 and lag 3, by shifting the values by 1, then 2 and 3 positions downwards. Next step is to plot the created dataset and analyze the position of the values and the shape and direction that these form on the plot. The result is as follows:

The relationship of a time step with one ore more previous time steps is one important characteristic of time series data. Figure 2 shows the relationship between each observation and 3 lags of it. That is, the relationship of each observation with the 3 previous observations. These plots show a strong positive correlation between observations and their lagged values, as the points are gathered from the bottom left to the top right corner along the diagonal. More specifically, the strength is given by the multitude of points tightly clustered to the diagonal.

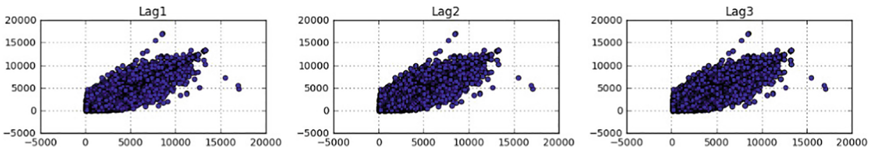


Fig. 2. Lag scatter plots at lag1 (t−1), lag2 (t−2) and lag3 (t−3).

5 Proposed Model

For a better understanding of how our NARX model works, we introduce the concept of a Neural Network (NN). Similar to the human brain, NNs have a "learning" process, which means they get inputs from the real world and produce a requested output by using a set of mathematical operations. NNs have in their composition an Input layer, a Hidden layer and an Output layer, each containing "nodes" (neurons) allowing the information to flow in a specified way. For the sake of clarity we will call this type a Simple Neural Network (SNN), as shown in Fig. 3(a). Here, all the nodes from 2 consecutive layers are fully connected to each other by edges. These edges are marked with "weights", which express the importance of the inputs to the output. Furthermore, a function called "activation function" will determine the values that are passed to the output.

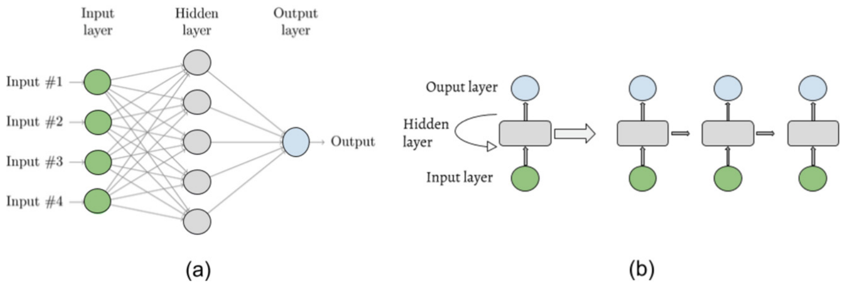


Fig. 3. Interaction of layers in a Simple Neural Network (a) and a Recurrent Neural Network (b)

Next, we present a more complex type of NN, namely Recurrent Neural Network (RNN), better fitting our dataset. RNNs are more similar to how human brain works, as they allow information to persist over time. These learn form input that contains current and past values and the information flows in a directed cycle, as shown in Fig. 3(b). Since electricity data is of time series nature, this type of NN can easily forecast future consumption based on past values.

NNs are able to solve 2 major type of problems. On one hand, these are capable of *classifying* objects based on their characteristics. On the other hand, NN have high performance on *regression* problems, such weather forecast, stock market prediction or in our case, electricity consumption forecasting [17].

Our model for this work is a Recurrent Neural Network, called Non-Linear AutoRegressive with eXogenous inputs (NARX). The strong relationship between points in the dataset and their past values, as described in Sect. 3, validates the choice of an RNN for forecasting future values. The nodes in an RNN form a directed cycle allowing information to persist in time. This information contains specific patterns and past features which can influence future behavior [17].

Beside information about past electricity consumption we feed simultaneously weather (temperature) data as inputs to NARX neural network. This model has been used previously on forecasting electricity price [10], however not on electricity consumption. Therefore, the contribution of this work is building a NARX capable of predicting electricity consumption based on past usage and weather condition.

The output of the model at given time, Y_t, for our model is calculated as a combination of the inputs and previous outputs. For the sake of simplicity we call the previous outputs as "steps". One or multiple steps can be maintained for a particular value. A representation of the network structure is presented in Fig. 4.

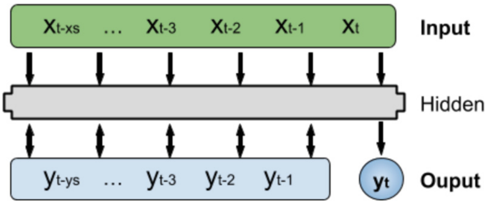


Fig. 4. Interaction of input, hidden and output layers in NARX network.

Here, the inputs are transmitted to the nodes in the hidden layer and activation function yields the values transmitted to the output layer. However, communication of the hidden and output layers is bidirectional. The current output is computed using its steps, and the formula is as follows:

$$Y_t = F_O(F_H(\sum_{i=0}^{xs} X_{t-i} \times Uh_i + \sum_{j=1}^{ys} Y_{t-j} \times Vh_j + Bh) \times \sum_{k=1}^{N} Wo_k + Bo) \quad (1)$$

where, F_O, F_H are output mapping function and activation function for the hidden layer respectively, Uh_i, Vh_j, Wo_k, Bh, Bo are weights and biases for the hidden and output layers and finally xs, ys represent the number of inputs and number of steps fed to the hidden layer from both the input and the output layer, and N the number of the output nodes.

First, we take the linear combination of the inputs and steps, we add the biases and apply the activation function F_H. Then, we apply again weights and biases and run the results through F_O function which will map our outputs. The type of activation functions and the size of the hidden layer will be both determined in the parameter optimization process. Based on Fig. 4 and Eq. 1, a high level representation of the hidden and output layer computations is given below:

$$H = F_H(X \otimes U + Y \otimes V + Bh) \quad (2)$$

$$O = F_O(H \otimes W + B_O) \quad (3)$$

In order to validate the performance of NARX neural network on the given dataset, we compute the MSE (Mean Square Error), computed as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - P_i)^2 \quad (4)$$

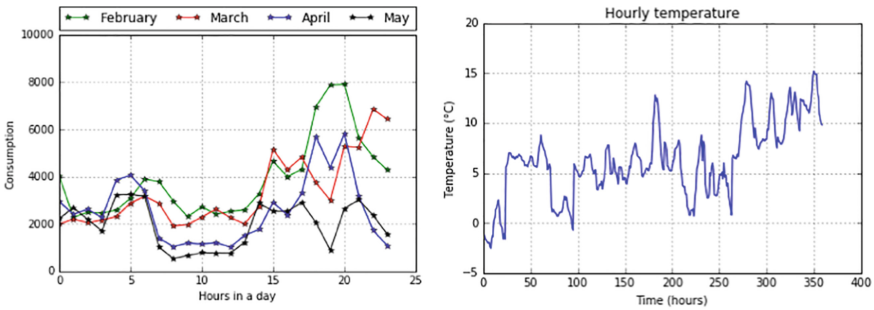
where, Y_i are the true consumption values, P_i are the predicted values and n is the size of the test set. The size of the test set will be determined based of the type of experiment that will be conducted. The goal is to train the network to carry out more accurate predictions, when using temperature data in addition as input. The next section presents the results of the conducted experiments.

6 Experimental Results

Electricity consumption is highly dependent on geographical position, seasonality, weather outliers and period of occupancy. Moreover, often we find a noticeable difference between the week and week end days. Therefore, we build a subset of the given dataset consisting only of Monday's consumption.

The dataset used for experiments, was provided by the local energy supplier, in a Norwegian Southwest city. In this region the weather is quite constant, with a few outliers (days or even weeks). However, the main source for heating and cooking, washing and cooling equipment is electricity, resulting in continuous electricity usage all over the year. We consider as influencer only the weather, as other factors such a number of inhabitants is not provided and for instance the houses construction material is less relevant (all being similar).

The decreasing trend is attested by Fig. 5, which represents the hourly average of the consumption of four days of Monday, in February, March, April and May. In this period the temperature is increasing and this leads to a valid dataset for our model, due to inversely proportional relationship between the usage and temperature.



(a) (b)

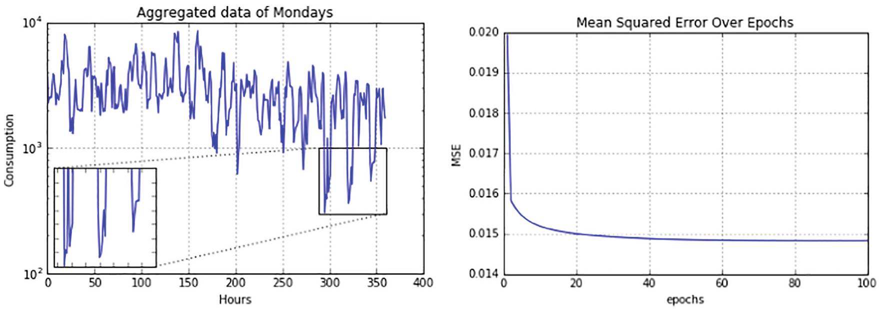
Fig. 5. The deceasing consumption (a) and increasing temperature over months (b).

The first experiment is conducted using the whole dataset which consists of plain time series data of the electricity consumption. In this case the neural network is built as follows: one input node, two hidden and one output node. As testing data we use the last Monday of May.

6.1 Temperature Effect

In the second experiment we increase the complexity of the forecasting problem by adding the temperature values to the input. Thus, the neural network has the following structure: two input nodes, two hidden and one output node. Here, we feed to the network a set of aggregated inputs over hours from all Mondays in the dataset, except the last one. The last Monday will be used as test data and the predictions will be made for each hour on that day.

This outperformed the first experiment (without temperature values). In Fig. 6 we notice a different consumption pattern in the last three Mondays of the dataset (the zoomed in box). However, the network learned successfully from the training set and produced the presented MSE in Fig. 6.



(a) (b)

Fig. 6. The outliers in the dataset (a) and the MSE of predictions (b).

7 Conclusion

In main contribution of this paper is to offer a technique of forecasting short- term electricity consumption, for a cheaper and more efficient finite resource management. We built a dataset containing values of similar days (Mondays), over a period which implies season change. First, the raw time series was fed to the NARX neural network and the values of the last day in the dataset were predicted. Second, we added measured temperature to the input and the model has been proven to carry out better results.

For future work, the same approach can be applied to a larger dataset, as the prediction based on only similar days of the week is quite successful. The same approach can be tested on medium and long term forecast, for creating a more general image of the future consumption pattern, however the main interest of this work was short-term forecasting, as it offers detailed information on future consumption.

References

1. Gupta, V., Pal, S.: An overview of different types of load forecasting methods and the factors affecting the load forecasting. Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET) **5**, 729–733 (2017)
2. Huang, B., Bai, X., Zhou, Z., Cui, Q., Zhu, D., Hu, R.: Energy informatics: fundamentals and standardization. ICT Express **3**(2), 76–80 (2017)
3. Zor, K., Timur, O., Teke, A.: A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting. In: International Youth Conference on Energy, August 2017

4. Pai, P.F., Hong, W.C.: Support vector machines with simulated annealing algorithms in electricity load forecasting. Energy Convers. Manag. **46**, 2669–2688 (2005)
5. Zhao, H.X., Magoulès, F.: A review on the prediction of building energy consumption. Renew. Sustain. Energy Rev. **16**(6), 3586–3592 (2012)
6. Ahmad, A.S., Hassan, M.Y., Abdullah, M.P., Rahman, H.A., Hussin, F., Abdullah, H., Saidur, R.: A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew. Sustain. Energy Rev. **33**, 102–109 (2014)
7. Lloyd, J.R.: GEFCom2012 hierarchical load forecasting: gradient boosting machines and Gaussian processes. Int. J. Forecast. **30**, 369–374 (2014)
8. Fan, C., Xiao, F., Zhao, Y.: A short-term building cooling load prediction method using deep learning algorithms. Appl. Energy **195**, 222–233 (2017)
9. Kiartzis, S.J., Bakirtzis, A.G.: A fuzzy expert system for peak load forecasting: application to the greek power system. In: Proceedings of the 10th Mediterranean Electrotechnical Conference, vol. 3, pp. 1097–1100 (2000)
10. Skarman, S.E., Georgiopoulous, M.: Short-term electrical load forecasting using a fuzzy ARTMAP neural network. In: Proceedings of SPIE, pp. 181–191 (1998)
11. Marino, D.L., Amarasinghe, K., Manic, M.: Building energy load forecasting using deep neural networks. In: 42nd Annual Conference of the IEEE Industrial Electronics Society, IECON 2016 (2016)
12. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. arXiv:1412.3555v1, [cs.NE], 11 December 2014
13. Marcjasz, G., Uniejewski, B., Weron, R.: On the importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks. Int. J. Forecast. **35**, 1520–1532 (2018)
14. Mocanu, E., Nguyen, P.H., Gibescu, M., Kling, W.L.: Deep learning for estimating building energy consumption. Sustain. Energy Grids Netw. **6**, 91–99 (2016)
15. Ghelardoni, L., Ghio, A., Anguita, D.: Energy load forecasting using empirical mode decomposition and support vector regression. IEEE Trans. Smart Grid **4**(1), 549–556 (2013)
16. Yao, R., Steemers, K.: A method of formulating energy load profile for domestic buildings in the UK. Energy Build. **37**(6), 663–671 (2005)
17. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning

Author Index

J. P. Liyanage et al. (Eds.): *Engineering Assets and Public Infrastructures in the Age of Digitalization*, LNME, pp. 947–950, 2020.
https://doi.org/10.1007/978-3-030-48021-9

W

X

Y

Z