UiT The Arctic University of Norway

Faculty of Science and Technology
Department of Technology and Safety

Performance Measurement System in Complex Environment: Observed and Unobserved Risk Factors

Rezgar Zaki

A dissertation for the degree of Philosophiae Doctor, September 2020

Performance Measurement System in Complex Environment: Observed and Unobserved Risk Factors

By

Rezgar Zaki

Thesis submitted in fulfilment of the requirement for the degree of
Doctor of Philosophy (PhD)

UiT The Arctic University of Norway

Faculty of Science and Technology
Department of Technology and Safety
2020

Abstract

World demand for energy leads industry to harvest energy in complex environment with harsh conditions and sensitive areas, such as the Arctic region – one of the last remaining wild places in the world – with potentially harmful consequences. Moreover, over the past few decades, the increasing trend of melting sea ice in the Arctic has provided increased access and has created new opportunities for economic development within metals and minerals, fisheries, cargo shipping, cruising, subsea telecom cables and pipelines. However, development of the Arctic resources is assumed to be technologically and economically challenging and risky. Studies reveal that, due to low temperatures, sea ice, polar low pressures, poor visibility, seasonal darkness limitations to the logistics of supplies, etc., Arctic operational conditions have significant effects on the performance of components and industry activities in various ways, including increasing failure rate and repair time, and can cause different types of production losses. The optimal functioning of technical systems involved in design and operation in the Arctic faces numerous challenges, in order to succeed in a globally competitive market with limited resources. The concept of the Performance Measurement System (PMS) is frequently used by industries and has been shown to be an essential concept for improving efficiency and effectiveness and supporting the design, planning, and managing of a company; PMS refers to output results obtained from a system that permits evaluation and comparison, relative to past results or other companies. PMS needs up-to-date and accurate performance information on its business. This performance information needs to be integrated, dynamic and accessible, to assist fast decision-making. However, performance terminologies and standards for the Arctic reveal that the Performance Indicators (PIs) measured by industries though important, are not enough and could still be improved by identifying more important indicators, which contribute to a successful PMS in the Arctic. Hence, the development and continuous improvement of PMSs and the identification of more PIs for judging performance of equipment in the Arctic are critical for industry success. Moreover, the quantification of performance is complex, as it involves various indicators with different perspectives at various hierarchical levels. The lack of correct sources of information and data on PIs and suitable statistical models and standard approaches are a barrier to the successful quantification of PIs. Operation and maintenance data are often collected from multiple and distributed units in different operational conditions, which can introduce heterogeneity into the data. Part of such heterogeneity can be explained by the observable risk factors, whose values and the way that they can affect the item's PIs are known. However, some factors which may affect PIs are typically unknown (unobserved risk factors), leading to unobserved heterogeneity. Nevertheless, many researchers have ignored the effect of observed and un-observed risk factors, and this may lead to erroneous model selection, as well as wrong conclusions and decisions. The statistics models must be able to quantify the effect of observed and unobserved risk factors on PIs and must be built based on correct assumptions that reflect the operational conditions.

In this thesis, a methodology for the monitoring and analysis of operation and maintenance performance is developed. The aim is to facilitate improvements and the optimization of decision-making for operation and maintenance in the Arctic. Firstly, a brief survey of technological and operational challenges in the Arctic region, from a performance point of view, is presented. Further, appropriate performance indicators/criteria that need to be measured for judging the performance of equipment/systems in the Arctic that contribute to a successful PMS will be discussed. Thereafter, the study focuses on improvement and modifying the available statistical approach for the prediction of PIs, considering the effect of observed and unobserved risk factors.

The thesis consists of two parts. The first part gives an introductory summary of the study, followed by a discussion of the appended papers and conclusions.

The second part consist of three appended papers. The first paper concerns the development of a model for improving safety performance measurement. The second paper is a study of the reliability performance indicator, while Paper C is concerned with the maintainability performance indicator.

Keywords: operation and maintenance, performance measurement, performance indicators, observed and unobserved covariates, proportional hazard model, proportional repair model.

Acknowledgment

This thesis is submitted as a partial fulfillment of the requirements for the degree of Doctor of Philosophy (PhD) at the Arctic University of Norway (UiT). The research work was carried out in the subject area of Operation and Safety Engineering in the period from January 2016 to November 2019 at UiT.

I wish to express my gratitude to my supervisor, Professor Abbas Barabadi, for introducing me to the research area, and for providing me with all the necessary facilities and continuous support during this study. Thank you!

I would like to express my thanks to Yngve Birkelund, Head of the Department, Gunn-Helene Turi, Bjørn-Morten Batalden, Jarle André Johansen, Javad Barabady, Lisbeth Klausen, Leif Hansen, Yuan Fuqing, Masoud Naseri, Marith Gabrielsen, Jinmei Lu, Kåre Johansen, Are Sydnes, Vegard Nergård, and Maria Sydnes, for providing me with all necessary facilities, their valuable guidance and support during my study.

I would also like to thank my co-supervisors, Amir Garmabaki, at the Luleå University of Technology, Sweden and Yonas Zewdu Ayele at Østfold University College for their invaluable comments and guidance for improvements to the thesis. I would also like to extend my thanks to Ali Nouri for his support and for sharing his professional experience during this study.

I would like to express my appreciation to all my colleagues and research fellows at the Department of Engineering and Safety at the UiT, including Amir Nadem, Bjarte Rød, Lise Lotte Evenseth, Johana Evelyn, Per Roald Leikanger, Brian Murray, Alyona Dekhtyareva, Vitaly Dekhtyarev, Mustafa Albara, Abolfazl Shojaeibarjouei and all those who created a warm and friendly atmosphere at UiT.

I wish to express my sincere gratitude to my daughters, Arina and Atena, for their support and understanding. Finally, I wish to express my deep and heartfelt gratitude to my mother, Khorshid, and my father, Abdullah, for their encouragement and inconceivable support throughout my entire life.

Rezgar Zaki

September 2020

Tromsø, Norway

Abbreviations

AFT	Accelerated failure time model
AIC	Akaike information criterion
AvGOF	The average values from the Kolmogorov-Smirnov goodness-of-fit test
AvPLOT	The average values from the correlation coefficient (PLOT) test
BIC	Bayesian information criterion
Conf.	Confidence
DMP	Dynamic Multi-Dimensional Performance Framework
DT.	Dump truck
EER	Evacuation, Escape and Rescue (EER)
EMPHM	Extension mixed proportional hazards model
FFL	Free-fall lifeboats
HLP	Helicopter landing pad
HR	Likelihood ratios
IEI	Icing effect index
IEID	Icing effect index on dependability
IEIP	Icing effect index for performability
IEIS	Icing effect index on survivability
IRIP	Icing risk index for performability
LCC	Life Cycle Cost
LKV	The average values from the likelihood value test
MFM	Mixture frailty model
MPHM	Mixed proportional hazard model
MTTR	Mean time to repair
NHPP	Non-homogeneous Poisson process
OAE	Overall Asset Effectiveness
OEE	Overall Equipment Effectiveness
OFE	Overall Factory Effectiveness
OPE	Overall Plant Effectiveness
P	Observed covariates
PEE	Production Equipment Effectiveness
PH	Proportional hazards
PHM	Proportional hazard model
PI	Performance Indicator
PMS	Performance Measurement System
POI	Probability of ice accretion
PRM	Proportional repair model
Q test	Cochran's Q test
RAM	Reliability, availability, and maintainability
RDF	Results and Determinants Framework

ROI	Return on Investment
ROS	Return on Sales
S.E.	Standard error
Std	Standard deviation
t	System/machine operation time or time to repair
TEEP	Total Equipment Effectiveness Performance
TBF	Time between failure
TRP	Trend renewal process
TTR	Time to repair
WCED	World Commission on Environment and Development

Notations

$\lambda_0(t)$	Baseline hazard rate
$\lambda_0 i(t)$	Baseline hazard rate of r'th stratum
$\mu_0(t)$	Baseline repair rate
g(A)	Gamma distribution
$\Gamma(\theta)$	Gamma function
$\mathrm{I_s}$	Icing effect index for safety
$\mathrm{I_m^{\alpha_m}}$	Icing effect index on survivability for the maintainability
$\mathrm{I_q^{\alpha_q}}$	Icing effect index on survivability for the quality
$\mathrm{I_r^{\alpha_r}}$	Icing effect index on survivability for the reliability
$\mathrm{I_{su}}$	Icing effect index on sustainability
$\mathrm{m}\left(t, z_i, z_j(t)\right)$	The probability density function of repair process
$M_0(t)$	The baseline maintainability function
$M_{0m}(\mathrm{t})$	The baseline maintainability function of the several repair processes
$\lambda_j\left(t, z_i, z_j(t)\middle\|A\right)$	The conditional failure rate of time-dependent and time-independent observed covariates and unobserved covariates
$\lambda_r\left(t, z_i, z_j(t)\middle\|A\right)$	The conditional failure rate time-dependent and time-independent observed covariates and unobserved covariates of r'th stratum
$\mathrm{M}\left(t, z_i, z_j(t)\middle\|A\right)$	The conditional maintainability function of time-dependent and time-independent observed covariates and unobserved covariates
$R_p\left(t, z_i, z_j(t)\middle\|A\right)$	The conditional reliability of time-dependent and time-independent observed covariates and unobserved covariates
$\mu\left(t, z_i, z_j(t)\middle\|A\right)$	The conditional repair rate of time-dependent and time-independent observed covariates and unobserved covariates
$\psi(z, z(t); P; \delta)$	The function of time-independent and time-dependent observed covariates
γ_k	The proportion of the repair tasks belonging to the kth repair process
δ and η	The regression coefficient observed covariates
η_k	The scale parameter
β_k	The shape parameter
$\mathrm{M}\left(t, z_i, z_j(t)\right)$	The unconditional maintainability function of time-dependent and time-independent observed covariates
$\mathrm{R}\left(t, z_i, z_j(t)\right)$	The unconditional reliability function of time-dependent and time-independent observed covariates
$\mathrm{R}_r\left(t, z_i, z_j(t)\right)$	The unconditional reliability of time-dependent and time-independent observed covariates of r'th stratum
$z_j(t)$	Time-dependent observed covariates
z_i	Time-independent observed covariates
θ	Variance
$\gamma_{\mathrm{Dep.}} \geq 0$	Weight vectors of dependability on performability
$\alpha_\mathrm{m} \geq 0$	Weight vectors of maintainability
$\alpha_\mathrm{q} \geq 0$	Weight vectors of quality

$\alpha_{r} \geq 0$	Weight vectors of reliability
$\beta_{sur.} \geq 0$	Weight vectors of safety
$\beta_{sur.} \geq 0$	Weight vectors of survivability
$\gamma_{su.} \geq 0$	Weight vectors of sustainability on performability

Some Basic Definitions

Availability	The ability of an item to be in a state to perform a required function under given conditions at a given instant of time or over a given time interval, assuming that the required external resources are provided (IEC, 2019).
Capacity performance	An item's ability to deliver according to design capacity and/or current demands (requirements/needs) in a fixed period of time with given production resources (Standard, 1998b).
Dependability	Collective term which describes availability performance and its influencing factors, namely reliability performance, maintainability performance, and maintenance support performance (IEC, 2019).
Failure	A fault is the state of an item characterized by inability to perform a required function, excluding the inability during preventive maintenance or other planned actions, or due to lack of external resources. A fault is often the result of a failure of the item itself but may exist without prior failure (IEC, 2019).
Failure rate	The failure rate is the limit, if it exists, of the quotient of the conditional probability that the instant of a failure of a non-repaired item falls within a given time interval (t, t+Δt) and the duration of this time interval, Δt, when Δt tends to zero, given that the item has not failed up to the beginning of the time interval (IEC, 2019).
Maintainability	The ability of an item, under given conditions of use, to be retained in, or restored to, a state in which it can perform a required function, when maintenance is performed under given conditions and using stated procedures and resources (IEC, 2019).
Maintenance	The ability of a maintenance organization, under given conditions, to provide upon demand the resources required to maintain an item, under a given maintenance policy (IEC, 2019).
Mean time between failures	The expectation of time between failures (IEC, 2019).
Mean time to repair	The expectation of the time to restoration (IEC, 2019).
Observed covariates	All those factors which may have an influence on the reliability characteristics of a system are called observed covariates. Observed covariates are also called explanatory variables. Examples of observed covariates are the operating environment (dust, temperature and humidity, etc.), the skill of operators, etc. (Klein et al., 2016).
Performance	The characteristics defining the ability of a measuring instrument to achieve the intended functions (IEC, 2019).
Reliability	The ability of an item to perform a required function under given conditions for a given time interval (IEC, 2019).

Risk	The combination of the probability of occurrence of harm and the severity of that harm (IEC, 2019).
Strata	The strata of a data set are obtained by grouping the data on the basis of discrete values of a single covariate or combinations of a set of covariates (Kalbfleisch and Ross, 1983)[1].
Sustainable development	Development that meets the needs of the present without compromising the ability of future generations to meet their own needs (Brundtland et al., 1987).
Supportability	Ability to be supported to sustain the required availability with a defined operational profile and given logistic and maintenance resources (IEC, 2019).

List of appended papers

Paper A Barabadi, Abbas; Zaki, Rezgar; Garmabaki, Amir Hossein Soleiman. Designing for performability: An icing risk index for Arctic offshore. Cold Regions Science and Technology 2016; Volum 124. ISSN 0165-232X.s 77 - 86.s

Paper B Zaki, Rezgar; Barabadi, Abbas; Nouri Qarahasanlou, Ali. Observed and unobserved heterogeneity in failure data analysis: A case study. Submitted to Part O: Journal of Risk and Reliability. Under review.

Paper C Zaki, Rezgar; Barabadi, Abbas; Garmabaki, Amir Hossein Soleiman; Nouri Qarahasanlou, Ali. A mixture frailty model for maintainability analysis of mechanical components: a case study. International Journal of Systems Assurance Engineering and Management 2019; Volum 10 (6). ISSN 0975-6809.s 1646 - 1653.s doi: 10.1007/s13198-019-00917-3.

Table of Contents

List of Tables

List of Figures

1

Part 1

Thesis summary

CHAPTER 1

1. INTRODUCTION

This chapter gives a short description of the research area; it includes the background, problem definition, purpose, research question, scope and limitations and ends with the structure of the thesis.

1.1. Background

1.1.1. Industry in the Arctic area

Energy is a key element for driving modern industries and people's quality of life. World demand for energy leads industry to harvest energy in complex environment with harsh conditions and sensitive areas, such as the Arctic region – one of the last remaining wild places in the world – with potentially harmful consequences. The Arctic region can be defined geographically by the Arctic Circle, its climate, vegetation and marine boundaries (Figure 1, a)

a b

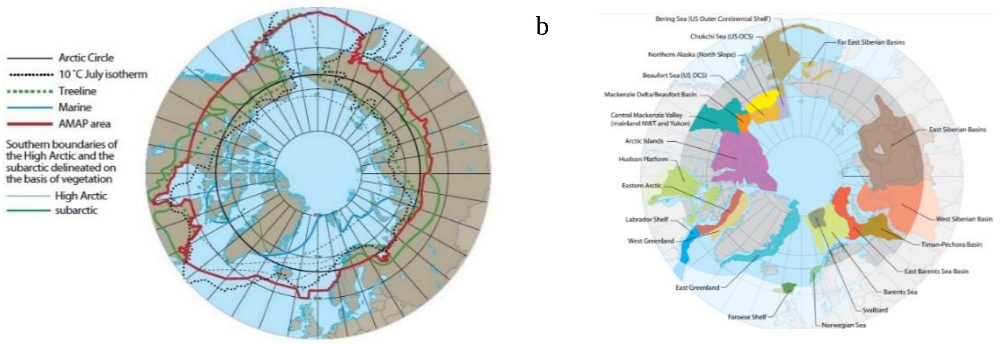


Figure 1. a) The Arctic region and its boundaries (Hansen and Van Oostdam, 2009), Figure 1.b) Major oil and gas provinces (OGP) and basins around the Arctic (AMAP, 2010b).

However, it is often delimited by the Arctic Circle, located at 66º, 32'N latitude (Perry and Andersen, 2012, Murray et al., 1998). The population comprises about four million permanent residents, and eight "Arctic States" have control over the various lands that compose the Arctic region: Canada, Denmark (as the sovereign of self-governing Greenland), Finland, Norway, Sweden, Iceland, Russia, and the United States (Fow, 2011). The Arctic is characterized by its harsh climate, with high variation in temperature and light, polar lows, short summers, large areas of permafrost, and extensive snow and ice cover in winter (ACI, 2005).

According to the United States Geological Survey assessment, the Arctic contains approximately 13 percent (90 billion barrels) of the world's undiscovered conventional oil resources and about 30 percent of its undiscovered conventional natural gas resources (Oil, 2011). Consequently, since the 1960s, when more intensive oil and gas activity started in the Arctic, over 440 exploration wells have been drilled (Council, 2015); currently, the Arctic produces about a tenth of the world's oil and a quarter of its gas (AMAP, 2010a). Given the large undiscovered petroleum resources and the reduced sea ice, increased future oil and gas production in the Arctic can be expected (Peters et al., 2011). (Figure 1, b). Over the past few decades, the increasing trend of melting sea ice in the Arctic has provided increased access and has created new opportunities for economic development within renewable energy, metals and minerals, fisheries, cargo shipping, cruising, subsea telecom cables, ports, pipelines and power grids (Koivurova, 2013, Meier et al., 2014, Serreze and Barry, 2011, Quillérou et al., 2015). (See Figure 2). However, the development of Arctic resources is assumed to be technologically and economically challenging and risky. The Arctic represents a new frontier, where existing technologies are tested to their limits (Kristoffersen and Langhelle, 2017). Studies reveal that, due to low temperatures, sea ice, polar low pressures, poor visibility and seasonal darkness, etc., the Arctic operational conditions have significant effects on the performance of components and industry activities in various ways, including increasing failure rate and repair time, and can cause different types of production losses (Barabadi, 2014, Gao et al., 2010, Trump et al., 2018). Operators in the Arctic also face greater complexity, since the environment is vulnerable, and communication and rescue operations' infrastructure has not been developed significantly.

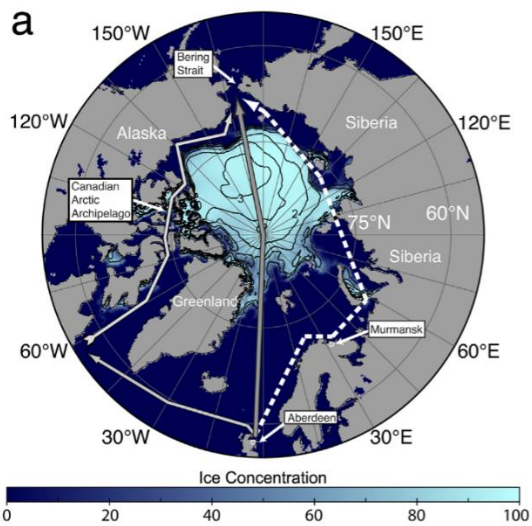


Figure 2. Model 2010–2019 (a) and 2030–2039 (b) sea ice concentration (%, shades of blue) and thickness (labeled contours) during the navigation period (June–October). The Arctic shipping routes are shown schematically: the Northern Sea Route (NSR) (dashed arrow), the North Pole Route (NPR) (DARK-GRAY ARROW), the Northwest passage (NWP) and the Arctic Bridge (AB) (LIGHT-GRAY ARROW) (Aksenov et al., 2017).

1.1.2. The need for a Performance Measurement System (PMS)

The optimal functioning of technical systems involved in the Arctic faces numerous challenges, in order to succeed in a globally competitive market with limited resources (Katic et al., 2011, García-Granero et al., 2018). It depends on the utilization of new knowledge, imagination, creativity and innovations (Zamecnik and Rajnoha, 2015). In this regard, operating more efficiently and effectively, in order to sustain competitiveness, reduce downtimes, costs, wastes, and enhance productivity, quality and safety, has been industries' major concern. The concept of a Performance Measurement System (PMS) is frequently used by industries to achieve such goals which refer to output results obtained from a system that permits evaluation and comparison, relative to past results or other companies (Katic et al., 2011, Franco-Santos et al., 2007). A PMS includes a hierarchical relationship of Performance Indicators (PIs), positioned in a strategic context for deviate detection, measures to describe the status potential, measures to track past achievements and measures to evaluate performance against strategic goals and initiatives (Lebas, 1995, Nanni et al., 1990). It enables decision-making processes to be supported by the gathering, elaborating and analysis of information (Vukšić et al., 2013). The most important reason for implementing a PMS is to quantify the value created by an engineering process or an action. In performance measurement research, phrases like “If you cannot quantify it, you cannot manage it” or “You are what you measure” are commonly heard (Garvin, 1994, Hauser and Katz, 1998). Measures help to identify areas of strengths and weaknesses and to decide on future initiatives, with the goal of improving a company’s performance (O’Neill Jr, 2006). Since the 1880s, different perspectives and PIs have been used within the performance measuring concept, including effectiveness, efficiency, financial, learning, growth, renewal, employee competences, internal and external structure, customer satisfaction, stakeholder contribution, capacity, people, future, etc. The end of the 1980s was a turning point in the performance measurement literature. Markets became competitive, and customers became more demanding, due to the globalization of trade (Kaplan, 1991, Hayes and Abernathy, 1980). This situation led to companies attempting to find more balanced, multi-criteria/indicators and integrated PMS frameworks, considering both financial and non-financial performance perspectives and internal and external performance perspectives. All these frameworks were concerned with what to measure, and they tried to answer the question of how to design a PMS.

1.2. Problem definition

PMSs have been shown to be an essential concept, to improve the efficiency and effectiveness and to support the design, planning and management of a company. PMS needs up-to-date and accurate performance information on its company and business. This performance information needs to be integrated, dynamic and accessible, to assist fast decision-making (Nudurupati et al., 2011). Measure-validation and the reliability of monitoring and analysing the performance of a system in the Arctic depend on two important questions, namely: What needs to be measured and how will it be measured?

An evaluation of the available PMSs shows that there is confusion over terms and criteria for the PIs of companies; various indicators are used for various industries, and the researchers have tried to define indicators in relationship to their area of specialty (Bourne et al., 2003, (Bititci et al., 1997) (Taticchi et al., 2010). These frameworks have been gradually modified and improved, and all have their relative benefits and limitations. Considering the unique and challenging Arctic operational conditions, with strict regulations and requirement for safety and the environment, the designed system or equipment must be available and safe, as well as economically viable. Such systems must be able to minimize environmental pollution and require the minimum quantity of raw material and energy. Without taking these challenges into account, design, maintenance and operation cannot be at an acceptable level of performance (Kumar et al., 2012, Markeset et al., 2015). Operational conditions in the Arctic can increase power losses, life cycle costs and safety hazards. Moreover, the less developed infrastructure in the Arctic creates several challenges, such as limitations to the logistics of supplies, material and personnel required for operation and maintenance activities (FURULY et al., 2013). However, performance terminologies and standards for the Arctic reveal that the PIs measured by industry, though important, are not enough and could still be improved by identifying important indicators, which contribute to a successful PMS in the Arctic. Hence, development and continuous improvement of PMSs and the identification of more PIs for judging performance in the Arctic are critical for industry success.

Moreover, quantification of performance is complex, as it involves various indicators with different perspectives at various hierarchical levels. Finding a proper approach or models to justify the impact of the external environment or factors influencing PIs is being identified as an important challenge in measuring PIs (Kayrbekova et al., 2011, Naseri et al., 2016, Markeset, 2008). The lack of effective information systems, with the correct sources of information and data on performance indicators and suitable statistical models and standard approaches, is a barrier to the successful quantification of PIs (Kennerley and Neely, 2002, Eccles, 1991, Norton and Kaplan, 1999). Operation and maintenance data are often collected from multiple and distributed units in different operational conditions, which can introduce heterogeneity into the data. Part of such heterogeneity can be explained by the observable risk factors, whose values and the way that they can affect the item's PIs are known. However, some factors which may affect PIs are typically unknown (unobserved risk factors), leading to unobserved heterogeneity. Nevertheless, many researchers have ignored the effect of observed and un-observed risk factors, and this may lead to erroneous model selection, as well as wrong conclusions and decisions. The statistics models must be able to quantify risk effect on PIs and must be built based on correct assumptions that reflect the operational conditions. The first element of the Figure 3 ("Why a Performance Measurement System?") includes contributions dealing with what is meant by PMS; these definitions have evolved over time, reflecting the evolution of the concept.

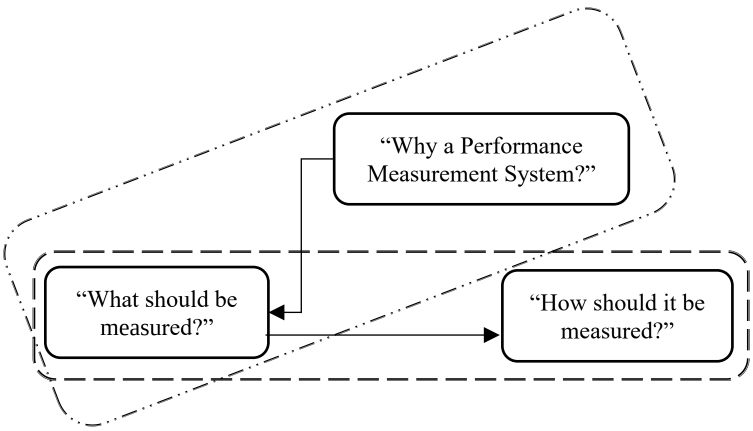


Figure 3. "What should be measured?" and "How should it be measured?" Two separate questions and categories that should arise for designing PMS.

The second element of the figure ("What should be measured?") deals with the appropriate performance indicators/criteria that need to be measured for judging the performance of equipment/systems in the Arctic. The third element ("How should it be measured?") deals with how to measure PIs.

1.3. Purpose and objectives

The purpose of this research is to study, analyse and suggest a methodology for the monitoring and analysis of operation and maintenance, taking into consideration the operational conditions in the Arctic. The main objective of the study is to suggest a PMS for the Arctic and modify the available statistical approach for the prediction of performability, considering the effect of observed and unobserved risk factors. More specifically, the following objectives are determined:

- To review the generic body of literature on performance measurement, to understand key concepts, definitions, aspects of criteria for measuring the performance in a company.
- To identify and discuss appropriate performance indicators/criteria that need to be measured for judging the performance of equipment/systems in the Arctic which contribute to a successful PMS.
- To contribute towards a clarifying vision of PMS for the Arctic.

- To discuss the effect of operational conditions on the performance indicators of systems/equipment in the Arctic.
- To develop models to quantify PIs, considering the observed and un-observed risk factors.

The models and framework developed in this work can be employed in facilities and technology activities to analyse the impact of operational conditions on the performance of systems/equipment and to assist calculations and predictions.

1.4. Research questions

To fulfil the above purpose, the following research questions (RQs) have been formulated:

- RQ1: How the concept of PMS has evolved over time and how it can improve the performability of a system?
- RO2: Which indicators/criteria should be considered to be measured for judging the performance of equipment/systems in the Arctic, and how can operational conditions affect the PIs of systems/equipment?
- RQ3: How to estimate the effect of operation conditions (observed and unobserved risk factors) on safety, reliability and maintainability performance of an item?

1.5. Scope and limitations

The scope of this research includes the operation and maintenance performance of equipment/systems in complex environment with harsh conditions and sensitive areas, such as the Arctic region. The focus of most of the available studies was on designing a PMS, with few studies illustrating the issues involved in the quantifying of the PIs. Hence, this study limited its focus more to illustrating the issues involved in quantifying PIs, such as safety, reliability, and maintainability performance, and not to the general concept.

1.6. Linkage of research questions and appended papers

The linkage between the research questions and the appended papers is shown in Table 1.

Table 1. Linkage between the research questions and the appended papers

	Paper A	Paper B	Paper C
RQ1	×		
RQ2	×	×	×
RQ3	×	×	×

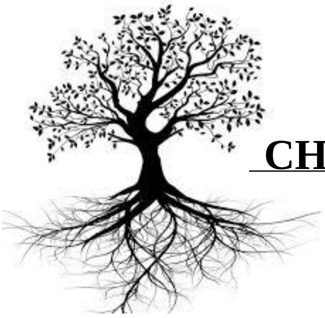
1.7. Structure of the thesis

This thesis consists of the research summary and three appended journal papers, in two parts. The first part consists of six chapters that give an introductory summary of the study and describe the relevant theoretical background to this research work, the literature review, analysis, results, and discussions, as well as the conclusions of the work.

The first chapter provides background information, the problem description and justification of the study, research purpose and questions.

The scientific and systematic approach followed in this study is described in Chapter 2, while Chapter 3 summarizes the appended papers. The fourth chapter presents the results of the research study. The areas of discussions focus on the stated research objectives. Finally, the contributions of the research work and suggestions for future work are presented in the fifth chapter.

The second part consists of three appended papers. Paper A concerns the development of a model for improving safety performance measurement. Paper B is a study of the reliability performance indicator, and Paper C concerns the maintainability performance indicator.



CHAPTER 2
2. RESEARCH METHODOLOGY

This chapter presents the research methodology. This includes the research approach, process, and strategy, followed by data collection activities and a discussion of the quality of the research.

Research is a common name for the search for knowledge. A systematic approach to solving a research problem is termed the research methodology (Kothari, 2004, Rajasekar et al., 2006). The research methodology selected is the link between thinking and evidence (Sumser, 2001). There are two types of research, basic and applied research. Basic research is the search for knowledge and understanding of a topic, and applied research is the research to address a specific concern or problem or to offer solutions. Considering the essence of the present research, it can be classified in the applied research group. Based on what the research is trying to accomplish, the purpose can be classified into three groups: describing a phenomenon (descriptive research), exploring a new topic (exploratory research), or explaining why something occurs (explanatory research). In this study, exploratory research is intended to generate new knowledge and a model regarding the effect of operational conditions on safety, reliability, and maintainability performance. Correspondingly, based on the research questions in this study, it can be concluded that this research can be grouped in the exploratory and descriptive classes, because it explores a new topic and describes a phenomenon (Neuman, 2007).

Research methodology has many dimensions, and the research methodology of the thesis can be explained in five main dimensions that are used to achieve the research aim and objectives of the thesis:

- The research approach
- The research process
- The technique for data collection
- The data analysis techniques
- The reliability and validity of the research

2.1. Research approach

The research approach can be classified into three main topics: deductive, inductive and abductive (Neuman, 2007). The deductive approach starts with an abstract and a logical relationship among concepts and then ends with empirical evidence. The aim of the deductive approach is to test theories. In an inductive approach, the research begins with observations and moves toward more abstract generalizations and ideas. The purpose of the inductive approach is to gain descriptions of characteristics and patterns, and the approach begins with collecting data on characteristics or patterns and ends with relating these to the research questions. Finally, abduction is a combination of deduction and induction (Ghodrati, 2005, Neuman, 2007, Alvarsson and Sköldberg, 1994).

In this study, the research started as a deductive approach, with a literature study, to gain a deeper understanding of how the concept PMS can improve the performability of a system, available statistical models, and the operational and technological challenges of the Arctic.

The result of the literature study shows that the existing frameworks and methods should be improved, to be more suitable for the prediction of safety, reliability, and maintainability performance in the Arctic region. Thereafter, a framework suggested for the performability of a system in the Arctic and models were improved, in order to analyse the historical data. The improved models were then applied in an inductive approach by studying the empirical data. Thereafter, the validity of the models was carried out, and conclusions were drawn, based on the experience gained from empirical case studies. As the research study started with a deductive approach, followed up by an inductive approach, it can be characterized as having an abductive research approach.

The research approach can also be classified as quantitative, qualitative and/or mixed (see e.g. (Ghodrati, 2005, Neuman, 2007, Sullivan, 2001). This research can be classified as using quantitative methods, because the data used were mostly statistical data collected from field data, databases, reports and interviews. Moreover, the outcomes were used to mention a final course of action.

2.2. Research process

A research process gives a series of action steps, along with the interconnections and sequencing of the steps, to effectively achieve the aims of a research study (Kothari, 2009). The research process mostly depends on which type of data and information the researcher is looking for Yin (2003). In the case of collecting and analysing empirical evidence, Yin (2003) describes five different research processes to apply; these include experiment, survey, archival analysis, history and case study. Where it is a case study, experiment and survey processes usually refer to the present situation, and archival analysis and history processes refer to the past conditions of the case under study (Yin, 2003).

The experimental research process implements principles existing in natural science that can be conducted in real life or laboratories. In survey research, during a short time period, the researcher asks numerous questions to people and then summarizes the answers in graphs or tables (Neuman, 2007, Harrison et al., 2017). Based on different types of research processes and considering the research questions and the approach of this study, it can be classified into the case study research process group. In case study research, the researcher examines, in detail, the features of a case or multiple cases over time (Neuman, 2007, Harrison et al., 2017). The study implemented scientific principles in the real life of a system/component, to understand further features and behaviours.

2.3. Data collection and analysis

The process of transforming data into useful information for decision-making support is an essential aspect of a piece of research. There are six main sources of information: documentation, archival data, interviews, direct observations, participant-observations and physical artefacts that are applicable for a case study (Yin, 2017, Marshall and Rossman, 2014). Moreover, data collection techniques can be classified into (Neuman, 2007):

- Qualitative (i.e., expressed as words, visual images, sounds, or objects); these include field research and historical-comparative research

- Quantitative (i.e., expressed as numbers); these include experiments, surveys, content analyses, and existing statistics

Based on the objectives of the study, historical data from offshore oil and gas and mining industries have been used, in order to study the effect of operational conditions on the performance of equipment (Table 2). The required data of this study were collected from expert opinions, design information, meteorological data, and reports of the maintenance, repair and inventory crew and the operators of machines. The main sources for collecting data and summaries of the case studies in the papers are listed in Table 2. The documentation consisted of different descriptions, policies, and procedures pertaining to maintenance programs and failure consequence categories. In addition, a survey was performed, in order to elicit expert judgments and to estimate an Icing Risk Index for performability.

Table 2. Data used in the study.

Paper No.	Industry	Period	Data type	Covariates	Source of evidence	RQs addressed
1	Oil and gas	-	Item geometry, design information, expert opinions, available experience related to the site, and meteorological data	-	Experiments, survey	RQ 1, 2
2	Mining	18 months	Time to repairs (TTRs)	Working shift, Weather condition, Precipitation, Temperature, Involved maintenance crew	Existing statistics: documentation, archival records, direct observation, and interview	RQ 2, 3
3	Mining	18 months	Time between failures (TBFs)	Company, Working shift, Weather condition, Road condition, Rock fragmentation	Existing statistics: documentation, archival records, direct observation, and interview	RQ 2, 3

To study the influence of the operating environment on the performance of equipment (paper B and paper C), the covariates were classified into observed and unobserved groups. The observed covariates whose effects on the failure and repair processes are known and their associated levels are recorded with the failure and repair data, such as "working shift", "weather condition", "precipitation", etc., extracted from different sources of evidence (Table 2 and Table 3). They can be time-dependent or time-independent. The Cox regression model family, such as the proportional hazards model (PHM), proportional repair model (PRM) and its extension, is the most dominant statistical approach for capturing the effect of observed covariates on the reliability and maintainability performance of an item.

Table 3. Sample data of reliability and maintainability

Failure and repair No.	TBFs and TTRs	Status	Observed covariates		
			Working shift	Weather condition	Precipitation
1	24	1	A	1	12
2	13	1	B	2	32

Unobserved covariates that cause to unobserved heterogeneity can be calculated by a mixture frailty model (Paper B). This model consists of three multiplicative factors: i) the baseline failure $(\lambda_0(t))$ dependent on time alone (TBFs), which is modelled by using appropriate distributions, ii) a positive multiplicative factor to describe the function of time-independent and time-dependent observed covariates, $(\psi(z, z(t); \eta; \delta))$ and iii) a positive multiplicative factor α_j that represents the effect of unobserved covariates.

$$\lambda_j(t; z; z(t); \alpha) = \alpha_j . \lambda_0(t)\psi(z, z(t); \eta; \delta) \tag{1}$$

The suggested framework and models present in paper B is based on the mixed proportional hazards model that provide appropriate tools for modeling observed and unobserved heterogeneity among failure data. The first step in analyzing the collected failure data of a repairable system in this study is to check the trend of the failure data. In the next step, the time dependency of observed covariates should be checked. Later, the failure data need to be investigated for unobserved covariates. Data sets without unobserved heterogeneity are analyzed using the classical proportional hazards model, including the proportional hazards model (when all observed covariates are time-independent) and the extension of the proportional hazards model (in the presence of time-dependent covariates). Moreover, data sets with unobserved heterogeneity are analyzed using the mixed proportional hazards model family.

In Paper C a mixture frailty model for maintainability is developed that are able to model the effect of observed and unobserved covariates, as well as identifying different repair processes in the repair dataset. When each repair process is regarded as an independent repair mode with a repair distribution in the presence of some specific observed or un-observed risk factors, then a mixture frailty model (MFM) can be used to predict maintainability more precisely. When A (effect of unobserved covariates) is distributed as gamma with mean one and variance θ, then maintainability becomes:

$$\mathrm{M}_\theta(t) = 1 - \left[1 - \theta ln\left\{1 - \mathrm{M}\left(t, z_i, z_j(t)\right)\right\}\right]^{-\frac{1}{\theta}} \tag{2}$$

If the observed covariate follows the exponential function in the presence of W time-independent observed covariates and M time-dependent observed covariates $\mathrm{M}\left(t, z_i, z_j(t)\right)$ can be written as:

$$\mathrm{M}\left(t, z_i, z_j(t)\right) = 1 - [1 - \mathrm{M}_0(\mathrm{t})]^{\exp\left[\sum_{i=1}^{W} p_i z_i + \sum_{j=1}^{M} \delta_j z_j(t)\right]} \tag{3}$$

where $\mathrm{M}_0(\mathrm{t})$is the baseline maintainability function dependent only on the time, as follows:

$$\mathrm{M}_0(t) = 1 - exp\left[-\int_0^t \mu_0(t')dt'\right] \tag{4}$$

Moreover, in order to quantify the effect of different types of ice on performability, the concept of an icing effect index (IEI) on performability is developed (Paper A). An IEI on performability can represent the consequences of icing on equipment. Thereafter, considering the probability of ice accretion and IEI on performability, the icing risk index for performability (IRIP) can be quantified as:

$$\mathrm{IRILP} = \mathrm{IEIP} \times \mathrm{POI} \tag{5}$$

In this paper, we assumed that Performability has five principal indicators (reliability, maintainability, quality, safety, and sustainability) and two dependent indicators (dependability and survivability) (Figure 4). Quantifying the IEIP is a bottom-to top process that starts by quantifying the IEI on survivability (IEIS).

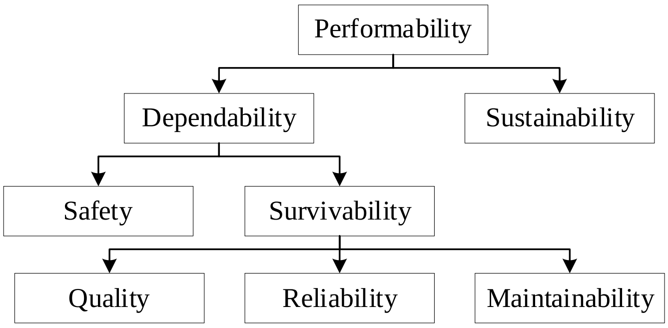


Figure 4. Performability concept (Misra, 2008b)

IEI on survivability $(IEIS) = I_q^{\alpha_q} \times I_r^{\alpha_r} \times I_m^{\alpha_m}$ (6)

IEI on dependability $\mathrm{IEID} = \mathrm{IEIS}^{\beta_{\mathrm{sur}}} \times \mathrm{I}_{\mathrm{S}}^{\beta_{\mathrm{s}}}$ (7)

IEI on performability $\mathrm{IEIP} = \mathrm{IEID}^{\gamma_{\mathrm{Dep}}} \times \mathrm{I}_{\mathrm{su}}^{\gamma_{\mathrm{su}}}$ (8)

where $\mathrm{I_q}, \mathrm{I_r}, \mathrm{I_m}$, IEIS, $IEID, I_s$ and I_{su} are the IEIs for the quality, reliability, maintainability, survivability, dependability, safety and sustainability and the parameters $\mathrm{\alpha_q}, \mathrm{\alpha_r}, \mathrm{\alpha_m}, \beta_{sur}, \beta_s, \gamma_{Dep}$ and γ_s are the weight vectors of quality, reliability, maintainability, survivability, dependability, safety and sustainability, respectively.

In the cases of free-fall lifeboats and helicopter landing pads in paper A a group of 13 experts has been asked to identify IEIs. Table 4 shows the IEIs of different types of ice on the performability indicators. IEIs will range from1to10; where1shows no icing effect on the indicators of the selected item and10 shows a very high effect.

Table 4. Final IEs on performability attributes for FFL and HLP

Selected item	Ice type	Quality	Reliability	Maintainability	Safety	Sustainability
Free-Fall lifeboat (FFL)	Sea spray icing	7	8	9	8	7
	Snow	7	7	5	7	3
	Glaze	4	7	5	6	2
	Rime	3	4	4	4	2
	Frost	2	2	3	2	1
	Sleet	1	1	2	2	1
Helicopter landing pad (HLP)	Sea spray icing	7	9	9	9	7
	Snow	6	7	8	7	6
	Glaze	4	6	8	6	5
	Rime	2	4	5	5	4
	Frost	1	3	5	4	3
	Sleet	1	2	2	2	1

Firstly, a group of 13 experts is selected based on a set of criteria (e.g., knowledge on cold-climate technology, icing phenomenon, and the equipment unit or the case of interest); once experts are selected, expert data are elicited through individual interview. In this regard, experts are asked to provide their opinions in the form of single-point estimates, probability distribution, parameters of a distribution, intervals, etc. After collecting the data, Cronbach's alpha coefficient is applied to measure the reliability of responses to the questionnaire. The value is $\alpha \geq 0.86$ and $\alpha \geq 0.82$ for FFL and HLP, respectively; this means that the reliability of responses is excellent. Once expert opinions are elicited, the geometrical mean is used as aggregation method to combine expert data and thus to obtain a single solution to extract the final IEI-based questionnaires.

Moreover Table 5 shows weight ranking for different performability indicators that were available in the design information of free-fall lifeboats and helicopter landing pad and we calculated the rank values based on a method that was developed by (Jiang and Ji, 2002).

Table 5. Attributes' weight ranking of FFL and HLP

Item	Survivability	Dependability	Performability
Free-Fall lifeboat (FFL)	$\alpha_r \geq \alpha_q \geq \alpha_m$	$\beta_s \geq \beta_{sur}$	$\gamma_{Dep} \geq \gamma_s$
Helicopter landing pad (HLP)	$\alpha_r \geq \alpha_m \geq \alpha_q$	$\beta_s \geq \beta_{sur}$	$\gamma_{Dep} \geq \gamma_s$

2.4. Reliability and validity of the research

According to Hannula (2002), measures produced by quantitative research should try to fulfil these criteria: validity, reliability, and relevance. Validity is the ability of a measure to measure what it is intended to measure. Reliability refers to the consistency of measurement results, including characteristics such as accuracy and precision. Relevance is the value and usefulness of the measure to its users (Hannula, 2002). In this study, the empirical data are used as case studies for safety, reliability, and maintainability performance analysis.

In order to assure the reliability of the study, the source of data (reports) is available for recollection and reanalysis. The data gathering processes are carried out as per established standards and methodology described in the literature. Furthermore, the theoretical concepts are explained, and the analysis approach is described in each paper with details, in order to guide other researchers. The application of the developed models for improving the performance of the systems in this study is clearly discussed especially through the conducting of illustrative case studies and are published in peer-reviewed journals.

CHAPTER 3

3. SUMMARY OF THE PAPERS

In this chapter, the appended papers and their logical relations and contributions are discussed.

Paper A

Title: Designing for performability: An icing risk index for Arctic offshore.

Purpose: To review and discuss the effects of different types of ice accretion on the performability of Arctic offshore facilities and to develop an Icing Risk Index.

Findings: The paper finds that designing for performability is an effective way to meet the design goal for a complex operational condition such as the Arctic region. Designing for performability in the harsh, sensitive, and remote Arctic area is a challenging task. It requires a range of tools to be employed and is dependent on a large amount of data and information. However, taking into consideration the unique Arctic operational conditions, most of the available tools need to be modified, and, in some cases, new tools should be developed. Ice accretion is one of the most hazardous operational conditions in cold regions. Hence, this study has reviewed the effect of different types of ice accretion on performability indicators and then developed an Icing Risk Index, which can be used to quantify the effect of different types of ice on performability. Further, its application is shown by means of a case study. The case study demonstrates how the Icing Risk Index can be applied to Arctic offshore facilities.

Paper B

Title: Observed and unobserved heterogeneity in failure data analysis: A case study.

Purpose: To develop a framework for reliability analysis in the presence of unobserved and observed covariates.

Findings: In many reliability studies, data sets are assumed to be homogeneous, with the failure data being independent and identically distributed. However, in reality, failure data are often collected from multiple and distributed units in different operational conditions, which can introduce heterogeneity into the data. Part of such heterogeneity can be explained and isolated by the observable covariates, whose values and the way that they can affect the item’s reliability are known. However, some factors, which may affect the reliability of the item, are typically unknown, leading to unobserved heterogeneity. These factors are categorized as unobserved covariates. In most reliability studies, the effect of unobserved covariates is neglected. This may lead to erroneous model selection for the time to failure of the item, as well as wrong conclusions and decisions.

In this paper, the required statistical tests and available models for observed or unobserved heterogeneity in the reliability analysis of failure data are reviewed, and then a systematic framework is developed to facilitate the application of these models. The framework is based on the mixed proportional hazards model and its extension, which provides an appropriate tool for modelling observed and unobserved heterogeneity under the different types of maintenance strategies. Further, its application is then shown by a case study.

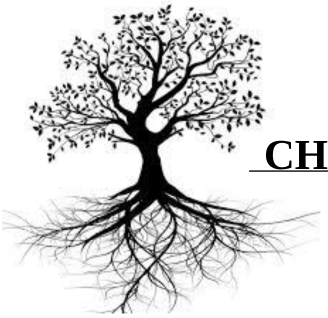
The result of the case study shows that, ignoring the effect of unobserved covariates, and using a Proportional Hazards Model (PHM) instead of a Mixed Proportional Hazards Model (MPHM), will underestimate or overestimate the effect of covariates. Hence, for any decisions on the operation and maintenance strategy, the effect of unobserved covariates should be considered.

Paper C

Title: A mixture frailty model for maintainability analysis of mechanical components: A case study.

Purpose: To evaluate the effect of observed and unobserved covariates on the maintainability of a component or a system, as well as identifying different repair processes in the repair dataset.

Findings: Existing studies regarding the maintainability analysis of historical repair data have simplified their analysis by considering a complex system as a single item. In these studies, the assumption is that all repair data represent an identical repair process for the item. However, these failure modes may have completely different repair processes and resources. In practice, mechanical systems are composed of multiple parts, with various failure mechanisms, which need different repair processes (repair modes) to return to the operational phase. These studies have viewed the historical data as a black box, with no information regarding the repair process and its operational conditions. Moreover, the relationship between the elements of maintainability is complex and mediated by many influence factors, such as ambient temperature, human factors, and dissimilarity in personality or skill level between maintenance crews, etc. As a main part of maintainability, logistics and spare parts constitute a complex activity that is time- and location-dependent. In this paper, the application of a Mixture Frailty Model (MFM) for maintainability analysis is discussed. MFM has the ability to model the effect of observed and unobserved covariates on maintainability. Moreover, it can capture different repair processes in a single database, by the use of a convex combination of their associated distributions. In the second part of the paper, the application of the developed model is illustrated by investigating the effect of observed and unobserved covariates on the maintainability of trucks at a copper mine. The results of analysis show that most identified observed covariates and unobserved covariate(s) have a significant effect on the maintainability of trucks.



CHAPTER 4

4. RESULTS AND DISCUSSION

This chapter discusses and presents the results of the research study. The areas of discussions focus on the stated research questions.

4.1. First research question

RQ1. How the concept of PMS has evolved over time and how it can improve the performability of a system?

The word “performance” comes from the French word *Parfournir*, meaning “to bring through, to carry out”. Performance is an act of performing, implementing, achieving and fulfilling the given tasks that need to be measured (Naz et al., 2016). The field of performance measurement has evolved over a long period and has been defined and redefined from different perspectives in different industries. For example, from an operations perspective, Bourne et al. (2003) defined it as the set of multi-dimensional performance measures (financial/non-financial and internal/external) that quantify the performance that has been achieved. Bititci et al. describe it as the reporting process that gives feedback to employees on the outcome of actions (Bititci et al., 1997). Neely et al. defined PMS as a balanced and dynamic system that enables the support of decision-making processes by gathering, elaborating and analysing information (Neely et al., 2002). As pointed out by Taticchi et al. the concept of ‘dynamicity’ can be referred to the need to develop a system that continuously monitors the internal and external context and reviews goals and priorities (Taticchi et al., 2010). Meanwhile, ‘balance’ refers to the need to use different indicators and perspectives that are tied together, giving a holistic view of the organization (Kaplan and Norton, 1996). Hence, computerization is a technique to deliver the dynamism of performance measurement, to continually capture, store, measure, interpret and visualize data and information (Srimai et al., 2011). Moreover from a management accounting perspective, measuring performance provides the company with the ability to check its position (to compare positions or monitor progress), communicate its position (to communicate performance internally and with the regulator), to confirm its priorities (to manage cost and actions), and to compel progress (as a means of motivation) (Neely, 1998).

A PMS framework includes the hierarchical relationship of performance indicators (PIs), positioned in a strategic context for the detection of deviations, measures to track past achievements, and measures to describe the status potential and evaluate performance against strategic goals and initiatives (Lebas, 1995). According to Table 6, Since the 1880s, different perspectives and performance indicators have been used within the performance measuring concept, including effectiveness, efficiency, financial, learning perspective, growth, renewal, employee competences, internal and external structures, stakeholder satisfaction, stakeholder contribution, capacity, people, future, etc.

Ghalayini and Noble (1996) believed that the literature concerning performance measurement evolved through two phases. The first phase, cost accounting orientation, started in the late 1880s and is known as the traditional phase. The second phase started after 1980 and attempted to present a balanced and integrated view of PMS. The cost accounting orientation approach tries to quantify performance and other improvement efforts in financial terms (Ghalayini and Noble, 1996). Financial measures, such as Return On Investment (ROI), Return On Sales (ROS), revenue per employee, revenue per unit production, cost variance analysis, standard costing and flexible budgets are some of the techniques that were used to measure performance in that era (Bourne et al., 2003, Khan and Shah, 2011). The traditional phase is criticized for ignoring clients and their needs (Ghalayini and Noble, 1996); it was internally rather than externally focused, backward-looking and historically focused. Therefore, these were not predictive measures, and the era is also criticized for not providing adequate information for a productivity measurement (Hayes and Abernathy, 1980, Kaplan, 2005). Due to these shortcomings in traditional measures, Nakajima (1988) introduced Overall Equipment Effectiveness (OEE). OEE is defined as a measure of total equipment performance, and it categorizes major losses or reasons for poor performance (Muchiri and Pintelon, 2008). OEE is a three-part analysis tool for equipment performance, based on its availability and performance and the quality rate of the output. The goal was to achieve zero breakdown and zero defects related to equipment. OEE has evolved to include other production losses that were not originally included (Nakajima, 1988). This has led to the development of new terminologies like Total Equipment Effectiveness Performance (TEEP), Production Equipment Effectiveness (PEE), Overall Plant Effectiveness (OPE), Overall Asset Effectiveness (OAE), and Overall Factory Effectiveness (OFE). The difference between these terminologies is based on the type of production losses, losses due to external and internal reasons and levels of effectiveness measurement, namely, equipment-level effectiveness, operational-level effectiveness and business-level effectiveness (Muchiri and Pintelon, 2008, Ljungberg, 1998). However, OEE and its measures – availability, performance speed, and quality rate – only reflect the internal effectiveness of a system and financial performance, while external effectiveness, which is characterized by customer satisfaction and measures that have a long-term effect on a company’s profitability, is missing. The end of the1980s was a turning point in the performance measurement literature, as it marked the beginning of the second phase. Markets became competitive, and customers were more demanding, due to the globalization of trade and the emergence of a world economy (Hayes and Abernathy, 1980, Kaplan, 1984). This situation led to companies attempting to find more balanced, multi-criteria and integrated PM frameworks, considering both financial and non-financial performance perspectives and internal and external performance perspectives. This trend has blended with established social and environment accounting. Consequently, it had led to the development of a number of performance measurement systems, since the 1990s, to focus on the customer’s and the stakeholder’s requirements, rather than only reflecting the shareholder’s economic profits (Garengo et al., 2005). Some of the most well-known and widely cited performance measurement systems are: Performance Measurement Matrix (Bourne et al., 2000), Balanced Scorecard (Kaplan and Norton, 1992), SMART (Lynch and Cross, 1991), AMBITE Performance Measurement Cube (Bradley, 1996), Quality Management Excellence Model (EFQM) (Lin and Shen, 2007), Performance Prism (Neely et al., 2002), Dynamic Multi-Dimensional Performance Framework (DMP) (Maltz et al., 2003), the Results and Determinants Framework (RDF) (Fitzgerald et al., 1991), Integrated Dynamic PMS (Ghalayini et al., 1997), and QUEST (Abran and Buglione, 2003). Table 6 indicate these measures and their criteria are of the business environment existing at that time.

There are standards and performance measurements that have been developed for operation and maintenance in the Arctic. DNV-OS-A201 (Gudmestad, 2010) provides general principles for the preparation of mobile units and offshore installations for intended operations in cold-climate conditions. This is provided for by setting functional requirements for functions, systems, and equipment, considered important to safety, which are intended to be in operation in cold-climate conditions.

Table 6. Balance PMS frameworks with indicators in during the second phase and their development over the years.

Year	Framework.	Indicator/Criteria
1989	Performance measurement matrix	Cost factors, Non-cost factors, External factors, Internal factors
1991	SMART Pyramid (performance pyramid)	Quality, Delivery, Process time, Cost, Customer satisfaction, Flexibility, Productivity, Marketing measures, Financial measures
1991	Results and determinants matrix	Financial performance, Competitiveness, Quality, Flexibility, Resource utilization, Innovation
2002	Performance prism	Stakeholder satisfaction, Strategies, Processes, Capabilities, Stakeholder contribution
1992	AMBITE performance measurement cube	Time, Cost, Quality, Flexibility, Environment
2002	Quality Management Excellence Model (EFQM)	Leadership, People, Policy strategy, Partnership & resources, Processes, People results, Customer results, Impact on society results and Business results
1992	Balanced Score Card (BSC)	Financial, Customer, Internal process, Learning and growth
1997	Integrated Dynamics PM System	Timeliness, Finance, Customer satisfaction, Human factors, Quality, Flexibility
1998	NORSOK Z-016	Reliability, Maintainability, Supportability
2001	Performance prism	Stakeholder satisfaction, Strategies, Processes, Capabilities, Stakeholder contribution
2003	Dynamic Multi-dimensional Performance framework (DMP)	Financial, Market, Process, People, Future
2003	BCS of Advanced Information. Services Inc. (AISBSC)	Financial perspective, Customer perspective, People, Infrastructures, and innovation
2008	System performability	Survivability, Dependability, Sustainability
2008	ISO 20815	Item availability, Production availability, Deliverability
2010	Production assurance performance	Capacity, Dependability, Customer demand
2010	Production performance	Economical, Functional, HSE

ISO20815 (IOS, 2008) introduced performance measures for production availability, which include availability of the item/system, production availability, and deliverability. ISO/TC 67/SC 8 (Blanchet et al., 2007) includes aspects of offshore petroleum activity, i.e. exploration, drilling, production, transportation, and support activities, to ensure that all oil and gas operations are carried out to an acceptable safety level. Production assurance concept NORSOK Z-016 (Standard, 1998a) is another widely used operational measurement. It was developed in 1998 for the oil and gas industry and is built on reliability, maintainability, and supportability. Later (Barabady et al., 2010) formulated the production assurance performance concept, which is a combination of the capacity performance, customer demand and dependability concepts. Markeset (2008) combined availability performance with functional performance, which is based on the capability, capacity and HSE (Markeset, 2008).

4.2. Second research question

RO2: Which indicators/criteria should be considered to be measured for judging the performance of equipment/systems in the Arctic, and how can operational conditions affect the performance indicators of a systems/equipment?

Translating the company's strategy into concrete performance indicators is one of the most frequent recommendations in the designing of PMSs (Globerson, 1985, Lingle and Schiemann, 1996, Neely and Bourne, 2000). Meanwhile, identifying the different PIs for each critical strategic area, structuring the indicators hierarchically and the ability to quantify the effect of the indicators on the company's performance are the main steps in designing a PMS.

The PMS suggested for the Arctic in this study classifies measures into four interrelated criteria, by which the author means that they reflect the performance of a system in the Arctic, each containing indicators and measures from a distinct perspective (see Figure 5). This includes strategic areas, such as financial or cost-related issues, health safety and environment related issues, processes-related issues, and maintenance task related issues, while at the same time comprising the internal and external aspects. These perspectives are termed as:

- Financial performance
- Safety performance
- Overall equipment effectiveness
- Sustainability performance

The PMS should facilitate the quantification of the relationships between indicators with respect to overall performance. Hence, cause and effect diagrams are created as a discussion tool to structure the indicators and formalize the hierarchical nature of the performance measurement system in the Arctic. Figure 5 illustrates the concept of the performance measuring system in the Arctic and its related indicators.

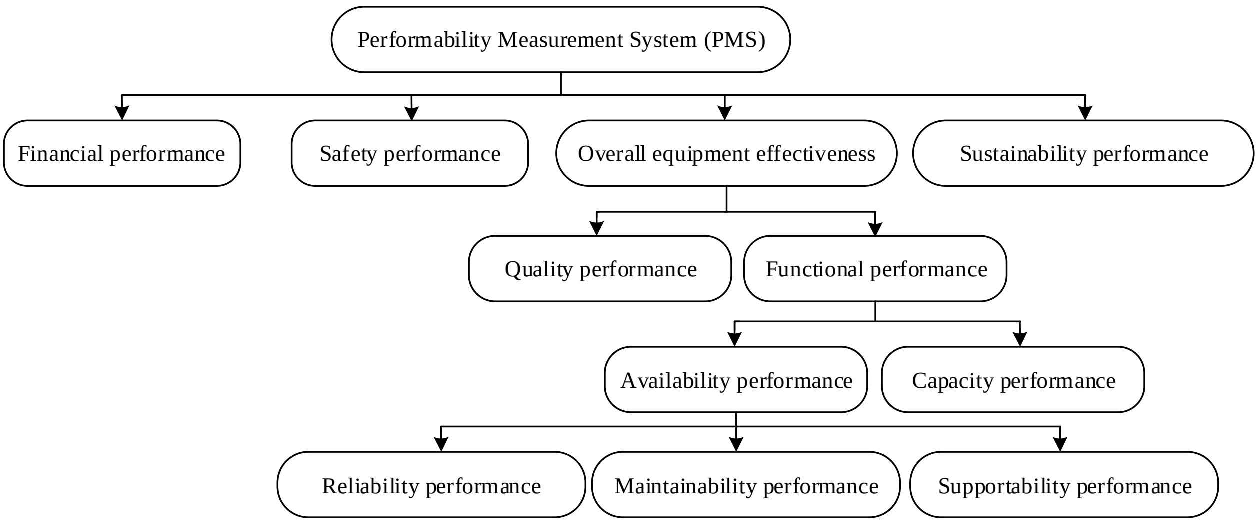


Figure 5. Performance Measuring System for the Arctic

Sustainability performance

The word "sustain" comes from the Latin *Sustenare*, meaning "to hold up" or to support, which has evolved to mean keeping something going or extending its duration (Sutton, 2004). The release of the report, Our Common Future by the World Commission on Environment and Development (WCED) , marked the starting point for the spread of the sustainable development concept (Mikkelsen and Langhelle, 2008). Our Common Future (1987) defined sustainable development as development that meets the needs of the present without compromising the ability of future generations to meet their own needs. It involves an attempt to combine a growing concern for the environment with socio-economic issues and to find a balance between social, environmental, and financial responsibilities (Ferrer-Balas et al., 2008, Brundtland et al., 1987, Butlin, 1989). Some central objectives have been identified, including: conserving and enhancing the resource base, social and economic risks' reduction, merging economics and the environment in decision-making, reorienting technology, reviving growth, education, human health and well-being. Designing for sustainability ensures that the demands of both the customer and society are met, while protecting the ecosystem (Mayyas et al., 2012). It requires integrating environmental issues into a system development process that also meets other demands, such as high quality at least cost (Keoleian and Menerey, 1994). Sustainable development and its principle require that products and systems use minimum material and energy and use non-hazardous materials throughout their entire life cycle. They should be designed for disassembly, designed for remanufacturing and designed for recycling and should be highly recyclable at the end of their life (Mayyas et al., 2012). Therefore, a sustainable company can improve its company's reputation and brand value and increase shareholder value or cost savings, by minimizing the use of material and energy. Moreover, sales may increase or customer loyalty may be strengthened, as there is a growing number of people who prioritize environmentally friendly products and services (Jan and Petra, 2016, Hopkins, 2002).

Implemented products and systems, especially in the Arctic, should comply with the principles of sustainability, to increase energy and material efficiencies, preserve ecosystem integrity, and promote human health, which in turn result in minimum life-cycle costs (Hallstedt et al., 2010). Without sustainability analysis, an overall performance evaluation cannot be comprehensive, particularly in the Arctic with its strict regulations and requirements for safety and the environment. A company needs environmental and social capital – alongside economic capital – to create value in the future (Jan and Petra, 2016). The social impact assessment should include the impact the company has on the local community and how the company contributes to the better health, education and safety of its employees and the local community. Industry activities in the Arctic create different hazards to the well-being and social cohesion of local communities, by exposure to noxious pollutants, as well as economic issues (Trump et al., 2018). As an example, the effect of icing and low temperatures in the Arctic on sustainability can be due to an increase in energy consumption, in the use of materials and in the use of processes and products that are used for ice protection and heating. De-icing technologies with a high consumption of energy have a negative impact on the sensitive environment and wilderness in the Arctic. The large power demand of offshore installations in the Arctic area is, in most cases, covered by their gas, and greenhouse gas emissions from power production are high. Moreover, the use of hazardous chemical ice protection causes degradation of environmental quality, increases produced waste and has serious environmental consequences (Shi et al., 2013).

<u>Safety performance</u>

The definition of "safety" can be the condition of being protected against financial, physical, political, social, educational, emotional, occupational, psychological, or any other types of consequences arising from accidents, harm, failure, damage, error, or any other event that could be undesirable (Misra, 2008c, Misra, 2008a). It is recognized in the literature that engineering products and systems can cause hazards during operation or maintenance, if they fail. They also generate financial losses, due to the disruptions in industrial processes, damage the production machinery, and harm the firm's reputation (Bottani et al., 2009). The prevention of an accident requires excellence in performance, which leads to reducing the chances of failure and the associated risk. Improving equipment and operational safety performance leads to eliminating or reducing the possibility of hazards (Vinnem, 2010). Hence, the design, development, manufacture, and maintenance of engineering products must strive for high safety performance and to reduce the probability of harmful consequences from flammable, toxic, and explosive hazards (Sultana et al., 2019). However, there is a balance to be struck between safety and the cost of achieving it (Misra, 2008c, O'Connor, 2008).

Working in a cold climate such as the Arctic can be dangerous for personnel. Yun and Marsden (2010) showed that, depending on Escape, Evacuation, and Rescue (EER) strategies for Arctic offshore facilities, the probability of success could fall from 90% in June to 50% in January (Yun and Marsden, 2010). At very low temperatures, electrical insulation starts to crack and exposes the conductors to the environment, and this creates a serious hazard for personnel. Low temperatures generate static electricity that destroys computers, making data unreliable. Moreover, wet snow or glaze causes slippery surfaces on handrails, ladders, decks, etc., constituting an important personnel safety hazard. Snow accumulation on valves inhibits manual operation and the ability to see position indications (Ryerson, 2011). Saltwater ice on antennas bridges insulators, causing arcing and loss of communication. Most researchers agree that the greatest hazard to infrastructure safety is sea spray-created superstructure icing. Large ice accretions can threaten the stability and integrity of offshore production facilities. The high weight caused by sea spray accumulation is an issue for buoyancy and stability, and can cause platform sinking; in addition, icing increases the wind resistance of the superstructure. Moreover, sea spray icing can cover boats, lifesaving apparatus, deck firefighting equipment, all of which are critical (Jones and Andreas, 2012, Orimolade et al., 2017).

Financial performance

Financial performance or cost-effectiveness is an essential element of the performance characteristics by which an item or product is evaluated, particularly in the competitive and uncertain environment with the requirement of environmental protection and social responsibility for present and future generations (Lassala et al., 2017, Taouab and Issor, 2019). This leads to many companies seeking methods to achieve competitive advantages, with respect to cost related to the dynamic environment, while showing concern for the environment and safety (Alsyouf, 2004). Financial performance requires equipment and services to be produced at the lowest possible cost. This refers to the extent to which it is technically feasible to reduce any input without decreasing the output, and without increasing any other input (Commission, 2013). Several Life Cycle Cost LCC tools have been developed, to evaluate the cost-benefit or financial performance of an item, describing the costs of the item from the early planning stages to the end of use and, gives decision-makers information to find the correct balance or best solution in respect of cost and benefit (Misra, 2008c).

Designing for harsh climate condition areas such as the Arctic increases the LCC of a system or equipment and, consequently, increases business risk. This is due to the lack of infrastructure in the Arctic, limitations regarding the logistics of supplies, material, and personnel required for operation and maintenance activities, etc. Moreover, low temperature, sea ice, icebergs and icing, darkness, and polar lows, together with long distances, place demands on the technical solutions used.

Overall equipment effectiveness is classified into two perspectives, including functional and quality performance, by which the author means that they reflect the overall equipment effectiveness (OEE) of a system in the Arctic.

Quality performance

The quality of a product is a measure of its degree of conformity with applicable design specifications and workmanship standards (O'Connor, 2008). If quality can be thought of as the excellence of a product at the time it is delivered to the customer, reliability is used in the engineering context to describe the ability of a product to work without failure during its expected time in use (Misra, 2008b). A product's reliability, therefore, depends upon how well it is designed to withstand the conditions under which it will be used, the quality of manufacture, and how well it is used and maintained.

Quality can be classified into two types (Phadke, 1995): design qualities and manufacturing qualities. In design quality, understanding the environments involved and the stresses that can be applied can prevent wear-out failures and overstress failures. Materials that are common in more benign weather conditions require early assessment for material selection and performance aspects, such as accuracies, efficiency, and operational energy requirements in the design process, to confirm integrity under Arctic conditions over the full life cycle of the facility. Due to the lack of experience and data in the Arctic area, there are significant uncertainties with designing for quality performance, and it is a challenging process. On the other hand, manufacturing qualities pertain to the manufacturing processes used when producing products that incorporate desired design qualities. In the case of machine tools, such qualities would correspond to dimensional variances, surface roughness, and processing accuracy. Production facilities consist of complex subsystems and components and they employ materials, men, and machines. These elements may have inherent variability and attributable variability. Variation in parameters and dimensions leads to weakening, component mismatch, incorrect fits, vibration, etc.

Arctic conditions may provide a situation in which the process is incapable of acceptable operation within the design limits. For example, welds will cool faster in cold weather, which results in increased susceptibility to cracking, both during and after welding. Ice can reduce the quality of communication tools and sensors. For instance, wind vanes and temperature sensors can be affected by ice; studies show that, in icing conditions, wind speed errors can be as high as 30% (Laakso et al., 2003). Moreover, the most important contributor to variability is man himself (Misra, 2008b). Studies show that, in outdoor work in the winter, cold stress frequently reduces working ability by 70% for short periods (Anttonen and Virokannas, 1994). Long periods of exposure to the cold result in decreased cognitive performance, injury, hypothermia, loss of sensitivity, and reduced manual dexterity and grip (Holmér, 1994). These conditions can directly influence the variability of man's decisions or actions.

Capacity performance

Capacity performance can be defined as an item's ability to deliver according to design capacity and/or current demands (requirements/needs) in a fixed period of time with given production resources (Shahidul et al., 2013, Barabady et al., 2010). Full industrial capacity is an attainable level of output that can be achieved under normal input conditions (Klein et al., 1973). Cohen (Cohen, 1993) defines capacity as the ability or aptitude to perform a functional task, with its measure being described by capabilities such as ability, competence, and efficiency. Capacity can affect the efficiency and effectiveness of the operation (Isaza et al., 2015). In the OEE concept, the performance of capacity is a critical factor, neglecting which, particularly in the design phase, may lead to large losses over the operational phase. Capacity performance can be briefly defined as how well the available capacity is used (used capacity/available capacity) (Pomorski, 1997, Cesarotti et al., 2013).The necessity of a capacity performance indicator in the Arctic, particularly in the design period, is revealed.

In the Arctic, used and available capacity are influenced both positively and negatively. On the one hand, consider the same wind speed in two wind farms, one in a cold climate region and the other in a tropical region; since cold winds are denser, the available energy capacity in the Arctic wind farm is higher than in the tropical one. On the other hand, despite the higher capacity in the Arctic, installing, running, and maintaining a wind farm in such an area to use this capacity is much more difficult than in a tropical region. Although nearly a quarter of the global wind energy capacity is operating in cold climate regions, to benefit from the great resource of dense winds, wind turbines in Arctic wind farms are threatened by ice accretion, leading to safety concerns as well as power output reduction, influencing the available capacity, while it negatively affects the used capacity. Consequently, low capacity performance is expected in Arctic wind farms (Stoyanov and Nixon, 2020).

Reliability performance

According to (IEC, 2019), reliability is "the ability of an item to perform a required function over a specified time and under the specified conditions." The main aim of system or equipment reliability is to prevent the failures that cause stoppages and downtime or reduce the adequate functional performance of the system (IEC, 2019). Failures occur when the effect of the applied load (L) is greater than the resistance (R) of the component or material (L > R). The reasons why (L > R) occurs can range from poor design specification and material defects, through to, e.g., fabrication errors, degradation in operation, and poor maintenance. While the resistance R is related to the materials, the design and the in-service condition of the system, the load L can be any type of load: functional, environmental or accidental (Veritas, 2002, Freitag and McFadden, 1997). Environmental conditions can include operational environments as well as preoperational environments, when stresses imposed on parts during manufacturing assembly, inspection, testing, shipping, and installation may have a significant impact on equipment reliability.

In the Arctic, low temperature, icing, and humidity are main concerns that can change the properties of some materials and fluids, increase the failure rate and reduce equipment performance by decreasing its reliability. For example, the icing on structures and equipment will increase wind drag by changing dimensions and weight, shapes, and drag coefficients. Moreover, it can change their natural frequencies, which is a significant factor influencing the dynamic behaviour and control of the systems, leading to increased oscillatory stresses (Ryerson, 2011). For some materials such as plastic, low-temperature stress can change the material's properties and increase its failure rate. The serviceability of rubber components, e.g. tyres, inner tubes, cables, hoses, bushings and seals, is seriously affected by low temperatures (Freitag and McFadden, 1997). Snow infiltration and extreme temperatures lead to condensation in the electronics and, consequently, can lead to electrical failure (Laakso et al., 2003).

Tension forces from ice accretion in some materials, such as steel and cables, increase considerably (Freitag and McFadden, 1997, Misra, 2008d). Low temperature generates static electricity that destroys computers, making data unreliable. Engines and equipment operating during cold weather are subject to higher wear and increased breakage (Dutta, 1988). Very often more than one environmental factor may be acting on systems or equipment. These combined environmental factors may have more adverse effects on reliability than the effects of these individual environments (Misra, 2008d, Fikke et al., 2006).

Maintainability performance

Maintainability, as a characteristic of design and installation, can be defined as the probability that equipment will be retained in or restored to a specified condition within a given period of time (IEC, 2019). Maintainability performance is a design factor that decides the degree to which a product allows safe, accurate, quick and easy replacement of its component parts (Garmabaki et al., 2016, Kumar et al., 2015). Design for maintainability needs to consider human ergonomics, logistic management, design layout, the level of experience and training of the maintenance personnel and so on (Naseri and Barabady, 2016, Knezevic, 1993). The main attributes of maintainability are standardization, interchangeability, troubleshooting, removal/installation, ease of handling, accessibility, safety precautions and skill level (Barabady et al., 2010, Kumar et al., 2012, FURULY et al., 2013). In general, Arctic climate conditions can contribute to changing the maintainability performance of an item, by affecting i) the maintenance and operational crew, ii) components and maintenance tools, and iii) maintenance support. For example, in icing and low-temperature conditions, the maintenance and operational crew should wear warm clothes and gloves, which can increase their body dimensions and reduce mobility and hand dexterity. Longer periods of darkness during winter may cause human depression and reduce the efficiency of workers; the period of brightness during the summer may cause sleep problems (Brunvoll et al., 2010). Icing may change the accessibility of the failed item, by changing its appearance and shape, leading to improper accessibility (Ryerson, 2011). Improper accessibility can increase the access, replacement, and removal time of failed components. Lack of satisfactory access to the equipment requiring maintenance is the most common problem mentioned by maintenance personnel. Moreover, icing may adversely affect helicopter activities, which are important for the logistics of transporting people and materials. Crane, lifting or hoisting provision devices, which are the key elements for carrying out inspections, and maintenance of equipment can be affected. Sensors on test equipment (e.g. temperature sensors, accelerometers, etc.) can be affected by different types of ice, leading to measurement errors in inspections and repair processes (Ryerson, 2011). Lower temperatures may affect the performance of several materials, such as iron and steel, polymers and plastics used in maintenance tools, and they experience embitterment at cold temperatures {Markeset, 2008 #2128.

<u>Supportability performance</u>

An important aspect of customer satisfaction and performability is reducing the downtime and repair costs of the system/equipment. Supportability plays an important role in maintaining a system at a desired level of availability and can be defined as the inherent quality of a system – including design, technical support data, and maintenance procedures – to facilitate the detection, isolation, and timely repair/replacement of system anomalies (Kratz, 2003). There are numerous factors that contribute to the supportability level achieved by each system. These include logistics considerations, such as spare parts, personnel, procedures, test equipment, and integrated tools (Smith and Knezevic, 1996). It is generally accepted that the availability and location of spare parts has a great impact on the supportability of a product/system (Markeset and Kumar, 2005). Thus, supportability concerns are essential for producing efficient and cost-effective systems. In the Arctic area, the remote geographical location from customers and suppliers, the cold and harsh climate and insufficient and inconvenient infrastructure can affect the efficiency and effectiveness of the logistics of required maintenance support services and the delivery of supplies (Barabadi, 2012, Gao and Markeset, 2007).

4.3. Third research question

RQ3: How to estimates the effect of operation conditions (observed and unobserved risk factors) on safety, reliability, and maintainability performance of an item?

Paper A

Ice accretion is one of the most hazardous operational conditions in cold regions. It can increase repair time and failure rate, power losses, life cycle costs, and safety hazards. In order to quantify the effect of different types of ice on performability, the concept of an icing effect index (IEI) on performability is developed. An IEI on performability can represent the consequences of icing on equipment. Thereafter, considering the probability of ice accretion and IEI on performability, the icing risk index for performability (IRIP) can be quantified.

In this study, we assumed that Performability has five principal attributes (reliability, maintainability, quality, safety, and sustainability) and two dependent attributes (dependability and survivability), (see Figure 1 of Paper A). To quantify the IEIP, in the first stage, the icing effect index (IEI) on the principal attributes of performability should be quantified. The IEI will range from 1 to 10, where 1 shows no icing effect on performability attributes and 10 shows a very high effect. In the next stage, considering that performability attributes may place a different importance or weight on the overall performability of the item, the weight of the performability attributes needs to be quantified. The functionality and criticality of the item will decide the weight of the performability attributes. As can be concluded from Figure 1 of paper A, quantifying the IEIP is a bottom-to top process that starts by quantifying the IEI on survivability (IEIS). For this aim, considering the effect of the selected types of ice on survivability attributes (quality, reliability, and maintainability), and the weight vectors of these attributes, the IEIS can be developed as:

$$\mathrm{IEIS} = \mathrm{I}_{\mathrm{q}}^{\alpha_{\mathrm{q}}} \times \mathrm{I}_{\mathrm{r}}^{\alpha_{\mathrm{r}}} \times \mathrm{I}_{\mathrm{m}}^{\alpha_{\mathrm{m}}} \tag{9}$$

where $\mathrm{I}_{\mathrm{q}}, \mathrm{I}_{\mathrm{r}}, \mathrm{I}_{\mathrm{m}}$ are the IEIs for the quality, reliability, and maintainability, respectively. The parameters $\alpha_{\mathrm{q}} \geq 0$, $\alpha_{\mathrm{r}} \geq 0$ and $\alpha_{\mathrm{m}} \geq 0$ are the weight vectors of survivability attributes, where $\alpha_{\mathrm{q}} + \alpha_{\mathrm{r}} + \alpha_{\mathrm{m}} = 1$

After quantification of the IEIS, the IEI for safety needs to be developed. Safety is one of the most sensitive elements of performability, which can easily be affected by icing. Considering the IEI for safetyI_s, and IEIS, the IEI on the dependability index, IEID, of the item can be calculated as:

$$IEID = IEIS^{\beta_{sur}} \times I_S^{\beta_s} \tag{10}$$

Where $\beta_{sur} + \beta_s = 1$ and $\beta_{sur} \geq 0$ and $\beta_s \geq 0$ show the weight vector of survivability and safety on dependability. Finally, the IEIP can be calculated by:

$$IEIP = IEID^{\gamma_{Dep}} \times I_{su}^{\gamma_{su}} \tag{11}$$

Where $\gamma_{Dep} + \gamma_s = 1$ and I_{su} is the IEI on sustainability and $\gamma_{Dep} \geq 0$ and $\gamma_s \geq 0$ show the weight vector of sustainability and dependability on performability.

After estimation of the IEIP, the probability of ice accretion (POI) should be estimated. POI is quantified by physical models, using statistical data or based on expert opinions. Finally, having IEIP and POI, the icing risk index for performability (IRIP) can be calculated by:

$$IRIP = IEIP \times POI \tag{12}$$

A detailed discussion on different steps taken to calculate the IRIP is presented in Paper A (Figure 3, Paper A).

Paper B

The suggested framework in Paper B, (Figure 6) is based on the MPHM and its extensions, which provide an appropriate tool for modelling observed and unobserved heterogeneity under the different types of maintenance strategies.

As this figure shows, in the first step, the context should be established. In this step, all external and internal parameters to be taken into account when analysing failure data and setting the scope and assumptions for the reliability analysis should be defined. In the next step, failure data and all possible observed covariates associated with each failure should be collected. In general, the first step in analysing the collected failure data of a repairable system is to check the trend of the failure data. In the case that the data shown trend Nonhomogeneous Poisson Processes (NHPP) or Trend Renewal Process (TRP) can be used to model the baseline hazard rate. However, when there is no trend in the data, classical distribution, such as Weibull distribution, can be used to model the baseline hazard rate. However, some goodness-of-fit test, such as residual test, should be used to find the best fit distribution for failure data. In the next step, the time dependency of observed covariates should be checked. Later, the failure data need to be investigated for unobserved covariates (For detail see Paper B). Data sets without unobserved heterogeneity will be analysed using the classical PHM, including PHM (when all observed covariates are time-independent) and extension of the PHM (in the presence of time-dependent covariates). Moreover, data sets with unobserved heterogeneity will be analysed using the MPHM family. The application of the proposed methodology is demonstrated by a case study in Paper B.

Paper C

In this study, the application of a mixture frailty model (MFM) for maintainability analysis is discussed. An MFM is an extension of the proportional repair model, where unobserved and observed covariates have a multiplicative effect on the repair rate. If the observed and unobserved covariate follows the exponential function and gamma distribution, then maintainability becomes:

$$\mathrm{M}_{\theta}(t) = 1 - \int_0^{\infty} \left\{1 - \mathrm{M}\left(t, z_i, z_j(t)\right)\right\}^{A} . \frac{A^{\frac{1}{\theta}-1} e^{-\frac{A}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right) \theta^{\frac{1}{\theta}}} dA = 1 - \left[1 - \theta ln\left\{1 - \mathrm{M}\left(t, z_i, z_j(t)\right)\right\}\right]^{-\frac{1}{\theta}} \tag{13}$$

If the observed covariate follows the exponential function in the presence of W time-independent observed covariates and M time-dependent observed covariates, $\mathrm{M}\left(t, z_i, z_j(t)\right)$ can be written as:

$$\mathrm{M}\left(t, z_i, z_j(t)\right) = 1 - [1 - \mathrm{M}_0(\mathrm{t})]^{\exp\left[\sum_{i=1}^{W} p_i z_i + \sum_{j=1}^{M} \delta_j z_j(t)\right]} \tag{14}$$

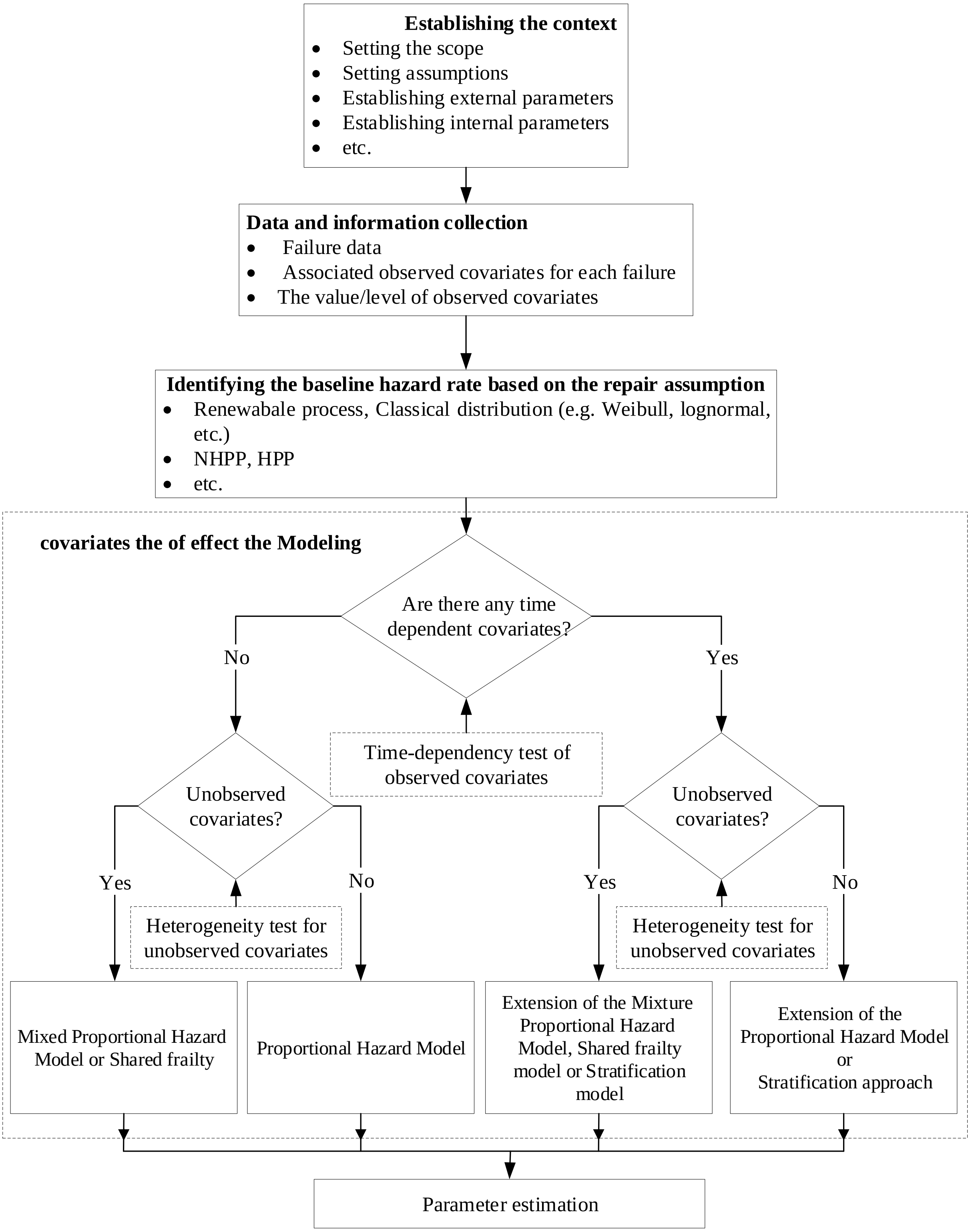


Figure 6. A framework for reliability model selection in the presence of observed and unobserved covariates

where $M_0(t)$ is the baseline maintainability function, dependent only on the time, as follows:

$$\mathrm{M}_0(t) = 1 - exp\left[-\int_0^t \mu_0\left(t'\right) dt'\right] \tag{15}$$

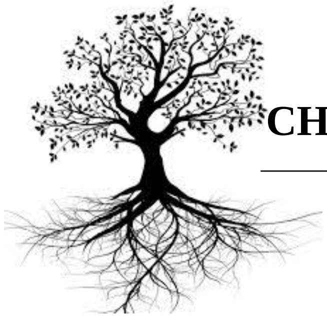
Moreover, if each repair process is regarded as an independent process with an individual, then mixture repair distribution can be used to model the maintainability baseline. Suppose a repair dataset of specific items consists of *N* repair processes, which require different maintenance tasks and repair actions comprised of several subsidiary tasks of unequal frequency and time duration. Under these conditions, the mixture baseline maintainability function, $(M_{0m}(\mathrm{t}))$, can be defined by mixing the $M_0(t)$ of the several repair processes as:

$$M_{0m}(\mathrm{t}) = \sum_{\mathrm{k}=1}^{\mathrm{N}} \gamma_{\mathrm{k}} . M_{0k} = \sum_{\mathrm{k}=1}^{\mathrm{N}} \gamma_{\mathrm{k}} . (1 - exp\left[-\int_0^t \mu_{k0}\left(t'\right) dt'\right]) \tag{16}$$

where M_{0k} is the baseline maintainability of the k^{th} repair process, and Υ_{k} is the proportion of the repair tasks belonging to the k^{th} repair process. The baseline maintainability function, if the repair rate for all repair processes follows 2-parameter Weibull distribution, is given by:

$$M_{0m}(\mathrm{t}) = \sum_{\mathrm{k}=1}^{\mathrm{N}} \gamma_{\mathrm{k}} . \left(1 - e^{-\left(\frac{t}{\eta_{\mathrm{k}}}\right)^{\beta_{\mathrm{k}}}}\right) \tag{17}$$

where β_{k} and η_{k} are the shape parameter and scale parameter of Weibull distribution for the k^{th} repair process. The application of the proposed model is demonstrated by a case study in Paper C.



CHAPTER 5

5. CONCLUSION, CONTRIBUTIONS AND FURTHER RESEARCH

In this chapter the conclusions, the author's contributions and possible future work are discussed.

Following the problem statement, objectives and research questions, the research continued with a review of the literature on performance measurement systems, to understand the key concepts, definitions, and aspects of criteria for measuring the performance in a company. Then a PMS for Arctic operational conditions was suggested. Thereafter, the research continued by developing models to quantify performance indicators, considering harsh climate conditions and risk factors.

5.1. Conclusions

The following conclusions have been reached:

- Due to low temperatures, sea ice, polar low pressures, poor visibility and seasonal darkness, etc., the Arctic operating condition has significant effects on the performance of component and industry activities in various ways, including increasing failure rate and repair, life cycle costs and safety hazards. Moreover, the less developed infrastructure in the Arctic creates several challenges, such as limitations to the logistics of supplies, material and personnel required for the operation and maintenance activities.
- Considering the unique and challenging Arctic operational conditions, with strict regulations and requirements for safety and the environment, the designed system or equipment must be available and safe, as well as economically viable. Such systems must be able to minimize environmental pollution and require the minimum quantity of raw materials and energy.
- PMS needs up-to-date and accurate performance information on its business. This performance information needs to be integrated, dynamic and accessible, to assist fast decision-making. Hence, the development and continuous improvement of PMSs and the identification of more PIs for judging performance in the Arctic are critical for industry success.
- Obtaining the correct sources of information and data on PIs and using the right tools or methods to measure the impact of the external environment or factors influencing PIs, is critical.
- The statistics models must be able to quantify the effect of covariates on PIs and must be built based on correct assumptions that reflect the operational conditions.
- Failure and repair data are often collected from multiple and distributed units in different operational conditions (e.g. operator skill, maintenance strategies, etc.), which can

introduce heterogeneity into the data. Part of such heterogeneity can be explained and isolated by the observable covariates, whose values and the way that they can affect the item's reliability or maintainability are known. However, some factors are typically unknown, leading to unobserved heterogeneity.

- In most reliability and maintainability studies, the effect of unobserved covariates is neglected for reasons of practicality and simplicity. This may lead to erroneous model selection for the time to failure or repair of the item, as well as wrong conclusions and decisions.
- In most maintainability analysis, researchers have simplified their analysis, by considering a complex system as a single item. In these studies, the assumption is that all repair data represent an identical repair process for the item. In practice, mechanical systems are composed of multiple parts, with various failure mechanisms, which need different repair processes (repair modes) to return to the operational phase; classical distribution, such as lognormal, which is only a function of time, may not be able to model such complexity.

5.2. Research contributions

Models and frameworks developed in this work can be employed in facilities and technology activities to analyse the impact of operational conditions on the performance of systems/equipment and to assist calculations and predictions. The research contributions are considered to be:

- Suggestion of a PMS for the Arctic.
- Development of an Icing Risk Index to quantify the effect of different types of ice on the performability of an item/system.
- Development of a methodology for reliability analysis in the presence of unobserved and observed covariates.
- Development of a Mixture Frailty Model (MFM) to estimate the effect of observed and unobserved covariates on the maintainability of a component or a system, as well as identifying different repair processes in the repair dataset.

5.3. Suggested future work

Based on the research presented in this thesis, the following points are suggested for future research:

• Development of applicable model for predicting the effect of operational conditions on sustainability performance.

• Development of applicable model for predicting the effect of operational conditions on financial performance.

• Improvement of the data collection system, in order to map and collect covariates that can affect performability.

REFERENCES

ABRAN, A. & BUGLIONE, L. 2003. A multidimensional performance model for consolidating balanced scorecards. *Advances in Engineering Software,* 34**,** 339-349.

ACI 2005. *Arctic climate impact assessment*, Cambridge University Press.

AKSENOV, Y., POPOVA, E. E., YOOL, A., NURSER, A. G., WILLIAMS, T. D., BERTINO, L. & BERGH, J. 2017. On the future navigability of Arctic sea routes: High-resolution projections of the Arctic Ocean and sea ice. *Marine Policy,* 75**,** 300-317.

ALSYOUF, I. 2004. *Cost effective maintenance for competitve advantages.* Växjö university press.

ALVARSSON, M. & SKÖLDBERG, K. 1994. Interpretation and Reflection; Philosophy of Science and Qualitative Method.

AMAP 2010a. Assessment 2007: Oil and Gas Activities in the Arctic-Effects and Potential Effects. . Volume 1.

AMAP 2010b. Assessment 2007: Oil and Gas Activities in the Arctic-Effects and Potential Effects. Volume 1.

ANTTONEN, H. & VIROKANNAS, H. 1994. Assessment of cold stress in outdoor work. *Arctic medical research,* 53**,** 40-48.

BARABADI, A. 2012. Reliability and Spare Parts Provision Considering Operational Environment: A Case Study. *International Journal of Performability Engineering,* 8**,** 497.

BARABADI, A. 2014. Reliability analysis of offshore production facilities under arctic conditions using reliability data from other areas. *Journal of Offshore Mechanics and Arctic Engineering,* 136**,** 021601.

BARABADY, J., MARKESET, T. & KUMAR, U. 2010. Review and discussion of production assurance program. *International Journal of Quality & Reliability Management,* 27**,** 702-720.

BITITCI, U. S., CARRIE, A. S. & MCDEVITT, L. 1997. Integrated performance measurement systems: a development guide. *International journal of operations & production management,* 17**,** 522-534.

BLANCHET, D., CROASDALE, K., SPRING, W. & THOMAS, G. ISO TC 67 SC 7 WG 8—An International Standard for Arctic Offshore Structures. Proceedings of the International Conference on Port and Ocean Engineering Under Arctic Conditions, 2007.

BOTTANI, E., MONICA, L. & VIGNALI, G. 2009. Safety management systems: Performance differences between adopters and non-adopters. *Safety science,* 47**,** 155-162.

BOURNE, M., MILLS, J., WILCOX, M., NEELY, A. & PLATTS, K. 2000. Designing, implementing and updating performance measurement systems. *International Journal of Operations & Production Management,* 20**,** 754-771.

BOURNE, M., NEELY, A., MILLS, J. & PLATTS, K. 2003. Implementing performance measurement systems: a literature review. *International Journal of Business Performance Management,* 5**,** 1-24.

BRADLEY, P. 1996. A performance measurement approach to the reengineering of manufacturing enterprises. *CIMRU, NUI Galway*.

BRUNDTLAND, G. H., KHALID, M., AGNELLI, S., AL-ATHEL, S. & CHIDZERO, B. 1987. Our common future. *In:* TOLBA, M. K. & BISWAS, A. K. (eds.) *Earth and Us: Population Resources Environment Development.* Elsevier.

BRUNVOLL, S., FURU, B. & PÅSCHE, A. 2010. Cold challenges-Health and Working Environment on Facilities in the Northern Areas. *Thelma AS.* Petroleum safety Authority Norway.

BUTLIN, J. 1989. Our common future. By World commission on environment and development. (London, Oxford University Press, 1987, pp. 383£ 5.95.). *Journal of International Development,* 1**,** 284-287.

CESAROTTI, V., GIUIUSA, A. & INTRONA, V. 2013. Using overall equipment effectiveness for manufacturing system design. *Operations Management***,** 51-79.

COHEN, J. M. 1993. *Building sustainbable public sector managerial, professional, and technical capacity: a framework for analysis and intervention,* Harvard Institute for International Development.

COMMISSION, P. 2013. On efficiency and effectiveness: some definitions. *Staff Research Note, Canberra***,** 1-14.

COUNCIL 2015. Realizing the Promise of US Arctic Oil and Gas Resources. *National Petroleum Council*.

DUTTA, P. K. 1988. Behavior of Materials at Cold Regions Temperatures. Part 1. Program Rationale and Test Plan. DTIC Document.

ECCLES, R. 1991. The performance measurennent manifesto. *Harvard business review,* 69**,** 131-137.

FERRER-BALAS, D., ADACHI, J., BANAS, S., DAVIDSON, C. I., HOSHIKOSHI, A., MISHRA, A., MOTODOA, Y., ONGA, M. & OSTWALD, M. 2008. An international comparative analysis of sustainability transformation across seven universities. *International Journal of Sustainability in Higher Education,* 9**,** 295-316.

FIKKE, S., RONSTEN, G., HEIMO, A., KUNZ, S., OSTROZLIK, M., PERSSON, P., SABATA, J., WAREING, B., WICHURE, B. & CHUM, J. 2006. COST 727: Atmospheric Icing on structures. *Measurements and data collection on icing: State of the Art, Publication of MeteoSwiss,* 75**,** 1422-1381.

FITZGERALD, L., JOHNSTON, R., BRIGNALL, T., SILVESTRO, R. & VOSS, C. 1991. *Performance measurement in service businesses,* Chartered Institute of Management Accountants London.

FOW, S. J. 2011. The Arctic Sustainability Principle. *Vermont Journal of Environmental Law,* 13**,** 537.

FRANCO-SANTOS, M., KENNERLEY, M., MICHELI, P., MARTINEZ, V., MASON, S., MARR, B., GRAY, D. & NEELY, A. 2007. Towards a definition of a business performance measurement system. *International Journal of Operations & Production Management,* 27**,** 784-801.

FREITAG, D. R. & MCFADDEN, T. T. 1997. *Introduction to cold regions engineering,* ASCE press New York.

FURULY, S., BARABADI, A. & BARABADY, J. 2013. Reliability Analysis of Mining Equipment Considering Operational Environments-A Case Study. *International Journal of Performability Engineering,* 9.

GAO, X., BARABADY, J. & MARKESET, T. 2010. An approach for prediction of petroleum production facility performance considering Arctic influence factors. *Reliability Engineering & System Safety,* 95**,** 837-846.

GAO, X. & MARKESET, T. Design for production assurance considering influence factors. Proceedings of the European Safety and Reliability Conference (ESREL 2007), Stavanger, 2007 2007. McGraw-Hill.

GARCÍA-GRANERO, E. M., PIEDRA-MUÑOZ, L. & GALDEANO-GÓMEZ, E. 2018. Eco-innovation measurement: A review of firm performance indicators. *Journal of cleaner production,* 191**,** 304-317.

GARENGO, P., BIAZZO, S. & BITITCI, U. S. 2005. Performance measurement systems in SMEs: A review for a research agenda. *International journal of management reviews,* 7**,** 25-47.

GARMABAKI, A., AHMADI, A. & AHMADI, M. 2016. Maintenance optimization using multi-attribute utility theory. *Current trends in reliability, availability, maintainability and safety.* Springer.

GARVIN, D. A. 1994. Building a learning organization. *BUSINESS CREDIT-NEW YORK-,* 96**,** 19-19.

GHALAYINI, A. M. & NOBLE, J. S. 1996. The changing basis of performance measurement. *International journal of operations & production management,* 16**,** 63-80.

GHALAYINI, A. M., NOBLE, J. S. & CROWE, T. J. 1997. An integrated dynamic performance measurement system for improving manufacturing competitiveness. *International Journal of production economics,* 48**,** 207-225.

GHODRATI, B. 2005. *Reliability and operating environment based spare parts planning.* Doctoral Thesis, Luleå University of Technology.

GLOBERSON, S. 1985. Issues in developing a performance criteria system for an organization. *International journal of production research,* 23**,** 639-646.

GUDMESTAD, O. 2010. Vessel winterization for cold climate operations. *The 5th Norway – Russia Arctic Offshore Workshop.* Joint Research and Innovation for the Petroleum industry working in the Arctic: University of Stavanger, Stavanger, NTNU, Trondheim.

HALLSTEDT, S., NY, H., ROBÈRT, K.-H. & BROMAN, G. 2010. An approach to assessing sustainability integration in strategic decision systems for product development. *Journal of Cleaner Production,* 18**,** 703-712.

HANNULA, M. 2002. Total productivity measurement based on partial productivity ratios. *International journal of production economics,* 78**,** 57-67.

HANSEN, J. C. & VAN OOSTDAM, J. 2009. AMAP Assessment 2009: Human Health in the Arctic.

HARRISON, H., BIRKS, M., FRANKLIN, R. & MILLS, J. 2017. Case Study Research: Foundations and Methodological Orientations. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 18.

HAUSER, J. & KATZ, G. 1998. Metrics: you are what you measure! *European Management Journal,* 16**,** 517-528.

HAYES, R. H. & ABERNATHY, W. J. 1980. Managing our way to economic decline. *Harvard Bus. Rev.;(United States),* 58.

HOLMÉR, I. 1994. Cold stress: Part I—Guidelines for the practitioner. *International Journal of industrial ergonomics,* 14**,** 139-149.

HOPKINS, M. J. D. 2002. Sustainability in the internal operations of companies. *Corporate Environmental Strategy,* 9**,** 398-408.

IEC 2019. IEC 60050 - International Electrotechnical Vocabulary.

IOS 2008. ISO/DIS 20815: Petroleum, petrochemical and natural gas industries, Production assurance and reliability management. International Organization for Standardization.

ISAZA, C. E., HERRERA KIT, P., LOZANO HERRERA, J., MENDEZ, K. & BALANZO, A. 2015. Capacity: A literature review and a research agenda. Rochester, NY: Social Science Research Network.

JAN, F. & PETRA, L. 2016. Sustainable Value as a Tool for Corporate Performance Management within New Public Management Framework. *WSEAS Transactions on Environment and Development, volume 12, issue: 12, 2016*.

JIANG, R. & JI, P. 2002. Age replacement policy: a multi-attribute value model. *Reliability Engineering & System Safety,* 76**,** 311-318.

JONES, K. F. & ANDREAS, E. L. 2012. Sea spray concentrations and the icing of fixed offshore structures. *Quarterly Journal of the Royal Meteorological Society,* 138**,** 131-144.

KALBFLEISCH, J. D. & ROSS, L. 1983. Prentice. 1980. The statistical analysis of failure time data. *New York: John Wiley& Sons, Inc. KalbfleischThe Statistical Analysis of Failure Time Data1980*.

KAPLAN, R. 1991. The evolution of management accounting, readings in accounting for management control. International Thomson Business Press.

KAPLAN, R. S. 1984. The evolution of management accounting. *Readings in accounting for management control.* Springer.

KAPLAN, R. S. 2005. How the balanced scorecard complements the McKinsey 7-S model. *Strategy & leadership,* 33**,** 41-46.

KAPLAN, R. S. & NORTON, D. P. 1992. The balanced scorecard: measures that drive performance. *Harvard business review,* 87**,** 172.

KAPLAN, R. S. & NORTON, D. P. 1996. Using the balanced scorecard as a strategic management system. Harvard business review Boston.

KATIC, D., MAJSTOROVIC, V. & COLAK, I. 2011. Performance measurement review. *Annals of DAAAM & Proceedings,* 22**,** 515-516.

KAYRBEKOVA, D., BARABADI, A. & MARKESET, T. 2011. Maintenance cost evaluation of a system to be used in Arctic conditions: a case study. *Journal of Quality in Maintenance Engineering,* 17**,** 320-336.

KENNERLEY, M. & NEELY, A. 2002. A framework of the factors affecting the evolution of performance measurement systems. *International Journal of Operations & Production Management,* 22**,** 1222-1245.

KEOLEIAN, G. A. & MENEREY, D. 1994. Sustainable development by design: review of life cycle design and related approaches. *Air & Waste,* 44**,** 645-668.

KHAN, K. & SHAH, A. 2011. Understanding performance measurement through the literature. *African Journal of Business Management,* 5**,** 13410-13418.

KLEIN, J. P., VAN HOUWELINGEN, H. C., IBRAHIM, J. G. & SCHEIKE, T. H. 2016. *Handbook of survival analysis,* CRC Press.

KLEIN, L. R., LONG, V., GREENSPAN, A., GREENWALD, D., EDMONSON, N. & PERRY, G. 1973. Capacity utilization: concept, measurement, and recent estimates. *Brookings Papers on Economic Activity,* 1973**,** 743-763.

KNEZEVIC, J. 1993. *Reliability, maintainability, and supportability: a probabilistic approach*, McGraw-Hill Companies.

KOIVUROVA, T. 2013. The dialectic of understanding progress in Arctic governance. *Mich. St. Int'l L. Rev.,* 22**,** 1.

KOTHARI, C. R. 2004. *Research methodology: Methods and techniques*, New Age International.

KRATZ, L. 2003. Designing and assessing supportability in DoD weapon systems: A guide to increased reliability and reduced logistics footprint. OFFICE OF THE SECRETARY OF DEFENSE WASHINGTON DC.

KRISTOFFERSEN, B. & LANGHELLE, O. 2017. Sustainable development as a global-Arctic matter: Imaginaries and controversies. *Governing Arctic Change.* Springer.

KUMAR, R., BARABADY, J., MARKESET, T. & KUMAR, U. 2012. Improving maintainability in extreme cold climatic conditions. *International Journal of Performability Engineering,* 8**,** 563.
KUMAR, U., AHMADI, A., VERMA, A. K. & VARDE, P. 2015. *Current Trends in Reliability, Availability, Maintainability and Safety: An Industry Perspective,* Springer.
LAAKSO, T., HOLTTINEN, H., RONSTEN, G., TALLHAUG, L., HORBATY, R., BARING-GOULD, I., LACROIX, A., PELTOLA, E. & TAMMELIN, B. 2003. State-of-the-art of wind energy in cold climates. *IEA annex XIX,* 24.
LASSALA, C., APETREI, A. & SAPENA, J. 2017. Sustainability matter and financial performance of companies. *Sustainability,* 9**,** 1498.
LEBAS, M. J. 1995. Performance measurement and performance management. *International journal of production economics,* 41**,** 23-35.
LIN, G. & SHEN, Q. 2007. Measuring the performance of value management studies in construction: critical review. *Journal of Management in Engineering,* 23**,** 2-9.
LINGLE, J. H. & SCHIEMANN, W. A. 1996. From balanced scorecard to strategic gauges: is measurement worth it? *Management review,* 85**,** 56.
LJUNGBERG, Õ. 1998. Measurement of overall equipment effectiveness as a basis for TPM activities. *International Journal of Operations & Production Management,* 18**,** 495-507.
LYNCH, R. L. & CROSS, K. F. 1991. *Measure up!: The essential guide to measuring business performance*, Mandarin.
MALTZ, A. C., SHENHAR, A. J. & REILLY, R. R. 2003. Beyond the balanced scorecard:: Refining the search for organizational success measures. *Long Range Planning,* 36**,** 187-204.
MARKESET, T. 2008. Design for high performance assurance for offshore production facilities in remote harsh and sensitive environments. *Opsearch,* 45**,** 275-290.
MARKESET, T. & KUMAR, U. 2005. Product support strategy: conventional versus functional products. *Journal of Quality in Maintenance Engineering,* 11**,** 53-67.
MARKESET, T., SÆLAND, A., GUDMESTAD, O. & BARABADY, J. 2015. 11 Petroleum production facilities in Arctic operational environments. *International Arctic Petroleum Cooperation: Barents Sea Scenarios***,** 184.
MARSHALL, C. & ROSSMAN, G. B. 2014. *Designing qualitative research*, Sage publications.
MAYYAS, A., QATTAWI, A., OMAR, M. & SHAN, D. 2012. Design for sustainability in automotive industry: A comprehensive review. *Renewable and sustainable energy reviews,* 16**,** 1845-1862.
MEIER, W. N., HOVELSRUD, G. K., VAN OORT, B. E., KEY, J. R., KOVACS, K. M., MICHEL, C., HAAS, C., GRANSKOG, M. A., GERLAND, S. & PEROVICH, D. K. 2014. Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity. *Reviews of Geophysics,* 52**,** 185-217.
MIKKELSEN, A. & LANGHELLE, O. 2008. *Arctic oil and gas: sustainability at risk?,* Routledge.
MISRA, K. B. 2008a. Dependability Considerations in the Design of a System. *Handbook of Performability Engineering.* Springer.
MISRA, K. B. 2008b. *Handbook of performability engineering,* Springer Science & Business Media.
MISRA, K. B. 2008c. Performability Engineering: An Essential Concept in the 21st Century. *Handbook of Performability Engineering.* Springer.

MISRA, K. B. 2008d. Reliability Engineering: A Perspective. *Handbook of Performability Engineering.* Springer.
MUCHIRI, P. & PINTELON, L. 2008. Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion. *International journal of production research,* 46**,** 3517-3535.
MURRAY, J. L., HACQUEBORD, L., GREGOR, D. J. & LOENG, H. 1998. Physical/geographical characteristics of the Arctic. *AMAP assessment report: arctic pollution issues***,** 9-24.
NAKAJIMA, S. 1988. Introduction to TPM: total productive maintenance.(Translation). *Productivity Press, Inc., 1988***,** 129.
NANNI, A. J., DIXON, J. R. & VOLLMANN, T. E. 1990. Strategic control and performance measurement. *Journal of Cost Management,* 4**,** 33-42.
NASERI, M. & BARABADY, J. 2016. On RAM performance of production facilities operating under the Barents Sea harsh environmental conditions. *International Journal of System Assurance Engineering and Management,* 7**,** 273-298.
NASERI, M., BARALDI, P., COMPARE, M. & ZIO, E. 2016. Availability assessment of oil and gas processing plants operating under dynamic Arctic weather conditions. *Reliability Engineering & System Safety,* 152**,** 66-82.
NAZ, F., IJAZ, F. & NAQVI, F. 2016. Financial performance of firms: Evidence from Pakistan cement industry. *Journal of Teaching and Education,* 5**,** 81-94.
NEELY, A. & BOURNE, M. 2000. Why measurement initiatives fail. *Measuring business excellence,* 4**,** 3-7.
NEELY, A. D. 1998. *Measuring business performance*, Profile books.
NEELY, A. D., ADAMS, C. & KENNERLEY, M. 2002. *The performance prism: The scorecard for measuring and managing business success*, Prentice Hall Financial Times London.
NEUMAN, W. L. 2007. *Basics of Social Research: Quantitative And Qualitative Approaches*, Pearson/Allyn and Bacon.
NORTON, D. P. & KAPLAN, R. 1999. *The Balanced Scorecard: translating strategy into action*, Institute for International Research.
NUDURUPATI, S. S., BITITCI, U. S., KUMAR, V. & CHAN, F. T. 2011. State of the art literature review on performance measurement. *Computers & Industrial Engineering,* 60**,** 279-290.
O'CONNOR, P. D. 2008. A Practitioner's View of Quality, Reliability and Safety. *Handbook of Performability Engineering.* Springer.
O'NEILL JR, R. 2006. Moving from performance measurement to performance management. *Public Management,* 333611**,** 3.
OIL, A. 2011. Natural Gas Resources: United States Energy Information Administration.
ORIMOLADE, A. P., GUDMESTAD, O. T. & WOLD, L. E. 2017. Vessel stability in polar low situations. *Ships and Offshore Structures,* 12**,** S82-S87.
PERRY, C. M. & ANDERSEN, B. 2012. New strategic dynamics in the Arctic region. *Institute for Foreign Policy Analysis. Feb*.
PETERS, G., NILSSEN, T., LINDHOLT, L., EIDE, M., GLOMSRØD, S., EIDE, L. & FUGLESTVEDT, J. 2011. Future emissions from shipping and petroleum activities in the Arctic. *Atmospheric Chemistry and Physics,* 11**,** 5305-5320.
PHADKE, M. S. 1995. *Quality engineering using robust design*, Prentice Hall PTR.
POMORSKI, T. Managing overall equipment effectiveness [OEE] to optimize factory performance. 1997 IEEE International Symposium on Semiconductor Manufacturing Conference Proceedings (Cat. No. 97CH36023), 1997 1997. IEEE, A33-A36.

QUILLÉROU, E., JACQUOT, M., CUDENNEC, A. & BAILLY, D. 2015. The Arctic: opportunities, concerns and challenges.
RAJASEKAR, S., PHILOMINATHAN, P. & CHINNATHAMBI, V. 2006. Research Methodology. *ArXiv preprint physics/0601009*, 23.
RYERSON, C. C. 2011. Ice protection of offshore platforms. *Cold Regions Science and Technology,* 65, 97-110.
SERREZE, M. C. & BARRY, R. G. 2011. Processes and impacts of Arctic amplification: A research synthesis. *Global and planetary change,* 77, 85-96.
SHAHIDUL, M., SYED SHAZALI, S., ABDULLAH, Y., TING, C., HISHAMUDDIN, A., AZRIN, M. & ADZLAN, A. 2013. Measuring Machinery Capacity Utilization and Its Impact on Manufacturing Performance and Environment. *Journal of manufacturing Operations Research and Sustainability*, 7-12.
SHI, X., VENEZIANO, D., XIE, N. & GONG, J. 2013. Use of chloride-based ice control products for sustainable winter maintenance: A balanced perspective. *Cold Regions Science and Technology,* 86, 104-112.
SMITH, C. & KNEZEVIC, J. 1996. Achieving quality through supportability-part I: concepts and principles. *Journal of Quality in Maintenance Engineering,* 2, 21-29.
SRIMAI, S., RADFORD, J. & WRIGHT, C. 2011. Evolutionary paths of performance measurement: an overview of its recent development. *International Journal of Productivity and Performance Management,* 60, 662-687.
STANDARD, N. 1998a. Regularity management & reliability technology. *Norwegian Technology Standards Institution, Oslo, Norway*.
STANDARD, N. 1998b. Regularity management & reliability technology. *Norwegian Technology Standards Institution, Oslo, Norway,* Z-016.
STOYANOV, D. & NIXON, J. 2020. Alternative operational strategies for wind turbines in cold climates. *Renewable Energy,* 145, 2694-2706.
SULLIVAN, T. J. 2001. *Methods of social research*, Harcourt College Publishers Fort Worth, TX.
SULTANA, S., ANDERSEN, B. S. & HAUGEN, S. 2019. 3Identifying safety indicators for safety performance measurement using a system engineering approach. *Process Safety and Environmental Protection*.
SUMSER, J. 2001. *A guide to empirical research in communication: Rules for looking*, Sage.
SUTTON, P. 2004. What is sustainability. *Eingana, the journal of the Victorian Association for Environmental Education*, 1-7.
TAOUAB, O. & ISSOR, Z. 2019. Firm Performance: Definition and Measurement Models.
European Scientific Journal January
TATICCHI, P., TONELLI, F. & CAGNAZZO, L. 2010. Performance measurement and management: a literature review and a research agenda. *Measuring business excellence,* 14, 4-18.
TRUMP, B. D., KADENIC, M. & LINKOV, I. 2018. A sustainable Arctic: Making hard decisions. *Arctic, Antarctic, and Alpine Research,* 50, e1438345.
VERITAS, D. N. 2002. Risk based inspection of offshore topsides static mechanical equipment. Oslo: DetNorske Veritas.
VINNEM, J. E. 2010. Risk indicators for major hazards on offshore installations. *Safety science,* 48, 770-787.
VUKŠIĆ, V. B., BACH, M. P. & POPOVIČ, A. 2013. Supporting performance management with business process management and business intelligence: A case analysis of integration and orchestration. *International journal of information management,* 33, 613-619.
YIN, R. K. 2003. *Case Study Research: Design and Methods*, SAGE Publications.

YIN, R. K. 2017. *Case study research and applications: Design and methods,* Sage publications.
YUN, G. & MARSDEN, A. 2010. Methodology for estimating probability of success of Escape, Evacuation, and Rescue (EER) strategies for arctic offshore facilities. *Cold Regions Science and Technology,* 61**,** 107-115.
ZAMECNIK, R. & RAJNOHA, R. 2015. Strategic business performance management on the base of controlling and managerial information support. *Procedia Economics and Finance,* 26**,** 769-776.

Part 2

Appended Papers

Paper A

Barabadi, Abbas; Zaki, Rezgar; Garmabaki, Amir Hossein Soleiman. Designing for performability: An icing risk index for Arctic offshore. Cold Regions Science and Technology 2016; Volum 124. ISSN 0165-232X.s 77 - 86.s

Cold Regions Science and Technology 124 (2016) 77–86

Contents lists available at ScienceDirect

Cold Regions Science and Technology

journal homepage: www.elsevier.com/locate/coldregions

Designing for performability: An icing risk index for Arctic offshore

Abbas Barabadi [a,*], A.H.S. Garmabaki [b], Rezgar Zaki [a]

[a] Department of Engineering and Safety, UiT The Arctic University of Norway, Tromsø, Norway
[b] Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden

ARTICLE INFO

Article history:
Received 14 June 2015
Received in revised form 29 November 2015
Accepted 31 December 2015
Available online 26 January 2016

Keywords:
Icing hazard
Maintainability
Quality
Reliability
Safety
Sustainability

ABSTRACT

Ice accretion affects the performability of offshore production facilities in various ways, including repair time and failure rate. It can increase power losses, life cycle costs, and safety hazards. There are few studies and limited systematically collected information about the impact of ice accretion on performability and its attributes (reliability, maintainability, quality, safety, and sustainability) for Arctic offshore production facilities. This paper will discuss the effects of different types of ice accretion on the performability of Arctic offshore production facilities. Then, to quantify their effect on the performability of offshore production facilities, an icing risk index is developed; its application is then shown by means of a case study.

1. Introduction

Over 25% of the world's undiscovered oil and gas petroleum reserves are expected to lie in the Arctic region, where approximately 84% of these reserves will be offshore (Gudmestad et al., 2007). The Arctic is known to be a harsh, sensitive, and remote area. Taking into consideration the unique and challenging Arctic operational conditions, the designed systems and selected technologies must be dependable and safe as well as economically viable (Gudmestad et al., 2007; Barabadi et al., 2015). Considering the high sensitivity of the area and its remoteness, which make any simple logistics a challenging task, the selected technology must be able to minimize environmental pollution and the need for spare parts and to require the minimum quantity of raw materials and energy.

Designing for production performance comprises appropriate approaches that can enable designers to meet these important goals. Production performance is described as the capacity of a system to meet the demand for delivery or performance. During recent years, production performance management programs have experienced faster development, and they increasingly play an important role in supporting the decision-making process (Gao et al., 2010). Moreover, some standards such as ISO20815 (2008) and NORSOK Z-016 (N. Z-016, 1998) have been developed to provide processes and activities, requirements and guidelines for systematic production performance management. The focus of these standards is on the reliability, availability, and maintainability of equipment. In most studies related to production performance, the main objective is to optimize reliability, availability, and maintainability to achieve the production assurance goals (Barabadi, 2011; Barabadi, 2012). Fig. 1 shows the performability concept of a system. As illustrated, an effective design should consist of all the attributes of performability simultaneously. A product or a system having these attributes is usually expected to perform well over its lifetime, incurring minimum life cycle costs.

Designing for performability in the harsh, sensitive, and remote Arctic area is a challenging task. It requires a range of tools to be employed and is dependent on a large amount of data and information. However, taking into consideration the unique Arctic operational conditions, most of the available tools need to be modified, and in some cases, new tools should be developed. At this point of development, there is not enough of the data and information (such as reliability or maintainability data and information) which is required for an accurate performability analysis.

One of the most challenging parts of performability analysis is quantifying the effect of operational conditions on performability (Barabadi, 2012; Kayrbekova et al., 2011). More specifically, ice accretion is one of the most hazardous operational conditions in cold regions; it can significantly affect the performability of equipment (Gudmestad et al., 2007; Ryerson, 2011). In a place like the North Sea, icing is considered more as a nuisance, but in the harsh Arctic climate condition, it can present many more operation, maintenance and safety problems (Gao et al., 2010; Ryerson, 2011). Furthermore, in the Arctic, icing can be very frequent, and, hence, according to the accident pyramid concept, it can lead to fewer but more serious accidents, such as those involving

* Corresponding author at: Safety and Environment, Department of Engineering and Safety, University of Tromsø, Tel.: +47 77660339; fax: +47 77644900.
E-mail address: abbas.b.abadi@uit.no (A. Barabadi).

http://dx.doi.org/10.1016/j.coldregions.2015.12.013
0165-232X/

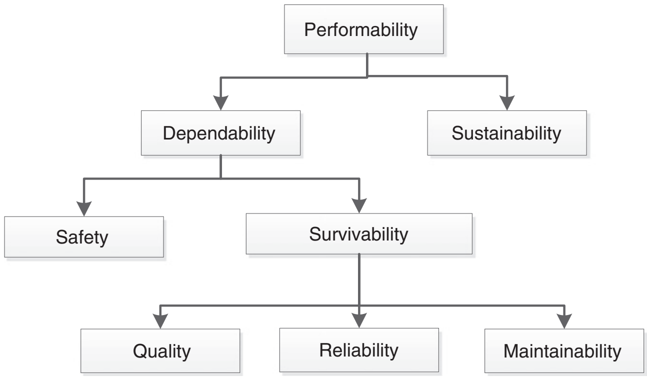


Fig. 1. Performability concept (Misra, 2008).

fatalities (Ryerson, 2011; Heinrich, 1941). Ice may build up in different forms (such as sea spray icing, frost, sleet, and glaze) based on the equipment shape, meteorological parameters (such as air temperature, wind speed, cloud liquid water content, cloud droplet spectra), and the elevation of equipment from sea level. Offshore platforms are complex systems with many items of equipment in various shapes, which makes them susceptible to different types of icing and icing problems (Ryerson, 2011).

Different types of icing may have different effects on performability attributes (sustainability, safety, quality, reliability, maintainability). Hence, to develop an effective practical solution and increase the performability of production facilities, one must have comprehensive knowledge of the different types of ice, how they form, and where they appear on offshore production facilities (Barabadi, 2012; Kayrbekova et al., 2011). In addition, it is very important to know how and how much they can affect the different concepts of performability. Considering the complexity of the icing effect on performability and its attributes, it is necessary to develop a standard factor for the identification, assessment, and prioritization of their risks. However, there are few studies and limited systematically collected information about the impact of ice on the performability of offshore production facilities. Recently, some researchers have focused on quantifying the effect of icing on offshore production facilities' performance in the Arctic (Ryerson, 2011). However, reliability and safety is their main focus, and these studies are not well detailed regarding the other concepts of performability. In this paper, the different types of ice accretion and their effect on the performability of offshore production facilities will be discussed. Then, in order to quantify the effect of different types of ice on performability, the concept of an icing effect index (IEI) on performability is developed. An IEI on performability can represent the consequences of icing on equipment. Thereafter, considering the probability of ice accretion and IEI on performability, the icing risk index for performability can be quantified. The rest of this paper is organized as follows: in Section 2, different types of ice accretion on offshore production facilities will be discussed; Section 3 will review the effect of icing on the different concepts of performability. Section 4 introduces the concept of an icing risk index for performability. The application of the method is illustrated by a case study in Section 5. Finally, Section 6 provides conclusions.

2. Physics of icing on Arctic offshore production facilities

Ice accretion is defined as the process of ice build-up on the surface of an object. Ice accretion on offshore production facilities can be categorized in two main groups: *i*) atmospheric icing and *ii*) sea spray icing or superstructure icing. Atmospheric icing is defined as the processes where falling or drifting raindrops, refrozen wet snow, or drizzle form accretions on an object that is exposed to the atmosphere (Ryerson, 2011; Farzaneh et al., 2008). Based on the procedure, feature and physical appearance, atmospheric icing can be categorized as (Ryerson, 2011; SAE, 2002):

- Glaze: This results from precipitating cold-water droplets that hit a surface and freeze upon impact. Up to 270 Metric Tons of glaze ice has been reported on a Canadian platform (Ryerson, 2011; Liljestrom and Lindgren, 1983), with thicknesses of up to 3 cm (Ryerson, 2011; Brown and Mitten, 1988).
- Snow: Snow accumulation up to 136 metric tons has been reported at a depth of 0.3 m on decks (Ryerson, 2011; Liljestrom and Lindgren, 1983).
- Rime: This results from droplets in fog, sea smoke, or cloud drops that hit a surface below 0 °C and freeze (Ryerson, 2011). Fett et al. (Fett et al., 1993) reported an accumulation of up to 10 cm on decks and 30 cm on railings in 12 h.
- Frost: This is the result of direct transformation of water vapor to ice and wet snow. Frost forms on windless clear nights on surfaces facing the sky (Ryerson, 2011).
- Sleet or ice pellets: Formed from raindrops that have been frozen before hitting surfaces, sleet accumulates loosely on horizontal surfaces such as decks, stairs, hatches, and helicopter landing pads (Ryerson, 2011).

In the case of sea spray icing, the sea spray droplets are carried by the wind and hit objects in their way. When the air temperature is colder than the freezing temperature of seawater, approximately around −2 °C, freezing spray occurs (Jones and Andreas, 2012). Waves, volume of spray flux, and salinity of seawater are important factors that affect rate of sea spray. Sea spray accumulation occurrence is very rapid when there are high winds, low air temperature, and low sea temperature.

Sea spray icing on stationary offshore structures differs significantly from sea spray icing on ships. Spray is generated on ships by heaving and pitching as the ship interacts with the waves it is moving through. Platform legs, bracing, blowout-preventer guidelines, mooring chains, marine risers, and flexible kill and choke lines in the splash zone 5–7 m above the sea are some potential areas for sea spray icing accumulation (Ryerson, 2011; Baller, 1983). Jones and Andreas (2009, 2012) developed a model to calculate the icing rate on cylinders with axes perpendicular to the wind direction. For more information, see also Horjen (2015). Fig. 2 shows the potential ice accretion areas on a drilling rig. For detailed information about different types of icing on a rig, see Ryerson (2009).

3. Icing effect on performability concepts

Different types of ice may have different effects on the performability attributes. For example, glaze is not an effective factor for the reliability or sustainability of a staircase in an offshore production facility, but it is a major hazard for safety and maintainability. Here, the effect of icing on performability attributes will be discussed briefly.

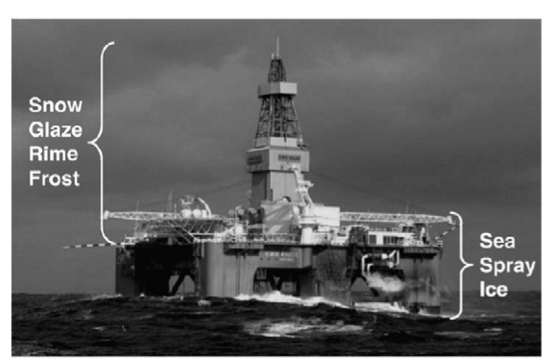


Fig. 2. Potential ice accretion areas on a rig (Ryerson, 2011).

3.1. A. Effects of icing on sustainability

A sustainable development is defined as a development that meets the needs of the present without compromising the ability of future generations to meet their own needs (Misra, 2008; Brundtland, 1987). The sustainability principle requires that products and systems use minimum material and energy throughout their entire life cycle. They should use non-hazardous materials and should be highly recyclable at the end of their life. The objective of sustainability is to increase energy and material efficiencies, preserve ecosystem integrity, and promote human health and happiness by merging design, economics, manufacturing, and policy. In general, to have a sustainable design, energy and material consumption should be minimized. In the Arctic offshore, in order to eliminate the problems of ice accretion, mechanical or electrical anti-icing and de-icing methods need to be taken into consideration (Farzaneh et al., 2008; Petrenko et al., 2011). Such methods negatively affect the sustainability by increasing energy and material consumption. For instance, the equipment and areas that require anti-icing measures should, as far as possible, be situated in protected locations, so that sea spray and weather cannot reach them. This may be accomplished by using fully enclosed spaces, semi-enclosures, recesses with removable “curtains” in front or similar, which can increase the material and energy usage significantly, and consequently, these protection methods will negatively affect the sustainability (DNV-OS-A201, 2013). In addition, such practice can be very expensive. For example, an investment of 5% of the cost of a 600 kW wind turbine has been estimated for the purchase and installation of anti-icing and de-icing systems (Laakso and Peltola, 2005). Or, for a windmill farm with medium icing severity, with an average of 30 icing days per year, the anti-icing and de-icing system payback time can be 5 years (Tammelin, 2005; Lamraoui et al., 2014). In addition, the large power demand of offshore installations in the Arctic area is, in most cases, covered by their own gas, and greenhouse gas emissions from power production are high. Ice-protection techniques with a high consumption of energy have negative impacts on the sensitive environment and wilderness in the Arctic. The use of hazardous chemical ice protection causes degradation of the environmental quality; it also increases the produced waste and serious environmental consequences (Shi et al., 2013).

3.2. B. Effects of icing on reliability

According to IEV191 (2015), reliability is “the ability of an item to perform a required function over specified time and under the specified conditions.” Most reliability studies in the Arctic offshore have focused on structural reliability (Ronalds et al., 2003; Der Kiureghian and Liu, 1986). However, there are still significant uncertainties with respect to the calculation of the loads of different types of ice (e.g. sea spray icing, atmospheric icing or floating sea ice) and their ice load effects on structures (Eik, 2011). The effects of icing on the reliability of equipment can be categorized as i) static ice loads, ii) wind action on iced structures and equipment, iii) dynamic effects, iv) damage caused by falling ice, and v) low-temperature stress on material.

Ice accretion will increase the stress (e.g. tension forces, oscillatory stresses) on structural components and mass imbalance. For example, when water sources are wind-driven and cables are oriented almost at right angles to the wind direction, ice accumulates on only one side. The torsional weak cable then rotates down, or twists, because of the weight of the ice accumulating on the side, and more ice accumulates on the new exposed face. This process, if occurring for a long-enough time, can cause cables to rotate multiple times with a spiral of ice enveloping them. Antennas and antenna structures can easily be overloaded by accreted ice. In particular, small fastening details are weakened when increased load is added on top of other actions, because the ice may easily double the normal load (Liljestrom and Lindgren, 1983). Ice from overhead may sag or fall causing unexpectedly high ice loads on lower structures.

Icing on structures and equipment will increase wind drag by changing dimensions and weight, shapes, and drag coefficients. For more information about wind action on iced structures see ISO12494 (2001). Moreover, it can change their natural frequencies, which is a significant factor influencing the dynamic behavior and control of the systems, leading to increasing oscillatory stresses. These stresses could cause fatigue in supports under the main deck and, potentially, loss of a rig (Ryerson, 2011). It should be mentioned that shedding and breaking of ice might cause important dynamic vibrations and stresses. For some materials such as plastic, low-temperature stress due to the icing can change the material's properties and increase their failure rate. Snow infiltration and extreme temperature lead to condensation in the electronics and, consequently, can lead to electrical failure (Laakso et al., 2003).

3.3. C. Effects of icing on quality

The quality of a product is a measure of its degree of conformance to applicable design specifications and workmanship standards (Misra, 2008). Offshore production facilities are made up of complex subsystems and components and they employ materials, men, and machines. These elements may have inherent variability and attributable variability. Variation of parameters and dimensions leads to weakening, component mismatch, incorrect fits, vibration, etc. These issues can increase the failure. Hence, it is necessary to establish an effective quality control plan to control the quality of these elements to an irreducible economic minimum. Quality can be classified into two types: design quality and manufacturing quality. In design quality, materials that are going to be used require early assessment for material selection and performance aspects to confirm their integrity over the life cycle of the facility. Hence, a set of the tests need to be established to check the quality of design. Understanding the physical environments involved and the stresses that can be applied at the site can prevent wear-out failures and overstress failures. As mentioned, due to the lack of experience and appropriate ice accretion models in the Arctic area, there are significant uncertainties with respect to calculation of ice effects such as ice load on different components. Such uncertainties can make a design quality a challenging process.

With regard to machine tools, the source of variation is the natural limits of capability that every process has. Ice accretion may provide a situation in which the process is incapable of acceptable operation within design limits. For example, welds will cool faster in cold weather, which results in increased susceptibility to cracking both during and after welding. Ice can reduce the quality of communication tools and sensors. For instance, wind vanes and temperature sensors can be affected by ice; studies show that, in icing conditions, wind speed errors can be as high as 30% (Laakso et al., 2003).

The most important contributor to variability is man himself (Misra, 2008). Studies show that in outdoor work in the winter, cold stress frequently reduces working ability by 70% for short periods (Anttonen and Virokannas, 1994). Long periods of exposure to the cold results in decreased cognitive performance, injury, hypothermia, loss of sensitivity, and reduced manual dexterity and grip (Holmér, 1994). These conditions can directly influence the variability of man's decisions or actions to a very large extent.

3.4. D. Effects of icing on maintainability

The formal definition of “maintainability,” according to IEV191 (2015), is “the ability of an item under given conditions of use, to be retained in, or restored to, a state in which it can perform a required function, when maintenance is performed under given conditions and using stated procedures and resources.” Maintainability measures the ease with which and the time in which a system can be restored to an operational level after failure. In general, the icing can contribute to changing the maintainability performance of an item by affecting

i) the maintenance and operator crew, ii) components and maintenance tools, and iii) maintenance support. In icing conditions, the maintenance and operator crew should wear warm clothes and gloves, which can increase the body dimensions and reduce mobility and hand dexterity. Slippery pathways due to icing can also reduce the maintenance crew's mobility. Poor visibility in a workspace due to the snow or iced windows makes it difficult to read technical data and manuals and increases the propensity to miss something or to repair something incorrectly and consequently it can reduce maintainability. Icing may change the accessibility of the failed item by changing its appearance and shape leading to improper accessibility. Improper accessibility can increase access, replacement, and removal time of failed components. Maintenance supervisors estimate that a 30% saving in overall maintenance time could be achieved if access to equipment were ideal or unrestricted. Effective communication is very important in maintainability. Ice on antennas can cause insufficient communication coverage, impeding conversations, between the operator and maintenance crew.

Ice accretion may cause problems for the safe and quick passage of both personnel and materials and increase the delivery time of spare parts. Icing may cause stoppages in helicopter activities, which are important for the logistics of transporting people and materials. Every maintenance activity needs tools such as cranes or lifting and hoisting devices to carry out tests, inspections, and repairs. The reliability and availability of maintenance tools can be adversely affected by ice. Iced crane components could jam the windlass, causing cables to jump pulleys or to jam in guides (Ryerson, 2011). Sensors on test equipment (e.g. temperature sensors, accelerometers, etc.) can be affected by different types of ice, leading to measurement errors in inspections and repairs process.

3.5. E. Effect of icing on safety

Stability and integrity of offshore production facilities can be threatened by large ice accretions. Icing can cause slippery surfaces on handrails, ladders, decks, etc., constituting an important personnel safety hazard. Ice accretions can cover boats' lifesaving apparatus, and deck firefighting equipment, which are vital and critical pieces of equipment. Ice on burner booms can lead to explosion, fire, or accumulation of toxic gases if the ice is over the burner boom's load rating (Ryerson, 2011). Icing on valves inhibits manual operation and the ability to see position indications. Ice falling from on high can hit the operator or maintenance crew. The overall probability of the success of an escape, evacuation, and rescue (EER) strategy will reduce significantly in the presence of ice on a production facility (Timco and Dickins, 2005). Yun and Marsden (2010) showed that, depending on the rescue (EER) strategy, the probability of success can fall from 90% in June to 50% in January. Mechanical de-icing methods in use generally require access to the iced equipment, which can provide safety hazards. For more information about the effect of icing on safety, see Ryerson (2011).

4. Icing risk index for performability

Very generally speaking, risk is the potential of losing something of value. Here, we are going to quantify the risk of ice accretion on the performability of offshore production facilities. Risk has two elements: the consequence of an accident and the probability of an accident occurring, and it can be defined as

$$Risk = \text{probability } of\ accident\ occurring \times consequences\ in\ the\ case\ of\ occurrence \tag{1}$$

Here, to quantify the consequences of icing on performability, the concept of an icing effect index for performability (IEIP) has been developed. The IEIP will range from *1* to *10*, where *1* shows no icing effect on the performability of the selected item (component, subsystem or system) and *10* shows a very high effect. Considering that different types of ice may change the performability of a selected item and its attributes in different ways, the IEIP should be quantified for each type of ice separately. Having an IEIP for each type of ice will enable designers and managers to identify appropriate mitigation methods for each one.

Performability has five principal attributes (reliability, maintainability, quality, safety, and sustainability) and two dependent attributes (dependability and survivability). To quantify the IEIP, in the first stage, the icing effect index (IEI) on the principal attributes of performability should be quantified. The IEI will range from *1* to *10*, where *1* shows no icing effect on performability attributes and *10* shows a very high effect. In the next stage, considering that performability attributes may place a different importance or weight on the overall performability of the item, the weight of the performability attributes needs to be quantified. The functionality and criticality of the item will decide the weight of the performability attributes. For a critical item, which needs to be repaired as soon as it fails, maintainability is an important factor that has a great effect on its survivability and consequently on its performability. Hence, its maintainability has a high weight compared to an item which has several redundancies and it can be repaired later in suitable opportunity. The weight factors of performability attributes should be quantified, disregarding the icing effect.

As can be concluded from Fig. 1, quantifying the IEIP is a bottom-to-top process that starts by quantifying the IEI on survivability (IEIS). For this aim, considering the effect of the selected types of ice on survivability attributes (*quality*, *reliability*, and *maintainability*), and the weight vectors of these attributes, the *IEIS* can be developed as:

$$IEIS = I_q^{\alpha_q} \times I_r^{\alpha_r} \times I_m^{\alpha_m} \tag{2}$$

where I_q, I_r, I_m are the IEIs for the quality, reliability, and maintainability, respectively. The parameters, $\alpha_q \geq 0$, $\alpha_r \geq 0$ and $\alpha_m \geq 0$, are the weight vectors of survivability attributes, where $\alpha_q + \alpha_r + \alpha_m = 1$. As mentioned, the criticality and required survivability of an item will decide the weight vectors of reliability, maintainability, or quality.

After quantification of the *IEIS*, the *IEI* for safety needs to be developed. Safety is one of the most sensitive elements of performability, which can easily be affected by icing. For instance, even a tiny amount of glaze or sleet can affect the safety significantly. Considering the IEI for safety, I_s, and *IEIS*, the IEI on the dependability index, *IEID*, of the item can be calculated as

$$IEID = IEIS^{\beta_{Sur.}} \times I_S^{\beta_S} \tag{3}$$

where $\beta_{Sur.} + \beta_s = 1$ and the $\beta_{Sur.\geq 0}$ and $\beta_{s\geq 0}$ show the weight vector of survivability and safety on dependability. Finally, the IEIP can be calculated by

$$IEIP = IEID^{\gamma_{Dep.}} \times I_{su.}^{\gamma_{su.}} \tag{4}$$

where $\gamma_{Dep.} + \gamma_s = 1$ and I_{su} is the IEI on sustainability and $\gamma_{Dep.\geq 0}$ and $\gamma_{su.\geq 0}$ show the weight vector of sustainability and dependability on performability.

After estimation of the IEIP, the probability of ice accretion (POI) should be estimated. POI is quantified by physical models, using statistical data or based on expert opinions. Finally, having IEIP and POI, the icing risk index for performability (IRIP) can be calculated by

$$IRIP = IEIP \times POI \tag{5}$$

Fig. 3 shows the different steps taken to calculate the IRIP. As this figure shows, after identification of the items, all relevant data should be collected including item geometry, design information, expert opinions, available experience related to the site, and meteorological data. Using collected data, the IEIP, POI, and IRIP for the selected type of ice will be calculated. Finally, the IRIP should be checked against acceptable criteria to see whether it is acceptable or not. In the case where risk is not

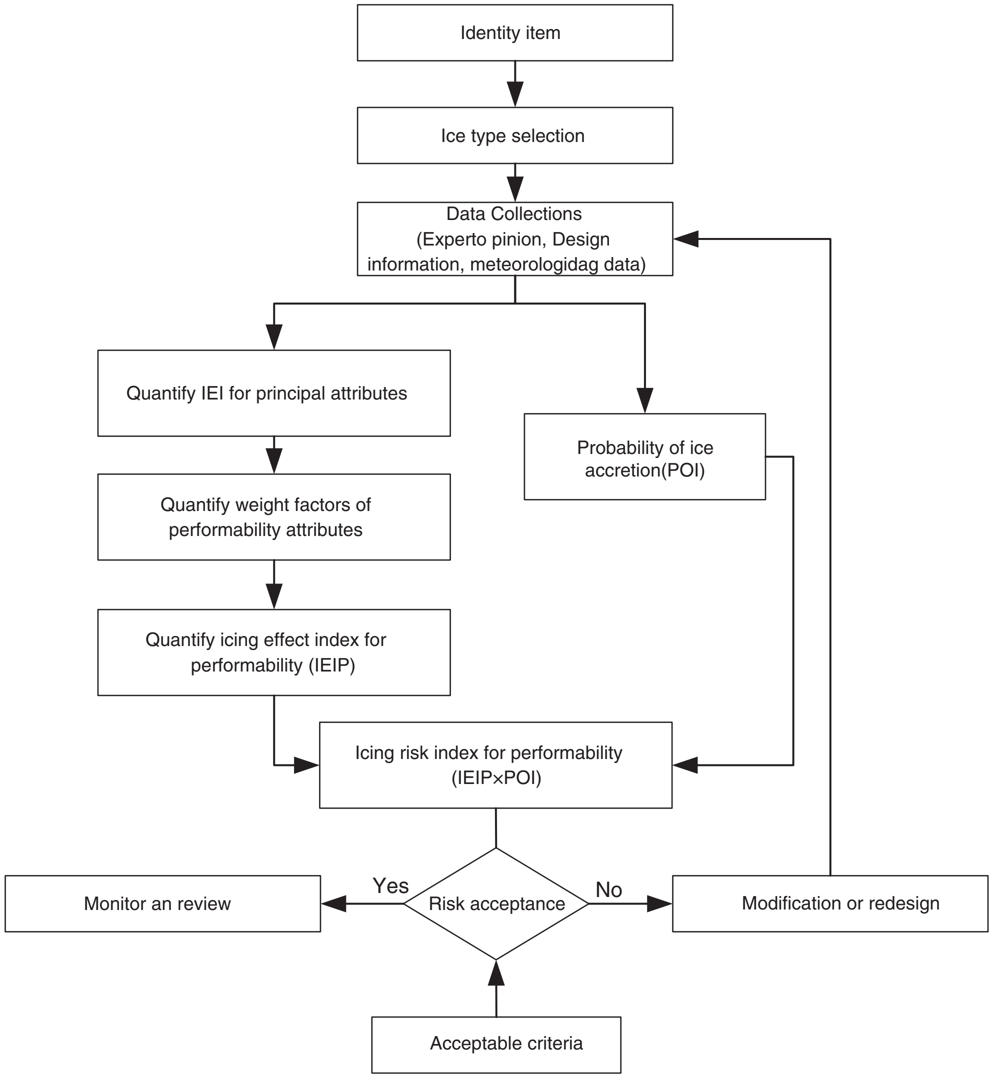


Fig. 3. Icing risk index for performability.

acceptable, some mitigation techniques should be applied. The mitigation techniques can be carried out *i*) to reduce the POI, for instance by using ice-phobic coating techniques; *ii*) to reduce IEIP, for example, by training operators and maintenance crew; or *iii*) to reduce both POI and IEIP. Moreover, having the POI and IEIP, the result can be shown on the risk matrix. A risk matrix is a matrix that is used during risk assessment to define the various levels of risk as the product of the probability categories and consequences categories. It should be mentioned that the number of categories should reflect the needs of the analysis. Hence, based on the goals and regulations, each company needs to create their own risk matrix.

One of the main challenges for calculating the *IRIP* is the collection of correct and sufficient information and data regarding ice effect and weight factors of performability attributes. As previously mentioned, at this stage of the work, due to limited engineering experience in the Arctic, this type of data is difficult to collect. Here, based on expert opinions and design information, a methodology is developed.

4.1. IEIS estimation using expert opinion

Here, the assumption is that the experts can provide the value for the IEI of principal performability attributes. To this aim, a formal expert judgment process can be followed, which consists of three main phases, namely, expert selection, elicitation of expert opinions, and aggregation of expert opinions (Meyer and Booker, 2001). Firstly, a number of experts should be selected based on a set of criteria (e.g., knowledge on cold-climate technology, icing phenomenon, and the equipment unit or the case of interest); once experts are selected, expert data are elicited through a formally defined procedure such as individual interview or Delphi approach (Meyer and Booker, 2001). In this regard, experts may be asked to provide their opinions in the form of single-point estimates, probability distribution, parameters of a distribution, intervals, etc. (Bedford and Cooke, 2001). Once expert opinions are elicited, an aggregation method is chosen to combine expert data and thus to obtain a single solution (e.g., single distribution function) that will be used by the decision-maker. Among different mathematical methods to combine expert data, weighted-arithmetic and weighted-geometric averaging techniques are the less complex ones, which are widely used in different applications of expert opinions (Clemen and Winkler, 2007). However, such methods require some weighting factors to be defined for experts, which is a challenging task. Several techniques such as equal-weighting, performance-based weighting, and computation of experts' weighting factors based on a set of predefined criteria are suggested in several studies (Clemen and Winkler, 2007; Cooke, 1991).

After estimation of IEI for principal attribute of performability, the IEI for dependent attribute should be estimated. Hence, weight factors for

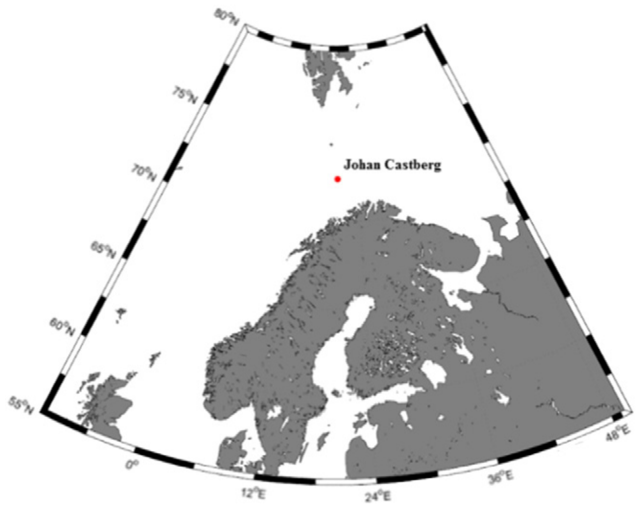


Fig. 4. Map of the locations of Johan Castberg (Statoil, 2012).

principal attribute should be estimated. Here, it is assumed that the values of weighted factors or their ranking for different performability attributes are available based on the design information. In the case where only the ranks of the weight vectors are available, the rank values should be quantified. Jiang and Ji (2002) developed a method that can be used for this aim. Here, the approach is discussed by means of an example; let us assume that for survivability of an item, the weight vector α_2 (reliability) is greater than α_3(maintainability) followed by α_1 (quality), i.e. $\alpha_2 \geq \alpha_3 \geq \alpha_1 ; \alpha_i > 0$. Let us define the following weight ratio as

$$\begin{cases} w_{3,1} = \alpha_3/\alpha_1 \\ w_{2,3} = \alpha_2/\alpha_3 \end{cases} \tag{6}$$

Based on Eq. (6), it is clear that $w_{3,1}, w_{2,3} \in [1, \infty)$. By simultaneously solving Eq. (6) along with weight constraint, $\alpha_1 + \alpha_2 + \alpha_3 = 1$, we have the following weight function:

$$\begin{cases} \alpha_1 = \dfrac{1}{w_{3,1} + w_{2,3} + 1} \\ \alpha_2 = \dfrac{1}{1/(w_{2,3}w_{3,1}) + (1/w_{2,3}) + 1} \\ \alpha_3 = \dfrac{1}{w_{2,3} + (1/w_{3,1}) + 1} \end{cases} \tag{7}$$

By considering Eq. (7) and $w_{3,1}, w_{2,3} \in [1, \infty)$, we are able to define the feasible area of the attributes' weights. In the next step, by analyzing Eq. (7), the extreme points of the feasible area can be found. The extreme points occur at the maximum or minimum point of the weight parameter as follows:

1. Weight factor α_1 achieves its minimum value, 0, when $w_{3,1} = w_{2,3} = \infty$, i.e., $\lim_{\substack{w_{31}\to\infty \\ w_{23}\to\infty}} \alpha_1 = 0$. Hence, the corresponding weight vector will be $(\alpha_1, \alpha_2, \alpha_3) = (0, 1, 0)$. In addition, it reaches its maximum when $w_{3,1} = w_{2,3} = 1$. The corresponding weight vector in this case is equal to $(\alpha_1, \alpha_2, \alpha_3) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.
2. Weight factor α_2 attains its minimum value, $\frac{1}{3}$, when $w_{3,1} = w_{2,3} = 1$, $\lim_{\substack{w_{31}\to 1 \\ w_{23}\to 1}} \alpha_2 = \dfrac{1}{3}$, and the corresponding weight vector will be $(\alpha_1, \alpha_2, \alpha_3) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. Also α_2 reaches its maximum when $w_{3,1} = w_{2,3} = \infty$, i.e., $\lim_{\substack{w_{31}\to\infty \\ w_{23}\to\infty}} \alpha_2 = 1$ and the corresponding weight vector will be $(\alpha_1, \alpha_2, \alpha_3) = (0, 1, 0)$.
3. Weight factor α_3 reaches its minimum value, 0, when $w_{3,1} = 1$, $w_{2,3} = \infty$, $\lim_{\substack{w_{31}\to 1 \\ w_{23}\to\infty}} \alpha_3 = 0$, and the corresponding weight vector $(\alpha_1, \alpha_2, \alpha_3) = (0, 1, 0)$. Also this parameter touches its maximum at $w_{3,1} = \infty$, $w_{2,3} = 1$, $\lim_{\substack{w_{31}\to\infty \\ w_{23}\to 1}} \alpha_1 = \dfrac{1}{2}$, and the weight vector in this case is $(\alpha_1, \alpha_2, \alpha_3) = \left(0, \frac{1}{2}, \frac{1}{2}\right)$.

As a result, the above analysis yields the following three extreme points (weight vectors): $W_1 = (\alpha_1, \alpha_2, \alpha_3) = (0, 1, 0)$, $W_2 = (\alpha_1, \alpha_2, \alpha_3) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, $W_3 = (\alpha_1, \alpha_2, \alpha_3) = \left(0, \frac{1}{2}, \frac{1}{2}\right)$.

5. Case study

The Johan Castberg field is newly discovered; drilling began in 2012 in the Norwegian Barents Sea (Fig. 4). Johan Castberg lies at a water depth of 360–390 m; it is located at a distance of about 200 km from the nearest land, which is Ingøya in Måsøy in Finnmark (Statoil, 2012).

On petroleum installations, such as drilling rigs, drill ships, semi-submersible platforms, etc., the presence of explosive and combustible substances increases the potential risk of fires and explosions. Hence, an effective escape, evacuation, and rescue (EER) plan is a key issue in such activities. EER addresses the entire process by which personnel are removed from a major accident event to an ultimate place of safety

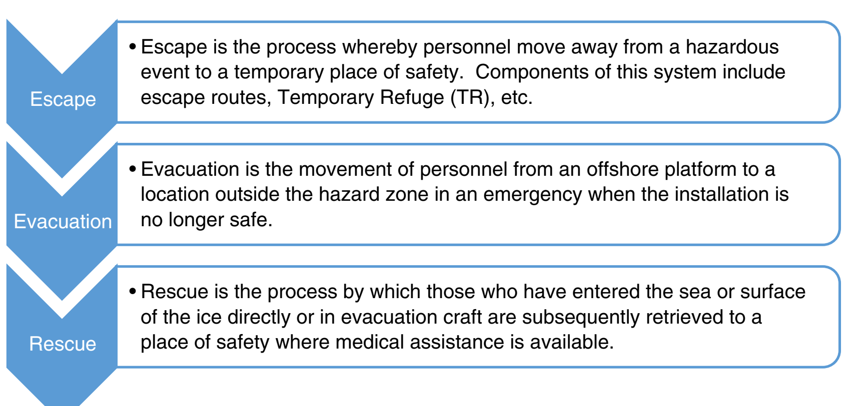


Fig. 5. The components of EER.

Fig. 6. West Hercules drilling rig used by Statoil in the Johan Castberg area, equipped with free-fall lifeboats (Seadrill, 2015).

(Yun and Marsden, 2010). The components of EER are described in Fig. 5.

In the Johan Castberg area, escape is not expected to be impacted by ice accretion. However, evacuation and rescue can be significantly affected by icing. In this paper, we focus on evacuation. Evacuation is further categorized as primary, secondary, and tertiary (Yun and Marsden, 2010; Thomson, 2015).

- Primary evacuation is the method of evacuation that can be carried out in a fully controlled manner under the direction of the person in charge. Primary evacuation in the Arctic can be carried out by craft such as helicopters, ice breaking platform supply vessels, and air cushioned vehicles. It is the preferred method of leaving the installation in an emergency.
- Secondary evacuation is a controlled means of removing personnel from the installation which can be carried out independent of external support.
- Tertiary escape constitutes "direct escape to sea"; i.e., if people cannot find their way to a lifeboat, they climb down to sea level, or even jump from a high level. Tertiary escape is clearly not advisable in the Arctic with its harsh climate conditions.

For further analysis, this paper considers helicopters for primary evacuation and lifeboats for secondary evacuation. They have been chosen due to their potential for successful evacuation in the Johan Castberg field, as is evidenced in similar environments, and to form a basis for demonstration of the tool and method (Marsden et al., 2011).

Table 2
Attributes' weight ranking of FFL and HLP.

Item	Survivability	Dependability	Performability
Free-fall lifeboat (FFL)	$\alpha_r \geq \alpha_q \geq \alpha_m$	$\beta_s \geq \beta_{sur.}$	$\gamma_{dep.} \geq \gamma_{su.}$
Helicopter landing pad (HLP)	$\alpha_r \geq \alpha_m \geq \alpha_q$	$\beta_s \geq \beta_{sur.}$	$\gamma_{dep.} \geq \gamma_{su.}$

Evacuation options involving helicopters play a major role in most evacuation plans when the evacuation can be performed in a fully controlled manner. However, in Johan Castberg, their operation is restricted by adverse weather conditions such as strong winds, low air temperatures, or atmospheric icing. Helicopters cannot be flown into known or forecast severe icing conditions (Peck et al., 2002). Besides, in the case of moderate icing conditions, the helicopter landing pad (HLP) installed on the cramped topsides of offshore installations will suffer to some degree from icing effects (Aviation, 2013). For example, frost on a helicopter landing pad creates slippery conditions. Slippery conditions on landing pads, which have no safety railings, could cause personnel to fall, sliding of the helicopter on the pad, and difficulty tying down the helicopter (Ryerson, 2009).

Regarding the lifeboats' application, different lifeboat types have been developed for offshore industry (Marsden et al., 2011). According to the available experience and studies, free-fall lifeboats (FFL) seem to be the preferred design (Fig. 6). Some of the ice accretion hazards related to FFL in the Johan Castberg area includes that

- the lifeboat is covered in snow or atmospheric icing during storage;
- the launching equipment is covered in snow or atmospheric icing;
- sea spray icing occurs shortly after launch;
- sea spray icing causes a significant amount of layer over time.

5.1. Qualifying IEI for principal attributes of FFL and HLP

Here, the aim is to identify the IEIs of different types of ice on the principal performability attributes of HLP and FFL based on expert opinions. In this case, a group of 13 experts has been asked to identify IEIs. After collecting the data, Cronbach's alpha coefficient is applied to measure the reliability of responses to the questionnaire. The value is $\alpha \geq 0.86$ and $\alpha \geq 0.82$ for FFL and HLP, respectively; this means that the reliability of responses is excellent. Thereafter, the geometrical mean is used to extract the final IEI-based questionnaires. Table 1 shows the final IEI for the principal attributes of performability in selected cases.

5.2. Qualifying weight vector of performability attributes for FFL and HLP

In this case, for FFL and HLP, only the ranking of the weight factor of each attribute is available (Table 2). As shown in Table 2, the ranking of the weight factors of the dependability and performability attributes

Table 1
Final IEI on performability attributes for FFL and HLP.

Selected item	Ice type	Quality (I_q)	Reliability (I_r)	Maintainability (I_m)	Safety (I_s)	Sustainability ($I_{su.}$)
Free-fall lifeboat (FFL)	Sea spray icing	7	8	9	8	7
	Snow	7	5	5	6	3
	Glaze	4	5	5	5	2
	Rime	3	4	4	4	2
	Frost	2	2	3	2	1
	Sleet	1	1	2	2	1
Helicopter landing pad (HLP)	Sea spray icing	7	9	9	9	7
	Snow	6	7	8	7	6
	Glaze	4	6	8	6	5
	Rime	2	4	5	5	4
	Frost	1	3	5	4	3
	Sleet	1	2	2	2	1

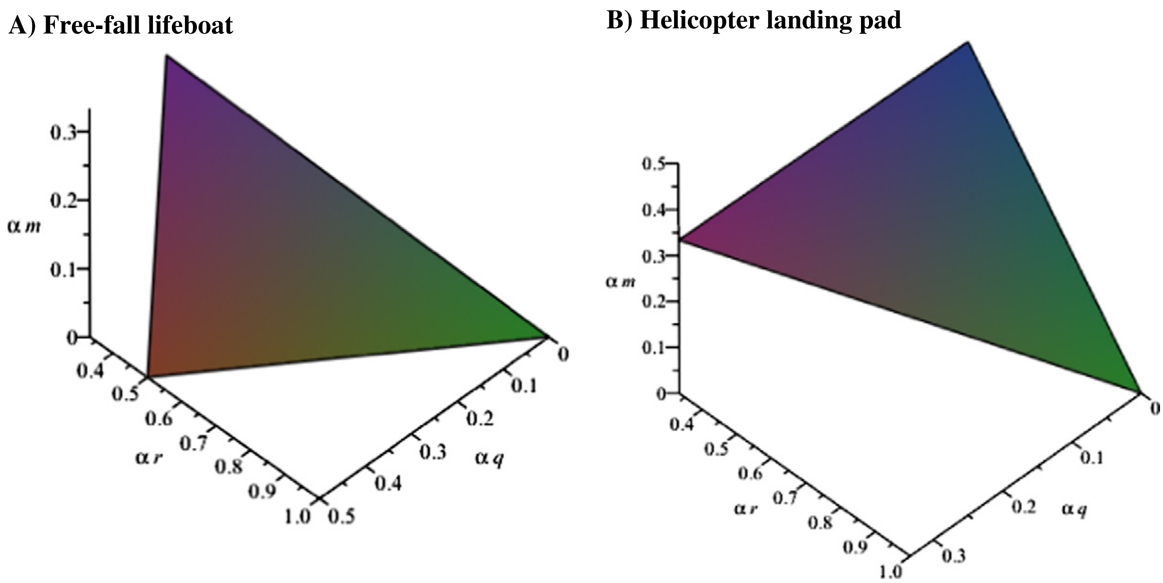


Fig. 7. The possible area for weighted vector of survivability of FFL and HLP.

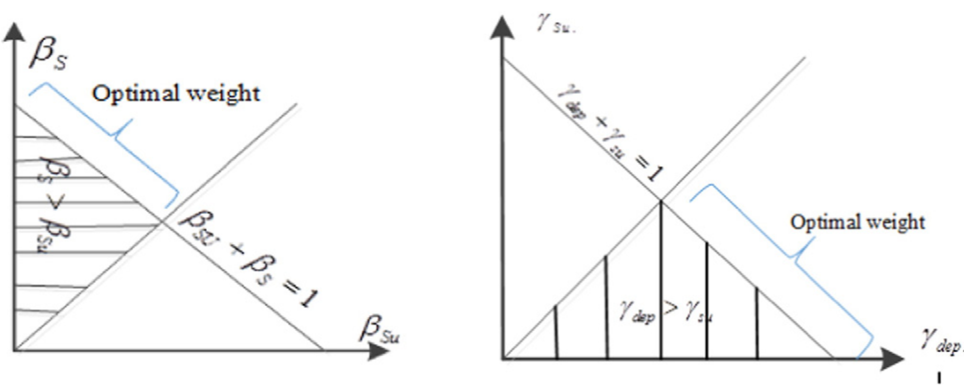


Fig. 8. The possible area for weighted vectors of dependability and performability.

are the same for FFL and HLP. However, in FFL survivability, quality has more weight than maintainability, in contrast with HLP, where maintainability has more weight than quality. Using the method discussed in Section 4.1 and the ranking of performability attributes in Table 2, the feasible values for the weighted factors can be calculated.

The feasible areas for the weight factors of survivability, dependability, and performability are shown in Figs. 7 and 8. As the weight rankings of the attributes are the same for dependability and performability in both cases, the feasible area will also be the same. The optimal values for survivability will be on the vertex of Fig. 7, and for dependability and performability, they will be on a line marked as "optimal weight" in Fig. 8. It should be mentioned that the final selected value for weight vectors needs to be approved by a group of experts or designers.

Table 3
Final selected weight values for performability attributes of FFL and HLP.

Item	Survivability	Dependability	Performability
Free-fall lifeboat (FFL)	$(\alpha_q,\alpha_r,\alpha_m)=(0,1,0)$	$(\beta_s,\beta_{sur.})=(\frac{3}{4},\frac{1}{4})$	$(\gamma_{dep.},\gamma_{su.})=(1,0)$
Helicopter landing pad (HLP)	$(\alpha_q,\alpha_r,\alpha_m)=(0,\frac{1}{2},\frac{1}{2})$	$(\beta_s,\beta_{sur.})=(\frac{1}{2},\frac{1}{2})$	$(\gamma_{dep.},\gamma_{su.})=(1,0)$

Different scenarios have been defined based on the feasible weight area to find the worst scenarios, where selected weight provides the highest icing impact on performability-dependent attributes. Table 3 shows the final selected weight vectors.

Table 4
The IEI on survivability, dependability, and performability for different types of ice.

Selected item	Performability element	Sea spray icing	Atmospheric icing				
			Snow	Glaze	Rime	Frost	Sleet
Free-fall lifeboat (FFL)	IEIS $(\alpha_q,\alpha_r,\alpha_m)=(0,1,0)$	8.0	5.0	5.0	4.0	2.0	1.0
	IEID $(\beta_s,\beta_{sur.})=(\frac{3}{4},\frac{1}{4})$	8.0	5.7	5.0	4.0	2.0	1.7
	IEIP $(\gamma_{dep.},\gamma_{su.})=(1,0)$	8.0	5.7	5.0	4.0	2.0	1.7
Helicopter landing pad (HLP)	IEIS $(\alpha_q,\alpha_r,\alpha_m)=(0,\frac{1}{2},\frac{1}{2})$	9.0	7.5	6.9	4.5	3.9	2.0
	IEID $(\beta_s,\beta_{sur.})=(\frac{1}{2},\frac{1}{2})$	9.0	7.2	6.4	4.7	3.9	2.0
	IEIP $(\gamma_{dep.},\gamma_{su.})=(1,0)$	9.0	7.2	6.4	4.7	3.9	2.0

Likelihood	IEIP				
	0-2	2-4	4-6	6-8	8-10
A: Almost certain – Is expected to occur in most circumstances					
B: Likely – Will probably occur in most circumstances	Frost, Sleet	Rime	Glaze Snow		
C: Possible – Might occur at some time				Sea spray icing	
D: Unlikely – Could occur at some time					
E: Rare – May occur only in exceptional circumstances					

Fig. 9. Risk matrix for FFL. Red: extreme risk; dark yellow: very high risk; yellow: high risk; dark green: moderate risk; green: low risk; gray: very low risk.

5.3. Icing risk index for performability

As illustrated in Fig. 3, having the IEI of the principal attributes and related weight vectors, the IEIS, IEID, and IEIP should be calculated using Eqs. (2), (3), and (4), respectively. The results are shown in Table 4.

In the next stage, in order to quantify the IRIP, the probability of accretion of different types of ice on FFL and HLP should be estimated. Rigs generally are not moving, and spray is generated only by wave motion against rig supports; this suggests that in the Arctic area, atmospheric ice could potentially contribute more to ice-related rig safety than sea spray (Ryerson, 2009). Previous studies have demonstrated a variety of views regarding atmospheric ice accretion versus sea spray. Makkonen (1984) suggests that, in the Arctic, sea spray is about 50% of the source for ice, the remaining being atmospheric sources. Most sea spray occurs 15–20 m above the sea surface, but it can be lifted as high as 60 m (Jorgensen, 1982). However, in such heights, the liquid water content due to the sea spray is low and thus presents less probability of sea spray icing on HLP and FFL, which are usually located well above the ocean surface. Here, five different categories are defined for the probability of ice accretion on HLP and FFL; thereafter, the group of experts is asked to identify the class for each type ice.

- A: Almost certain—Is expected to occur in most circumstances.
- B: Likely—Will probably occur in most circumstances.
- C: Possible—Might occur at some time.
- D: Unlikely—Could occur at some time.
- E: Rare—May occur only in exceptional circumstances

Considering the IEIP in Table 4 and the POI, the IRIP can be plotted on a risk matrix (Figs. 9 and 10). As shown in Figs. 9 and 10, here five categories are defined for IEIP.

This matrix can be used by designers or managers to see whether the current design is acceptable or whether some mitigation method should be implemented to reduce the IRIP. For example, for HLP, although the probability of sea spray icing is very low, it still provides a high risk for evacuation, which is unacceptable.

6. Conclusion

Designing for performability is an effective way to meet the design goal for a complex operational condition such as the Arctic region. Designing for performability implies less environmental pollution, reduced material and energy requirements, waste minimization, and finally, conservation and efficient utilization of available resources, which in turn results in minimum life cycle costs. Ice has significant effects on the performability of equipment. To manage and minimize icing effects on performability, it is necessary to study how, when, how much, and which type of ice will be accumulated on different items. Thereafter, their effects should be quantified by an appropriate approach. This paper has reviewed the effect of ice on performability elements and then developed a concept for an Icing Risk Index for performability, which can be used to quantify the effect of different types of ice on performability. Moreover, it can be used to compare the different design solutions regarding the effect of the icing. The result of the case study shows that the performability of free-fall lifeboats is significantly affected by sea spray ice (IEIP = 8), while sleet (IEIP = 1.7) has the minimum effect on its performability among the different types of ice. The same results are obtained for helicopter landing pads, where sea spray ice (IEIP = 9) has the maximum effect and sleet the minimum effect (IEIP = 2). Furthermore, the IRIP of FFL shows that the risk of icing for most types of icing (except frost and sleet for FFL and sleet for HLP) is in an unacceptable zone at a high or very high risk level. Hence, an

Likelihood	IEIP				
	0-2	2-4	4-6	6-8	8-10
A: Almost certain – Is expected to occur in most circumstances					
B: Likely – Will probably occur in most circumstances	Sleet	Frost	Rime	Snow Glaze	
C: Possible – Might occur at some time					
D: Unlikely – Could occur at some time					Sea spray icing
E: Rare – May occur only in exceptional circumstances					

Fig. 10. Risk matrix for HLP. Red: extreme risk; dark yellow: very high risk; yellow: high risk; dark green: moderate risk; green: low risk; gray: very low risk.

effective de-icing or anti-icing solution should be considered; this can vary for different types of icing.

References

Anttonen, H., Virokannas, H., 1994. Assessment of cold stress in outdoor work. Arctic Med. Res. 53, 40–48.
Aviation, A.U.C., 2013. CAP 437—standards for offshore helicopter landing areas. Civil Authority.
Baller, H., 1983. Rig winterization to allow year-round drilling off northern Norway. Oil Gas J.;(United States) 81.
Barabadi, A., 2011. Production Performance Analysis: Reliability, Maintainability and Operational Conditions. University of Stavanger, Norway (Doctoral Thesis).
Barabadi, A., 2012. Reliability and spare parts provision considering operational environment: a case study. Int. J. Performability Eng. 8, 497–506.
Barabadi, A., Gudmestad, O.T., Barabady, J., 2015. RAMS data collection under Arctic conditions. Reliab. Eng. Syst. Saf. 135, 92–99.
Bedford, T., Cooke, R., 2001. Probabilistic Risk Analysis: Foundations and Methods. Cambridge University Press.
Brown, R., Mitten, P., 1988. Ice accretion on drilling platforms off the east coast of Canada. Presented at the International Conference on Technology for Polar Areas. Tapir Publishers, Trondheim, Norway, Trondheim, Norway.
Brundtland, G.H., 1987. Our Common Future: World Commission on Environmental Development. The Brundtland-Report Oxford University Press, Oxford, UK.
Clemen, R.T., Winkler, R.L., 2007. Aggregating probability distributions. Advances in Decision Analysis: From Foundations to Applications. Cambridge University Press, New York, pp. 154–176.
Cooke, R.M., 1991. Experts in Uncertainty: Opinion and Subjective Probability in Science. Oxford University Press, New York, United States.
Der Kiureghian, A., Liu, P.-L., 1986. Structural reliability under incomplete probability information. J. Eng. Mech. 112, 85–104.
DNV-OS-A201, 2013. Winterization for Cold Climate Operations. Det Norske Veritas As.
Eik, K., 2011. Sea-ice management and its impact on the design of offshore structures. Cold Reg. Sci. Technol. 65, 172–183.
Farzaneh, M., Volat, C., Leblond, A., 2008. Anti-icing and De-icing Techniques for Overhead Lines. Springer.
Fett, R.W., Englebretson, R., Perryman, D., 1993. Forecasters Handbook for the Bering Sea, Aleutian Islands, and Gulf of Alaska DTIC Document .
Gao, X., Barabady, J., Markeset, T., 2010. An approach for prediction of petroleum production facility performance considering Arctic influence factors. Reliab. Eng. Syst. Saf. 95, 837–846.
Gudmestad, O., Løset, S., Alhimenko, A., Shkhinek, K., Tørum, A., Jensen, A., 2007. Engineering aspects related to Arctic offshore developments. St. Petersburg, Lan 255.
Heinrich, H.W., 1941. Industrial Accident Prevention. A Scientific Approach. the University of California: McGraw-Hill book Company, Incorporated.
Holmér, I., 1994. Cold stress: part I—guidelines for the practitioner. Int. J. Ind. Ergon. 14, 139–149.
Horjen, I., 2015. Offshore drilling rig ice accretion modeling including a surficial brine film. Cold Reg. Sci. Technol. 119, 84–110.
IEV191, 2015. International electrotechnical vocabulary (IEV), Chapter 191:dependability and quality of service. Available: http://std.iec.ch/iec60050 (November).
ISO12494, 2001. Atmospheric Icing of Structures. ISO, 2013, Geneva, Switzerland.
ISO20815, 2008. Petroleum, Petrochemical and Natural Gas Industries – Production Assurance and Reliability Management Switzerland, 2008.
Jiang, R., Ji, P., 2002. Age replacement policy: a multi-attribute value model. Reliab. Eng. Syst. Saf. 76, 311–318.
Jones, K.F., Andreas, E.L., 2012. Sea spray concentrations and the icing of fixed offshore structures. Q. J. R. Meteorol. Soc. 138, 131–144.
Jorgensen, T.S., 1982. Influence of Ice Accretion on Activity in the Northern Part of the Norwegian Continental Shelf: Otter Group.
Kayrbekova, D., Barabadi, A., Markeset, T., 2011. Maintenance cost evaluation of a system to be used in Arctic conditions: a case study. J. Qual. Maint. Eng. 17, 320–336.
Laakso, T., Peltola, E., 2005. Review on Blade Heating Technology and Future Prospects. BOREAS VII, Saariselka, Finland.
Laakso, T., Holttinen, H., Ronsten, G., Tallhaug, L., Horbaty, R., Baring-Gould, I., et al., 2003. State-of-the-art of wind energy in cold climates. IEA annex XIX. 24.
Lamraoui, F., Fortin, G., Benoit, R., Perron, J., Masson, C., 2014. Atmospheric icing impact on wind turbine production. Cold Reg. Sci. Technol. 100, 36–49.
Liljestrom, G., Lindgren, P., 1983. Ice weight a hazard to loaded semi. Offshore 43, 133–137.
Makkonen, L., 1984. Atmospheric Icing on Sea Structures. Engineer Research And Development Center Hanover Nh Cold Regions Research and Engineering Lab CRREL-MONO-84-2.
Marsden, A., Totten, M., Spring, W., 2011. Feasibility of escape, evacuation and rescue for facilities in Arctic shear zone environments. OTC Arctic Technology Conference.
Meyer, M.A., Booker, J.M., 2001. Eliciting and Analyzing Expert Judgment: a Practical Guide. 7 SIAM.
Misra, K.B., 2008. Handbook of Performability Engineering. Springer Science & Business Media.
N. Z-016, 1998. Regularity Management & Reliability Technology. Norwegian Technology Centre.
Peck, L., Ryerson, C.C., Martel, C.J., 2002. Army Aircraft Icing. Cold Regions Research and Engineering Laboratory ERDC/CRREL TR-02-13.
Petrenko, V.F., Sullivan, C.R., Kozlyuk, V., 2011. Variable-resistance conductors (VRC) for power-line de-icing. Cold Reg. Sci. Technol. 65, 23–28.
Ronalds, B., Anthony, N., Tuty, S., Fakas, E., 2003. Structural reliability of monopods under storm overload. J. Offshore Mech. Arct. Eng. 125, 114–118.
Ryerson, C.C., 2009. Assessment of Superstructure Ice Protection as Applied to Offshore Oil Operations Safety. Engineer Research And Development Center Hanover Nh Cold Regions Research and Engineering Lab ERDC/CRRELTR-09-4.
Ryerson, C.C., 2011. Ice protection of offshore platforms. Cold Reg. Sci. Technol. 65, 97–110.
SAE, 2002. Aircraft Inflight Icing Terminology. Aerospace Information Report, AIR45504. SAE, Society of Automotive Engineers, Warrendale PA Report AIR45504.
Seadrill, 2015. West hercules. Available http://www.seadrill.com/media-centre/image-library/our-rigs/west-hercules.aspx (18th Sep.).
Shi, X., Veneziano, D., Xie, N., Gong, J., 2013. Use of chloride-based ice control products for sustainable winter maintenance: a balanced perspective. Cold Reg. Sci. Technol. 86, 104–112.
Statoil, 2012. Major new oil discovery in the Barents Sea. Available http://www.statoil.com/en/NewsAndMedia/News/2012/Pages/08Jan_Havis.aspx (18 Sep. 2015).
Tammelin, B., 2005. Wind Turbines in Icing Environment: Improvement of Tools for Siting, Certification and Operation, New Icetools. Finnish Meteorological Institute.
Thomson, J., 2015. High Integrity Systems and Safety Management in Hazardous Industries: Butterworth–Heinemann.
Timco, G., Dickins, D., 2005. Environment guidelines for EER systems in ice-covered waters. Cold Reg. Sci. Technol. 42, 201–214.
Yun, G., Marsden, A., 2010. Methodology for estimating probability of success of escape, evacuation, and rescue (EER) strategies for arctic offshore facilities. Cold Reg. Sci. Technol. 61, 107–115.

Paper B

Zaki, Rezgar; Barabadi, Abbas; Barabady, Javad; Nouri, Ali. Observed and unobserved heterogeneity in failure data analysis: A case study. Submitted to Part O: Journal of Risk and Reliability. Under review

Observed and Unobserved Heterogeneity in Failure Data Analysis

Rezgar Zaki[1], Abbas Barabadi[1], Javad Barabadi[1], Ali Nouri Qarahasanlou*[2]

[1]Department of Engineering and Safety, The Arctic University of Norway (UiT), Tromsø, Norway
[2] Faculty of Technical & Engineering, Imam Khomeini International University, Qazvin, Iran

Abstract: The results of reliability analysis for heterogeneous data can differ substantially from those in a homogeneous case. Covariates can introduce observed and unobserved heterogeneity among data failures collected from specific equipment working at different locations under various operational and environmental conditions (e.g. operator skill, maintenance strategies, low temperature, etc.). In most reliability studies, observed heterogeneity due to observed covariates is discussed. However, unobserved heterogeneity for unobserved reasons, which may have a significant impact on reliability, is neglected. This can lead to erroneous model selection for the time to failure of the item, as well as wrong conclusions and decisions. A systematic approach, to model the unobserved covariate in the area of reliability analysis, is lacking. In this study, the required statistical tests and available models for observed and unobserved heterogeneity in the reliability analysis of failure data are reviewed; a methodology is then developed to facilitate the application of these models. The methodology is based on the mixed proportional hazards model and its extension, which provides an appropriate tool for modeling observed and unobserved heterogeneity under the different types of maintenance strategies. In the second part of the study, the application of the proposed methodology is shown by the investigation of observed and unobserved heterogeneity in the failure data of the chain part from three excavators in service at the Golgohar Sirjan Iron Mine in Iran.

Keywords: Reliability, Observed Covariate, Unobserved Covariate, Gamma Frailty Models, Mixed Proportional Hazard

1- Introduction

In many reliability studies, data sets are assumed to be homogeneous, where the failure data are independent and identically distributed (1–3). However, in reality, they are often working at different locations under various operational and environmental conditions (e.g. operator skill, maintenance strategies, low temperature, etc.) (4). This may introduce heterogeneity into the data (5,6). In general, differences in failure intensity are called heterogeneities and can be due to either observed or unobserved influence risk factors, which are called covariates (7–9). Covariates describe the item's characteristics or the environment in which the item operates (10), and they may have different levels. For example, as a covariate on the reliability of a pump, vibration may be of high, low or medium levels (11). Observed covariates may have different levels and effects, and they are recorded with the failure data. They can be time-dependent or time-independent. Time-dependent observed covariates vary continuously with time. Unobserved covariates are independent variables that may have a significant impact on the failure time of equipment; however, they are not available in the failure database (12).

*- E-mail address: Alinoorimine@gmail.com (Phone: +989149405561)

Unobserved covariates may lead to unobserved heterogeneity (11,13,14). For example, in a production process, some pumps may have a soft foot problem, due to a defect in the installation process. The soft foot problem will put the bearing in an over-stressed situation; this should be considered a covariate for reliability analysis. In the case that there is no information regarding soft foot in the failure database of the bearing, an unobserved covariate should be defined, to capture the effect of soft foot on the reliability of the bearing. In general, due to the quality of manufacturing, installation, operation and maintenance procedures, some items may become frailer, while others are more robust. In the presence of unobserved covariates, different items may have different levels of frailty. Unobserved covariates are typically unknown or not available for each item; hence, they cannot be explicitly included in the analysis. The result of our literature review revealed that, in many cases, unobserved covariates are eliminated during the failure data analysis (7,9,11,12,15). However, if unobserved covariates are neglected, the result of the reliability analysis only represents the reliability of items with an average level of frailty and not that of the individual items. High-risk items (high frailty) tend to fail earlier than low-risk items (low frailty) for unobserved reasons, and, thus, the population composition changes over time. Hence, in time, the analysis represents the item with low frailty, and the estimated reliability increases more with time than the reliability of a randomly selected item of the population (16,17).

The Cox regression model family, such as the proportional hazards model (PHM) and its extension, is the dominant statistical approach for capturing the effect of covariates on the reliability performance of an item (4,11,15,18–22). In PHM, the hazard rate of an item is the product of a baseline hazard rate and a positive functional term that describes how the hazard rate changes as a function of covariates. However, the PHM is very sensitive to the omission of the covariates and is unable to isolate the effect of unobserved covariates (11). The frailty model introduced by Clayton (23) and Vaupel et al. (14) is used to describe the influence of unobserved covariates in a proportional hazards model. A frailty model is a random effects model for time variables, where the random effect (the frailty) has a multiplicative effect on the hazard (13,21). Gamma distribution, inverse Gaussian or exponential distribution can be used to model the frailty (4,12,13,24).

Recently, some studies in the reliability field have used the frailty model to model the effect of missing covariates on the reliability of an item (4,7,22,25). However, a few of these are related to the application of a frailty model in reliability engineering with a focus on maintenance purposes. Asha et al. (26) incorporated the frailty model into load share systems and described the effect of observed and unobserved covariates on the reliability analysis. Xu and Li (27) obtained the stochastic properties of univariate frailty models, which are a special case of multivariate frailty models, and Misra et al.(28) used stochastic orders to compare frailty models arising from different choices of frailty distribution. Giorgio et al. (9) applied the model to a real set of failure time data of powertrain systems mounted on 33 buses, employed on urban and suburban routes in Italy. Slimacek and Lindqvist (29) implemented a frailty model and Poisson process to show unobserved covariates' effect on the reliability of wind turbines. Finkelstein (30) studied the ability to survive a single shock and the intensity of these shocks in time on system reliability. He noted that heterogeneity is a natural feature in many populations, and the frailty model gives an appropriate tool and flexible way to describe lifetimes.

However, these studies have not discussed how time-dependent covariates should be handled in the frailty model. Moreover, the required statistical tests for the investigation of observed and unobserved heterogeneity among the failure data are not discussed. The lack of both strong statistical knowledge among the analysts and systematic methodology are the main challenges for the effective application of a frailty model. To overcome these challenges, the main contribution of this paper is to present a systematic methodology for failure data analysis in the presence of unobserved and observed covariates. Moreover, the paper has reviewed the different models and statistical tests which are needed for an effective analysis of observed and unobserved heterogeneity. This paper is organized as follows. In Section 2, the basic concept is presented. Section 3 describes the proposed methodology. Thereafter, the application of the methodology is illustrated in the reliability analysis of a mining excavator in Section 4. Finally, Section 5 provides the conclusions.

2- Basic concept

The Cox regression models for reliability analysis considering the effect of covariates can be categorized into two main families: *i)* the mixed proportional hazard model family and *ii)* the proportional hazard model family. In these models, the hazard rate of an item is the product of a baseline hazard rate and a positive functional term that describes how the hazard rate changes as a function of unobserved and observed covariates. The baseline hazard rate is assumed to be identical and equal to the total hazard rate, when the observed and unobserved covariates have no influence on the failure pattern (11). The family of mixed proportional hazard models is able to handle the effect of unobserved covariates.

2-1- Mixed proportional hazard model (MPHM) family

In the mixed proportional hazards model, the baseline hazard acts multiplicatively on the *i)* observed covariate function $\psi(z;\eta)$ and *ii)* a time-independent frailty function α_j. Suppose we have a fleet of *j* items, the hazard function for an item at time $t > 0$ is:

$$\lambda_j(t;z;\alpha) = \alpha_j \lambda_0(t)\psi(z;\eta) \tag{1}$$

where $\lambda_0(t)$ is an arbitrary baseline hazard rate, dependent on time alone, z is a row vector consisting of the observed covariates associated with the item, η is a column vector consisting of the regression parameters for identified observed covariates, and α_j is a time-independent frailty function for item j and represents the cumulative effect of one or more unobserved covariates. The baseline hazard rate ($\lambda_0(t)$) may be either left unspecified or modeled using a specific parametric form such as Weibull distribution or Non-Homogeneous Poisson Process (NHPP).

According to the mixed proportional hazards model, the fleet of items (the population) is represented as a mixture, in which the $\lambda_0(t)$ and $\psi(z;\eta)$ are common to all items, although each item has its own frailty. The observed and unobserved covariates can affect the hazard rate, so that the actual hazard rate ($\lambda_j(t;z;\alpha)$) is either greater (e.g. in the case of higher vibration level or poor maintenance) or less (e.g. better training for operators, installation of a new ventilation system) than the baseline hazard rate. Moreover, equipment with $\alpha_j > 1$ is frailer and may have decreased time to failure. The items for which $\alpha_j < 1$ are less frail and tend to be more reliable.

Different functional forms of $\psi(z;\eta)$ and α_j may be used to model the observed and unobserved covariate functions. For example, the exponential form $exp(z\eta)$, the logistic form $log(1+\exp(z\eta))$, the inverse linear form $1/(1+z\eta)$, and the linear form $(1+z\eta)$ are some of the functions used for observed covariate functions (11,31,32). Moreover, gamma, inverse Gaussian and exponential distributions are used to model the frailty function (21,33,34).

Generally, the exponential distribution for $\psi(z;\eta)$ and gamma distribution, with the mean equal to one and variance of θ, are the most used functions for observed and unobserved distributions, respectively.

In Eq. (1), the assumption is that all covariates are time-independent. However, in reality, there are many cases where the covariates are time-dependent. It is of great practical importance to decide whether covariate effects are constant over time or their effects change (35). For example, a covariate that is used to represent crack growth may change over the operational time of the item. Such a covariate is time-dependent (11); hence, it should be modeled as a time-dependent covariate by using the crack propagation geometry. The hazard rate of an item in the presence of time-dependent covariates ($z(t)$) takes the following form (12):

$$\lambda_j(t;z;z(t);\alpha) = \alpha_j.\lambda_0(t)\psi(z,z(t);\eta;\delta) \tag{2}$$

where $z(t)$ is a row vector consisting of the observed time-dependent covariates associated with the item (e.g. ambient temperature, pressure on the failure time, etc.), η and δ are the corresponding regression coefficients (i.e. the effect size) of time-dependent and time-independent observed covariates. As Eq. (2) is an extension of Eq. (1) in this paper, Eq. (2) is named an extension mixed proportional hazards model (EMPHM).

Considering gamma distribution (with the mean to one and variance θ) for unobserved covariates and exponential function for observed covariates, the reliability function can be written as (12):

$$R_\theta(t;z;z(t)) = [1-\theta\ln\left(R_i(t;z;z(t))\right]^{-1/\theta} \tag{3}$$

where $R_i(t;z;z(t))$ is the item's reliability function and considering the existence of p_1 time-independent observed covariates and p_2 time-dependent observed covariates. It can be estimated by:

$$R_i(t;z;z(t)) = \left[exp\left(-\int_0^t \lambda_0(x)exp\left[\sum_{j=1}^{p_2}\delta_{sj}z_{sj}(t)\right]dx\right)\right]^{\exp[\sum_{i=1}^{p_1}\eta_{si}z_{si}]} \tag{4}$$

If all the covariates are time-independent, Eq. (4) reduces to:

$$R_i(t;z;z(t)) = R_0(t)^{\exp[\sum_{i=1}^{p_1}\eta_{si}z_{si}]} \tag{5}$$

where $R_0(t)$ is the baseline reliability function, and η_i represents regression parameters for *i* time-independent covariates (z_i). Thus, Eq. (3) can be written as:

$$R_\theta(t;z;z(t);\alpha) = \left[1-\theta\ln\left(R_0(t)^{\exp[\sum_{i=1}^{p_1}\eta_{si}z_{si}]}\right]^{-1/\theta} \tag{6}$$

which is named the mixed proportional hazard model (MPHM).

2-1-1- Shared frailty model

In some cases, a group of items share the same frailty value (36). For example, consider a company where identified excavators are utilized in two different shifts: night and day. Here, the shift can be considered a shared frailty. A shared frailty is a group-specific latent random effect that multiplies into the hazard function and will generate dependence between those items which share frailties (12). The distribution of the shared frailty is gamma, with mean 1 and variance to be estimated from the data. For data consisting of n groups, with the sth group comprised of n_s items *(s = 1,...,n)*, the shared frailty model can be written as:

$$\lambda(t; z; z(t); \alpha) = \alpha\lambda_{0s}(t)\exp\left[\sum_{i=1}^{p_1} \eta_{si} z_{si} + \sum_{j=1}^{p_2} \delta_{sj} z_{sj}(t)\right] \tag{7}$$

That is, for any member of the ith group, the standard hazard function is now multiplied by the shared frailty, α_s. In the case that there are no time-dependent covariates, Eq. (7) will reduce to:

$$\lambda(t; z; z(t); \alpha) = \alpha\lambda_{0s}(t)\exp\left[\sum_{i=1}^{p_1} \eta_{si} z_{si}\right] \tag{8}$$

2-1-2- Stratification approach model

In the case of the existence of time-dependent covariates ($z_j(t)$), the stratification approach can be used (15). In this approach, it is possible to classify a categorical or time-dependent covariate with several categories into different strata with different baseline hazards for each category. For example, in the case of modeling the effect of ambient temperature on the reliability of a pump which is installed outside, then the collected failure data can be categorized into four groups, based on the seasons (spring, summer, fall and winter). Figure 1 shows a graphical representation of this example, where i is the number of failures occurring within each stratum.

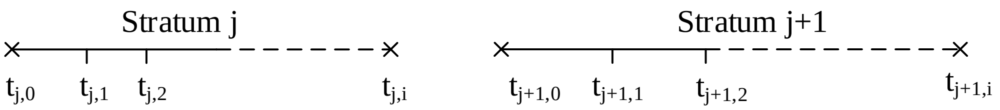


Figure 1: A graphical representation of the strata for a data set

In the stratification approach, the baseline hazard function differs for defined strata, but the regression coefficients are the same for all covariates. Hence, the hazard rate for the system in stratum r in the presence of unobserved covariates will be:

$$\lambda_r(t; z; \alpha) = \alpha_r.\lambda_{0r}(t)\exp\left[\sum_{i=1}^{p_1} \eta_i z_i\right] \tag{9}$$

where λ_{0r} is the baseline hazard for stratum r, and, if there is no unobserved covariate, then Eq. (9) will be reduced to:

$$\lambda_r(t; z) = \lambda_{0r}(t)\exp\left[\sum_{i=1}^{p_1} \eta_i z_i\right] \tag{10}$$

2-2- Proportional hazard model family

The main assumption in the proportional hazard model family is that all influence covariates are identified and there is no omission of covariates (no unobserved covariates). In the case of there being no unobserved covariates' effect on the hazard rate of the item, then, in Eq. (3), gamma distribution will be equal to 1, and the successive equation is:

$$\frac{\alpha^{\frac{1}{\theta}-1}e^{-\frac{\alpha}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)\theta^{\frac{1}{\theta}}}=1 \tag{11}$$

and the hazard rate can be written as:

$$\lambda(t;z;z(t))=\lambda_0(t)\exp\left[\sum_{i=1}^{p_1}\eta_i z_i+\sum_{j=1}^{p_2}\delta_j z_j(t)\right] \tag{12}$$

In the literature, this model is mainly referred to as an extension of the proportional hazard model (EPHM) (15,37). Proportionality assumption implies that the effect of a covariate is independent of time and the ratio of any two hazard rates is constant with respect to time, i.e.:

$$\frac{\lambda_1(t;z_1;\alpha)}{\lambda_2(t;z_2;\alpha)}=\frac{\lambda_0(t)\exp\left(\eta_1 z_1\right)}{\lambda_0(t)\exp\left(\eta_2 z_2\right)}=\exp(\eta_1 z_1-\eta_2 z_2)=Constant \tag{13}$$

where $z_1$and z_2 are any two different sets of time-independent observed covariates assumed to be associated with the item. If there is no time-dependent covariate, then Eq.(14) will reduce to the PHM, as follows:

$$\lambda(t;z)=\lambda_0(t)\exp\left[\sum_{i=1}^{p_1}\eta_i z_i\right] \tag{14}$$

2-3- Parameter estimation

For EMPHM, given the relationship between the hazard rate and the reliability functions, it can be shown that the conditional (item) reliability function, $R(t;z;z(t)|\alpha)$, conditional on the frailty α, is (12):

$$R(t;z;z(t)|\alpha)=\{R(t;z;z(t))\}^{\alpha} \tag{15}$$

The unconditional (population) reliability function can then be estimated by integrating out the unobserved α. If α has probability density function *g(α)*, then the population or unconditional reliability function is given by:

$$R_\theta(t;z;z(t))=\int_0^\infty\{R(t;z;z(t))\}^{\alpha}g(\alpha)d\alpha \tag{16}$$

where we use the subscript θ to emphasize the dependence on the frailty variance θ. The relationship between the reliability function and the hazard function still holds unconditional on α, and, thus, we can obtain the population hazard function using (12):

$$\lambda_\theta\big(t;z;z(t)\big)=-\frac{d}{dt}R_\theta(t;z;z(t))[R_\theta(t;z;z(t))]^{-1} \tag{17}$$

Having the gamma distribution with unobserved covariates (12):

$$R_\theta(t; z; z(t)) = [1 - \theta ln\{R(t; z; z(t))\}]^{-1/\theta} \tag{18}$$

Having the event times (t_{0i}, t_i, d_i) for $i = 1, , n$ with the ith observation corresponding to the time span $(t_{0i}, t_i]$, with either the failure occurring at time t_i $(d_i = 1)$ or the failure time being right-censored at time t_i $(d_i = 0)$, the likelihood function for reliability data is given by:

$$LnL = ln \prod_{i=1}^{n} \frac{\{R_{\theta i}(t_{0i}, z_i, z_i(t)\}^{1-d_i}\{f_{\theta i}(t_i, z_i, z_i(t)\}^{d_i}}{R_{\theta i}(t_i, z_i, z_i(t))} \tag{19}$$

where, $f_{\theta i}$ is the probability density function.

In a shared frailty model for failure data of unrepairable units, suppose we have data for $i = 1,...,n$ groups, with $j = 1,...,n_i$ observations per group, consisting of the trivariate response $(t_{0ij}, t_{ij}, d_{ij})$, which indicates the start time, end time, and failure/censoring for the jth item from the ith group, while the shared frailties follow a gamma distribution, L_i can be expressed compactly as (12):

$$L_i = \left[\prod_{j=1}^{n_i}\{\lambda_{ij}(t_{ij})\}^{d_{ij}}\right]\frac{\Gamma(1/_\theta + D_i)}{\Gamma(1/_\theta)}\theta^{D_i}\left\{1 - \theta\sum_{j=1}^{n_i} Ln\frac{R_{ij}(t_{ij})}{R_{ij}(t_{0ij})}\right\}^{-1/_\theta + D_i} \tag{20}$$

where $D_i = \sum_{j=1}^{n_i} d_{ij}$. Given the unconditional group likelihoods, we can estimate the regression parameters and frailty variance θ, by maximizing the overall log-likelihood $LnL = \sum_{i=1}^{n} \ln L_i$. In shared-frailty Cox models, the estimation consists of two steps. In the first step, the optimization is in terms of θ alone. For fixed θ, the second step consists of fitting a standard Cox model via penalized log-likelihood, with the v_i introduced as estimable coefficients of dummy variables identifying the groups. The same approach can be used to estimate the likelihood functions for EPHM, MPHM and PHM. For more information, see (7,11,12).

In a minimally repaired system (NHPP), the times between failures (TBF) are not independent and identically distributed random variables (except for the special case of constant failure intensity, that is, of homogeneous Poisson process), and the log-likelihood function relative to "m" minimally repaired systems, whose failure intensity is given by Eq. (2), results in (12):

$$lnL = ln\left[\prod_{j=1}^{m}\int g(\alpha)\left(\prod_{i=1}^{n_j} f(t_{i,j}|t_{i-1,j}; z, z(t), \alpha)\right)\frac{R(T_j; z, z(t), \alpha)}{R\left(t_{n_j,j}; z, z(t), \alpha\right)}d\alpha\right] \tag{21}$$

where $g(\alpha)$ denotes the probability density function (PDF) of the frailty parameter α, n_j is the number of observed failures of the j-th system, $t_{i,j}(i = 1, \dots, n_j; j = 1, \dots, m)$ is the i-th failure time of the j-th system observed up to T_j, $t_{o,j} = 0$. Conditional PDF of the failure time $t_{i,j}$, present the previous failure time $t_{i-1,j}$, as (12):

$$f_T(t_{i,j}|t_{i-1,j}; z, z(t), \alpha) = \alpha\lambda_0(t_{i,j})\psi(z, z(t); \eta; \delta)\frac{R(t_{i,j}; z, z(t), \alpha)}{R(t_{i-1,,j}; z, z(t), \alpha)} \tag{22}$$

The reliability function, $R(t; z; z(t)|\alpha)$, conditional on the frailty, α, is (12):

$$R(t; z; z(t)|\alpha) = \left[exp\left(-\int_0^t \lambda_0(t_{i,j})\psi(z, z(t); \eta; \delta)dx\right)\right]^\alpha \tag{23}$$

The log-likelihood function, relative to "m" repairable systems subject to perfect repairs (renewal process), whose hazard function is given by Eq. (2), results in:

$$lnL = ln\left[\prod_{j=1}^{m}\int g(\alpha)\left(\prod_{i=1}^{n_j} f(t_{i,j}|t_{i-1,j}; z, z(t), \alpha)\right)R(X_j; z, z(X_j), \alpha)d\alpha\right] \tag{24}$$

where, here, $x_{i,j} = t_{i,j} - t_{i-1,j}$ $(i = 1, \dots, n_j; j = 1, \dots, m)$ is the i-th time between failures (TBF) of the j-th system, $X_j = T_j - t_{n_j,j}$, and the (unconditional) pdf of the TBF $x_{i,j}$ is:

$$f_X(x_{i,j}; z, z(t), \alpha) = \alpha\lambda_0(x_{ij})\psi(z, z(x_{ij}); \eta; \delta)R(x_{ij}; z, z(t), \alpha) \tag{25}$$

Only in the absence of unobservable heterogeneity does the log-likelihood function in Eq. (24) reduce to the log-likelihood in Eq. (19). Indeed, in the presence of unobserved heterogeneity, the same value of the frailty variable α characterizes the whole path of each repairable system, so that the whole conditional likelihood function $L_j|\alpha$ of each system "j", given α, must be multiplied by $g(\alpha)$ and hence integrated on α. In addition, it must be mentioned that the log-likelihood function in Eq. (24) becomes much more complex when the j-th system is observed to start from a generic time which is not a failure time.

3- Proposed framework

The systematic framework for reliability analysis in the presence of observed and unobserved covariates (or heterogeneity) is described in Figure 2. This methodology is based on four important steps:

- Establishing the context and data collection
- Identifying the baseline hazard rate, based on maintenance nature
- Modeling the effect of the covariates
- Parameter estimation

As Figure 2 shows, in the first step, the context should be established. All external and internal parameters to be considered when analyzing failure data and setting the scope and assumptions for the reliability analysis should also be defined. External context is the external environment in which the item will be working, such as ambient temperature, pressure, humidity, etc. Internal context is the internal conditions related to the item itself and the company’s running and maintenance of the item, including the repair and physics of failure, operator condition, maintenance crew, etc. Understanding the external and internal contexts is important, in order to identify the observed covariates. For example, based on the physics of failure, road condition can contribute to the failure of an excavator in a mine; hence, it should be considered a covariate in the reliability analysis of the excavator. In this step, the possible relationships between different covariates should be investigated, as well as the stratum for each of them.

In the next step, failure data and all possible observed covariates associated with each failure should be collected.

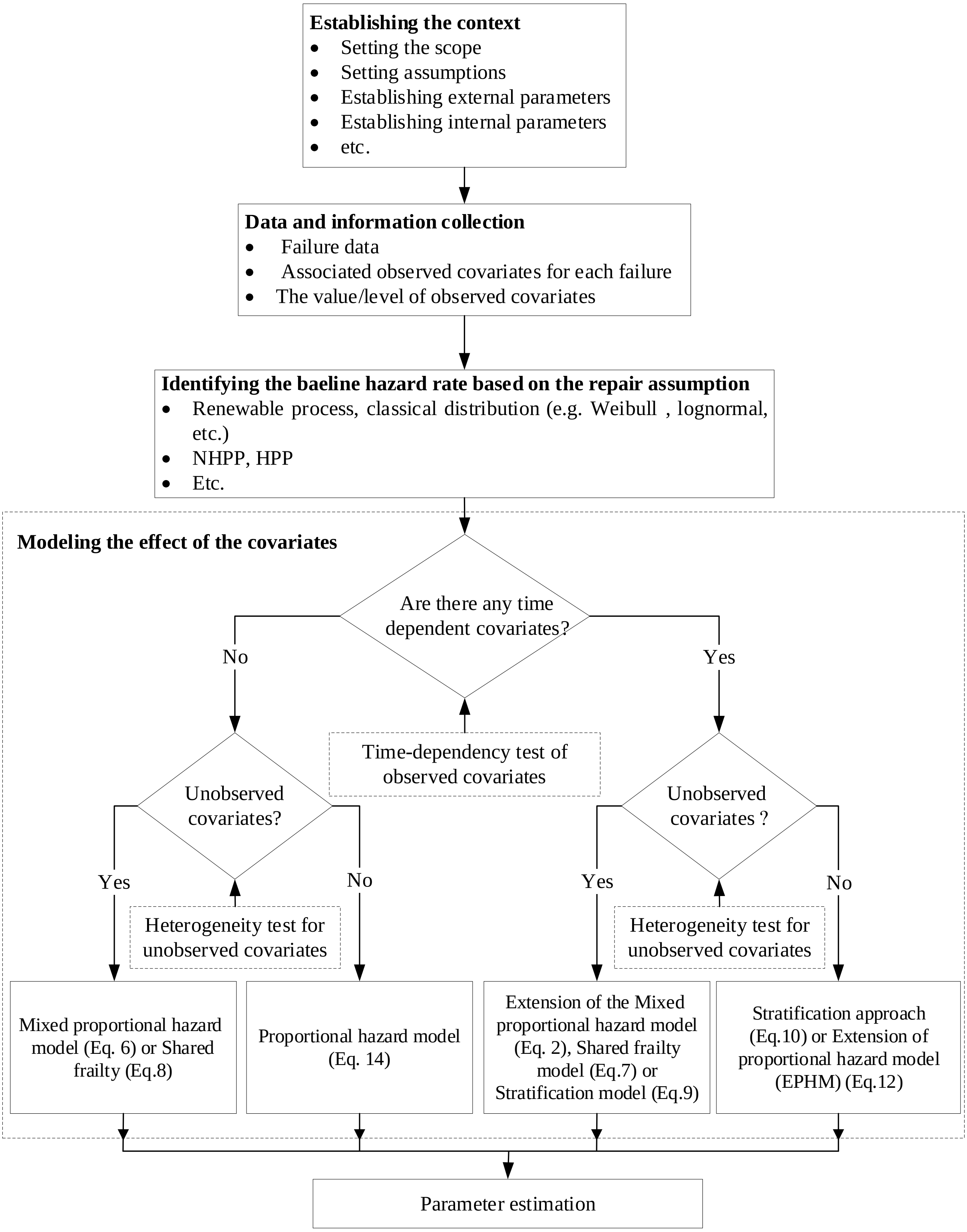


Figure 2: A framework for reliability model selection in the presence of observed and unobserved covariates

Thereafter, based on the nature of the failure data (e.g. trend behavior of the data) and the type of repair strategy, the appropriate baseline hazard should be selected for the data. For example, the common assumption for a repairable system can be i) perfect repair or good-as-new condition, ii) minimal repair or bad-as-old condition, or iii) jumps in the hazard rate after repair or different baseline hazard rate. Under the perfect repair strategy, the item is restored to as ‘good-as-new’ condition, and the main assumption is that the hazard rate is reset to that of a new system after maintenance. If the times between failures are independent and identically distributed (iid), it can be concluded that the item underwent perfect repair (21). In such cases, classical distribution, such as Weibull distribution, can be used to model the baseline hazard rate.

In the case of minimal repair (bad-as-old), an item has the same intensity function after repair as before the failure. The failure times when minimal repair is carried out can be thought of as a non-homogeneous Poisson process (NHPP). In other words, the baseline hazard rate will be modeled using a non-homogeneous Poisson model. However, it should be mentioned that, on some occasions, such as overhaul, the system may return to as ‘good-as-new’ condition. Under this condition, it is assumed that the NHPP is cyclic, with each cycle starting as a renewal process and, within the cycle, failure times follow the NHPP. In this case, the failure data will then be categorized by these occasions (for example, overhaul), and then a stratification approach is used to estimate the effect of each covariate, while the baseline hazard rate is modeled by an NHPP model. However, when a fleet of items is analyzed, after some time and undergoing several repairs, the baseline hazard rate will change. For example, in some cases, as the number of failures increases, the average failure time decreases; hence, the baseline hazard rate will not be identical for a particular failure number. Here, the failure data can be categorized based on the failure number; it can be used to define strata, and then the stratification approach can be used to model the fleet failure data.

In general, the first step in analyzing the collected failure data of a repairable system is to check the trend of the failure data. In the case of the data showing a trend, the NHPP or trend renewal process (TRP) can be used to model the baseline hazard rate. However, when there is no trend in the data, classical distribution, such as Weibull distribution, can be used to model the baseline hazard rate. However, some goodness-of-fit test, such as residual testing, should be used to find the best fit distribution for failure data. For more information regarding the trend test, see (7).

In the next step, the time dependency of observed covariates should be checked. Later, the failure data need to be investigated for unobserved covariates. Data sets without unobserved heterogeneity will be analyzed using the classical proportional hazards model, including the proportional hazards model (when all observed covariates are time-independent) and the extension of the proportional hazards model (in the presence of time-dependent covariates). Moreover, data sets with unobserved heterogeneity will be analyzed using the mixed proportional hazards model family. Also, the parameters of models could be estimated by different likelihoods for repairable and unrepairable systems (Eqs. (21) and (23)).

3-1- Time-dependency test of observed covariates

There are two general approaches for checking the time dependency of covariates: *i)* the graphical procedure and *ii)* the goodness-of-fit testing procedure (15). The developed graphical procedure can generally be categorized into three main groups: *i)* cumulative hazards plots, *ii)* average hazards plots and *iii)* residual plots (11). For example, in cumulative hazards plots, the data will be categorized according to the different level of the covariate that is to be checked for time dependency. Consider that a covariate can be categorized into r levels, in which the covariate is equal to z_r. Thereafter, the hazard rate can be written as:

$$\lambda_r(t;z;\alpha) = \alpha_s.\lambda_{0r}(t)\exp\left[\sum_{i=1}^{p_1}\eta_i z_i\right] \tag{26}$$

where $\eta_i z_i$ is the same as before, with $\eta_r z_r$ omitted, with $i=1,2,\ldots p_1$ and $j\neq r$. If the PH assumption is justified, then we will end up with:

$$\lambda_{0r}(t) = C_r\lambda_0(t), and\ C_r = \alpha_s \exp\,(\eta_r z_r) \tag{27}$$

A similar relation can be concluded for the cumulative baseline hazard rate. Hence, if the assumption of PH is justified, then the plots of the logarithm of the estimated cumulative baseline hazard rates against time for defined categories should simply be shifted by an additive constant, η_r. In other words, they should be approximately parallel and separated, corresponding to the different values of the covariates. Departure from parallelism of the above plots for different categories may suggest that z_r is a time-dependent covariate. For a review of other graphical approaches, see (11,15,38–40).

In the same way as the cumulative baseline hazard rate, a log–log Kaplan-Meier curve over different (combinations of) categories of variables can be used to check the assumption of PH. A log–log reliability curve is simply a transformation of an estimated reliability curve that results from taking the natural log of an estimated reliability probability twice. If we use a PHM or MPHM and plot the estimated log–log reliability curves for defined categories on the same graph, the two plots would be approximately parallel (11). In the residuals plot in the first step, the residual should be estimated by using the estimated values of the cumulative hazard rate, $H_0(t_i)$, and the regression vector η as:

$$e_i = -H_0(t_i)\exp\,(\eta_r z_r) \tag{28}$$

If the PH assumption is justified, then the logarithm of the estimated reliability function of e_i against the residuals should lie approximately on a straight line with slope -1 (11,41). A transformed plot of the partial residual suggested by Schoenfeld can also be used as an exploratory tool to detect the time-varying effects of a covariate, even when the a priori form of time dependence is unknown (42–44). The Schoenfeld Residuals Test is analogous with testing whether the slope of the scaled residuals on time is zero or not. If the slope is not zero then the proportional hazard assumption has been violated (44). When the covariates are quantitative, using graphical approaches is challenging, as it is difficult both to define different levels for quantitative covariates and to decide whether the plots are parallel or not. In such cases, it is better to use a goodness-of-fit testing procedure, such as the chi-squared goodness-of-fit test (3,45,46), the log rank test (3,45), the likelihood ratio test (3,45), score tests (45,47), the doubly cumulative hazard function (48), the Wilcoxon test (49) and generalized moments specification tests (50). For example, if the PH assumption is justified, the different two-sample tests, e.g. generalized Wilcoxon and log rank tests, should have the same results (11).

3-2- Heterogeneity test for unobserved covariates

Several statistical tests are available in the literature for identifying and quantifying the effects of unobserved heterogeneity. For example, Kimber (51) developed a Weibull-based score test for heterogeneity and then demonstrated its application in two case studies on infant nutrition. Under the assumption that the data follow a stratified proportional hazards model, where the hazard rate can differ within different strata, Gray (52) used the martingale residuals to test for variation over groups in reliability data. Commenges and Andersen (53) used marginal partial likelihood to develop a score test of homogeneity for reliability data, when the frailty model is used to model the covariates. The score test is valid for general distributions of the frailty variable, not only for the frequently used gamma distribution. In the meta-analysis, Cochran's Q test (Q test) is normally used to check the homogeneity among data sets. However, the Q test only checks the presence versus the absence of heterogeneity; it does not report on the extent of such heterogeneity. However, these statistical tests and their applications are limited, mainly due to their requirements, in terms of data and assumptions. Each test is the optimum one for detecting the heterogeneity of a specific form (54,55). For example, a shortcoming of the Q statistic is that it has poor power to detect true heterogeneity, among studies when the meta-analysis includes a small number of studies, and excessive power to detect negligible variability with a high number of studies. Recently, the I^2 index has been proposed to quantify the degree of heterogeneity in a meta-analysis (56). The likelihood ratio test, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are common tests for checking the hypothesis of the presence of heterogeneity against the null hypothesis of non-heterogeneity ($\hat{\theta} = 0$). In general, the AIC performs well when heterogeneity is small, but, if heterogeneity is large, the BIC will often perform better (7,10,57). For example, in the case of Weibull distribution for the baseline hazard rate, likelihood ratio can be written as:

$$R_H = 2\left(\ln L(\hat{\lambda}, \hat{\beta}, \hat{\eta}, \hat{\theta}) - lnL(\hat{\lambda}_0, \hat{\beta}_0, \hat{\eta}_0, 0)\right) \tag{29}$$

Here, $\hat{\lambda}$ and $\hat{\beta}$ are estimated parameters for Weibull distribution, $\hat{\eta}$ is the regression coefficient for observed covariates, and $\hat{\theta}$ can be interpreted as the degree of heterogeneity (7). These parameters can be estimated by maximizing the full likelihood function. On a 5% significance level, the null hypothesis (no heterogeneity) will be rejected if $R \geq 2.706$. Moreover, under the minimal repair strategy, a power law can be used to represent the intensity function. Under the assumption of the power law intensity function, in order to check whether a significant amount of heterogeneity among units exists, a three-step likelihood ratio test procedure can be performed (7). As the first step, the null hypothesis, say $H_0: \lambda_1 = \lambda_2,\ \lambda_m = \lambda_0\ , \beta_1 = \beta_2,\ \beta_m = \beta_0$, should be tested against the alternative hypothesis, $H_1: \lambda_1 \neq \lambda_2,\ \lambda_m \neq \lambda_0\ , \beta_1 \neq \beta_2,\ \beta_m \neq \beta_0$. In the second step, common λ and uncommon β and third steps, uncommon λ, common β should be carried out (9).

4- Mixed proportional hazards modeling of unrepairable systems: Case study

As one of the most important machines in the mine, the excavator needs to have a strong undercarriage and chain to provide excellent reliability and durability while working on rocky ground or blasted rock. Thus, the case study refers to chain failure data of three Caterpillar 390DL excavators in service in the Golgohar Sirjan Iron Mine in Iran over two years. This mine is located in the southwest of Kerman Province, Iran, and it contains six ore bodies, named 1, 2, 3, 4, 5 and 6, spread over an area of 40 km^2. The main design characteristics (weight, size, maximum capacity, etc.) of the excavators are identical. The number of observed failures of the excavators, n_i, ranges from 24 to 51 to a total of $n_{T=} \sum_i^4 n_i = 103$.

We used trend tests and serial correlation tests to check the independent assumption and identically distributed (*iid*) assumption in the collected data. The serial correlation tests were performed by T-test (TSTA) and Ljung-Box-Q (LBQ) statistics (58). Trends were assessed by three analytical tests: MIL-Hdbk-189 (MIL), Laplace and Anderson-Darling (A-D) (59). The results of the trend and correlation tests for each Excavator machine are presented in Table 1. Thus, the result for all excavators show that the data were independent and identically distributed. Hence, the classical distribution can be used to model the baseline hazard rate.

Table 1- Trend and serial correlation tests of each system

System	Trend Tests				Serial correlation tests	
	Subject	**MIL**	**Laplace**	**A-D**	**TSTA**	**LBQ**
Excavator 1	Test Statistic	13.74	4.64	13.29	0.86	0.82
	P-Value	0	0	0		
	DF	50				

The Akaike information criterion (AIC) and Bayesian information criterion (BIC) can be used to find the best-fit distribution for the baseline hazard rate (60). The distribution candidate with the smallest AIC and BIC values is the best-fit distribution to model the baseline hazard rate.

According to the framework in Figure 2, in addition to failure data (TBF), all associated observed covariates should be collected. To this aim, the observed covariates should be identified. Table 2 shows the selected observed covariates. As the table shows, five observed covariates are identified which may affect the reliability of the excavators. The numbers in the brackets in Table 2 are used to nominate (formulate) the covariates.

Table 2: The identified observed covariates for the excavators

Covariate	Covariate level	Covariate	Covariate level
Working shift (z_s)	Morning shift [3]	Excavator code (z_e)	Machine No. 1 [1]
	Afternoon shift [2]		Machine No. 2 [2]
	Night shift [1]		Machine No. 3 [3]
Machine movement (z_m)	Small [<=3 m]	Rock type (z_r)	West [1]
	Medium [3-13 m]		Ore [2]
	Large [13 m <]	Precipitation (z_p)	Continuous covariate
		Temperature (z_t)	Continuous covariate

According to the framework in Figure 2, after collecting the data on failures and observed covariates, the time dependency of the covariate should be checked. Here, the graphical approach (a ln–ln reliability curve) is used to check the time dependency of all the covariates. Figure 3 shows the *–ln (-ln reliability)* against the *ln (analysis time)* for an observed covariate: namely, Machine movement (z_m). As this graph shows, the curves are approximately parallel; hence, the assumption of proportionality is correct for the data sets, and it can be concluded that the covariates are time-independent. The *–ln (-ln reliability)* for other covariates confirms the same result.

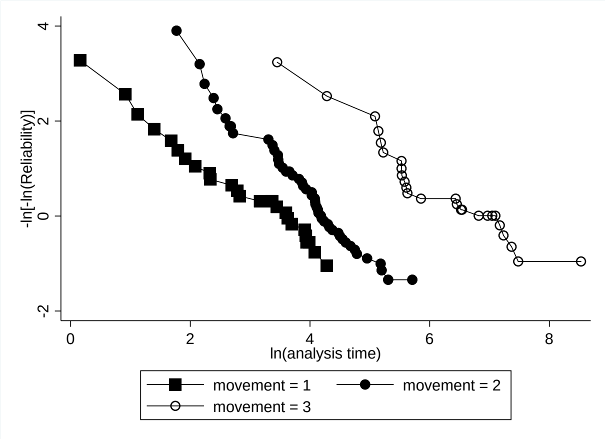


Figure 3: The log minus log graph for time between failures of the excavators, based on movement covariate

An analytical test was used to check the PH assumption in this study. Harrell and Lee's (1986) test is a variation of a test originally proposed by Shenfield (1982) and is based on the residuals (61). The PH testing approach is attractive because it provides a test statistic and p-value (P(PH)) for checking the PH assumption for a given predictor of interest. Thus, a more objective decision is provided by a statistical test rather than a graphical approach. The P(PH) is used for evaluating the PH assumption for that variable. Table 3 shows the statistics of the PH test for all covariates. The P(PH) values for α=0.05 are quite high for all variables satisfying the PH assumption.

Table 3: Analytical test approach results for PH assumption

Covariates	ρ	χ^2	Df.	P(PH)
z_s	-0.05595	0.3	1	0.5842
z_t	-0.05343	0.25	1	0.6182
z_p	0.00836	0	1	0.95
z_r	-0.10143	1.14	1	0.286
z_m	0.01796	0.03	1	0.8742
z_e	-0.00187	0	1	0.9852

In the next step of the framework, the presence of unobserved covariates (heterogeneity test) should be checked. For this, the best-fit distribution for the baseline hazard rate needs to be identified. The AIC and BIC procedures are applied to select the best-fit distribution for the baseline hazard rate, as well as to check the heterogeneity of data. Table 4 shows the values of the AIC and BIC for the different nominated distributions for the baseline hazard rate with the same covariates under two assumptions: *i)* with frailty and *ii)* without frailty. As the result in Table 4 shows, the Weibull MPHM is the most suitable model for the data, as it has the smallest AIC or BIC among all the models. Therefore, the model with unobserved heterogeneity can give a better estimation of the reliability of the excavators.

Moreover, we used the ratio test to also check the unobserved heterogeneity in unrepairable parts. Under the assumption of Weibull MPHM, while *i)* the gamma distribution represents the frailty model (with mean equal to one and variance equal to θ) and *ii)* there are no time-dependent covariates, the hazard rate can be written as:

$$\lambda(t;z;z(t);\alpha) = \frac{\alpha^{\frac{1}{\theta}-1} e^{-\frac{\alpha}{\theta}}}{\Gamma(\frac{1}{\theta})\theta^{\frac{1}{\theta}}}.(mt^{m-1})\exp\left[\sum_{i=1}^{p_1} \eta_i z_i\right] \tag{30}$$

Table 4: Goodness of fit of different reliability models

Model		Observations	d.f.	AIC	BIC	Log likelihood
With frailty	Exponential MPHM	103	8	315.1397	336.2175	-149.570
	Weibull MPHM	103	9	306.9757	330.6882	-144.488
	Gompertz MPHM	103	9	311.9349	335.6475	-146.968
Without frailty	Exponential PHM	103	7	313.3724	331.8155	-149.686
	Weibull PHM	103	8	314.8206	335.8984	-149.410
	Gompertz PHM	103	8	309.9349	331.0128	-146.968

For this, likelihood ratio tests were performed, as below:

$$R_H = 2\left(\left(\ln L(\hat{\lambda},\hat{\beta},\hat{\eta},\hat{\theta}) - lnL(\hat{\lambda}_0,\hat{\beta}_0,\hat{\eta}_0,0\,)\right)\right) = 9.84 \tag{31}$$

The p-value for R_H=9.84 will be equal to 0.001, which hints at the existence of an unobserved covariate's (unobserved heterogeneity) effect on the reliability of the excavators. Hence, the Weibull MPHM should be used to analyze the data. Software is available, which can estimate the parameters in MPHM, such as Stata, R and SAS. Table 5 and Table 6 show the results of the analysis in Stata.

Table 5: The result of covariates' effect analysis under the assumption of MPHM

Covariate	Coef.	Std. Error	Z	P>|Z|	[95% Conf. Interval]	
z_s	0.094	0.287	0.330	0.744	-0.469	0.656
z_t	0.018	0.028	0.630	0.528	-0.037	0.072
z_p	-2.615	9.615	-0.270	0.786	-21.460	16.230
z_r	0.547	0.456	1.200	0.230	-0.347	1.441
z_m	-3.274	0.744	-4.400	0.000	-4.732	-1.816
z_e	0.927	0.339	2.730	0.006	0.262	1.592

Table 6: The constant value, baseline and unobserved parameters estimation of MPHM

Parameters	Coef.	Std. Error	[95% Conf. Interval]	
Constant value	-5.709	1.896	-9.425	-1.994
Baseline Weibull ancillary parameter (m)	1.940	0.403	1.291	2.913
Variance of gamma distribution (θ)	1.576	0.714	0.649	3.829

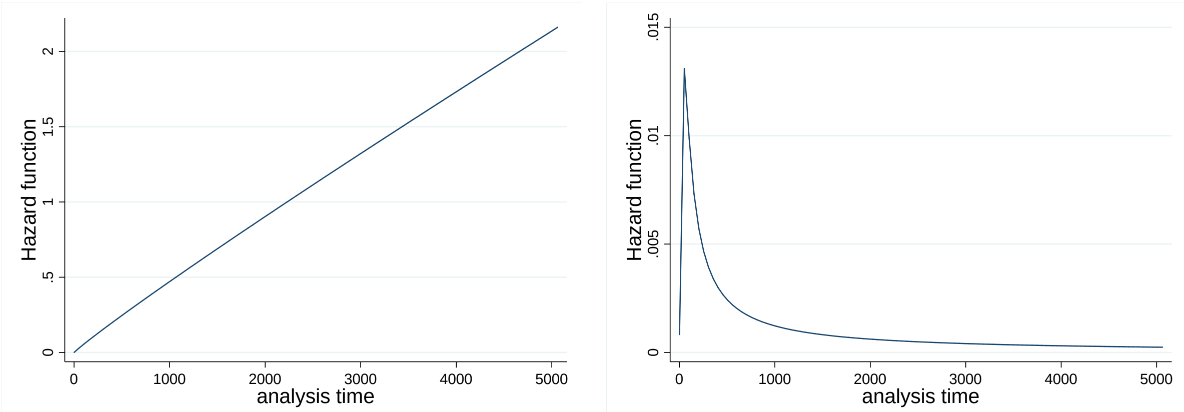
The result of the analysis in Table 7 shows that the constant value is -5.709, and Machine movement (z_m) and Excavator code (z_e) have a significant effect on the excavators' reliability. Based on Eq. (18), the unconditional reliability of Weibull MPHM can be written as:

$$R_\theta(t;z) = \left[1 - \theta ln\left(e^{\left(-(t^m)(\exp[\text{Constant value}+\sum_{i=1}^{p_1}\eta_i z_i])\right)}\right)\right]^{-1/\theta} \quad (32)$$

Having the regression coefficient for covariates, the unconditional reliability of the excavators will be equal to:

$$R_\theta(t;z) = \left[1 - 1.576 ln\left(e^{\left(-(t^{1.94})(\exp(-5.709-3.274z_m+0.937z_e))\right)}\right)\right]^{-1/1.576} \quad (33)$$

Figure 4 shows the unconditional and conditional hazard functions of the excavators. Figure 5(a) is the population hazard function, where the curve is unconditional on the frailty and is "averaged" over the frailty distribution, while Figure 5(b) is the individual hazard function, which is conditional on a frailty value of one $(\alpha_j = 1 \text{ in Eq. } (1))$. As Figure 5 shows, there is a big difference between the hazard rate of the excavators' population and that of an individual excavator.



(a) (b)

Figure 4: a) The unconditional (population) hazard function of the excavators on the mean of covariate, b) the excavator conditional (individual) hazard function on the mean of covariate

In the next step, in order to compare how much bias will be associated with analysis if the effect of unobserved covariates is ignored, analysis is performed on the assumption that there are no effects of unobserved covariates. The result of analysis, using Gompertz-PHM as the selected model, is shown in Table 7 and Table 8.

Table 7: The result of covariates' effect analysis under the assumption of Gompertz-PHM

Covariates	Coef.	Std. Error	z	P>|z|	[95% Conf. Interval]	
z_s	-0.125	0.136	-0.920	0.359	-0.392	0.142
z_t	0.009	0.014	0.620	0.536	-0.019	0.037
z_p	0.884	4.008	0.220	0.825	-6.971	8.739
z_r	0.269	0.252	1.070	0.285	-0.224	0.762
z_m	-1.606	0.172	-9.320	0.000	-1.943	-1.268

z_e	0.365	0.149	2.450	0.014	0.073	0.657

Table 8: The constant value, baseline and unobserved parameters' estimation of Gompertz-PHM

Parameters	Coef.	Std. Error	[95% Conf. Interval]	
Constant value	-2.844	0.728	-4.271	-1.416
Baseline Gompertz ancillary parameter (γ)	-0.001	0.000	-0.001	0.000

The results of the analysis under the assumption of Gompertz-PHM showed that Machine movement (z_m) and Excavator code (z_e) have a significant effect on the excavators' reliability. Having the regression coefficient for covariates, the reliability of the excavators will be equal to:

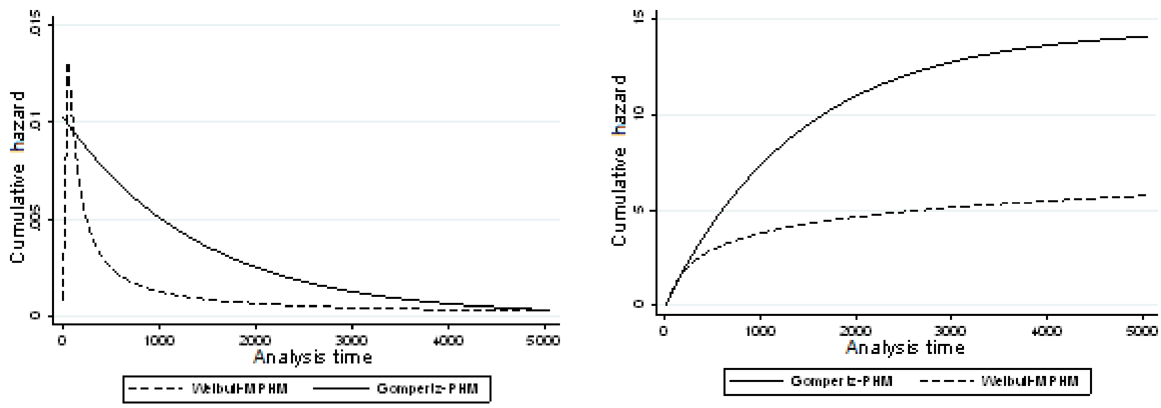
$$R(t;z) = \mathrm{e}^{\left(\left(\frac{e^{-0.001t}-1}{0.001}\right)\left(\exp\left(-2.884-1.606\mathrm{z_m}+0.365\mathrm{z_e}\right)\right)\right)} \tag{34}$$

As can be seen in Table 10, the regression coefficients of the observed covariates that have a significant effect on the hazard rate have different estimations in PHM than in MPHM.

Table 9: The difference between the effect of each covariate in PHM and MPHM

Covariates	Exp (coef.)		Differences	
	MPHM	PHM		
z_m	0.038	0.201	81%	Overestimate
z_e	2.527	1.440	-43%	Underestimate

Figure 5 compares the excavators' hazard and cumulative hazard rates in both models. As shown, after approximately 300 hours, different hazard rates are obtained, which hints that the unobserved covariates have a significant effect on the hazard rates of excavators; ignoring this factor may mislead a further decision on the operation and maintenance strategy.



(a) (b)

Figure 5: Comparison of the a) hazard function and b) cumulative hazard function under the Weibull-MPHM and Gompertz-PHM

5- Conclusion

The results of reliability analysis for heterogeneous data can differ substantially from those in a homogeneous case. In most cases, failing to account for heterogeneity would lead to significant differences in the estimation of the effects of covariates. Our recommendation is that all data sets should be checked for unobserved heterogeneity, using an appropriate statistical test. In the first step of the analysis of data sets with observed and unobserved heterogeneity, a time-dependency test of the observed covariates needs to be performed. Thereafter, the presence of unobserved covariates should be checked, using an appropriate statistical test. Finally, considering the type of repair strategy carried out on the item, the most appropriate model among the mixed proportional hazards model family should be selected. In our case study, the large variability in failure data and the differences in failure intensity of the excavators indicate heterogeneity among the collected data, which can be explained by observed and unobserved covariates. An analytical approach was used to check the trend and correlation of failure data. The result showed no trend and correlation among the data which could justify the *iid* assumption. Hence, the renewal process can represent the baseline hazard of excavators. The result of time-dependency and heterogeitysho tests (ratio test) indicated that all identified observed covariates are time-independent, and that there is an unobserved heterogeneity among the failure data. This means that some other factors, which were not included in this study, might have an effect on the excavators' reliability. Therefore, we need to further explore and model the effect of the unobserved factors, to enhance the accuracy of the estimation. Having these results and the developed framework (Figure 2), the mixed proportional hazards model (MPHM) was used to analyze the data. The result of analysis showed that two of the identified observed covariates have a significant effect on the hazard rate of the excavators. Ignoring the effect of unobserved covariates, and using PHM instead of MPHM, will underestimate the effect of Excavator type by 43 percent and overestimate the effect of Machine movement by 81 percent. Moreover, under the assumption of PHM, the baseline differs when MPHM is used to model the failure data. Thus, the failure rates of these models are completely different. Hence, for any decisions on the operation and maintenance strategy, the effect of unobserved covariates should be considered.

References:

1. Louit DM, Pascual R, Jardine AK. A practical procedure for the selection of time-to-failure models based on the assessment of trends in maintenance data. Reliability Engineering & System Safety. 2009;94(10):1618–28.
2. Liu X. Survival models on unobserved heterogeneity and their applications in analyzing large-scale survey data. Journal of biometrics & biostatistics. 2014;5.
3. Lawless JF. Statistical Models and Methods for Lifetime Data. John Wiley & Sons; 2011. 662 p.
4. Finkelstein M. Imperfect repair and lifesaving in heterogeneous populations. Reliability Engineering & System Safety. 2007 Dec 1;92(12):1671–6.
5. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Statistics in medicine. 1999;18(20):2693–708.
6. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in medicine. 2002;21(11):1539–58.
7. Garmabaki A, Ahmadi A, Block J, Pham H, Kumar U. A reliability decision framework for multiple repairable units. Reliability Engineering & System Safety. 2016;150:78–88.
8. Guida M, Giorgio M. Reliability analysis of accelerated life-test data from a repairable system. IEEE Transactions on Reliability. 1995;44(2):337–46.

9. Giorgio M, Guida M, Pulcini G. Repairable system analysis in presence of covariates and random effects. Reliability Engineering & System Safety. 2014 Nov 1;131:271–81.

10. Lawless JF. Regression methods for Poisson process data. Journal of the American Statistical Association. 1987;82(399):808–15.

11. Kumar D, Klefsjö B. Proportional hazards model: a review. Reliability Engineering & System Safety. 1994;44(2):177–88.

12. Gutierrez RG. Parametric frailty and shared frailty survival models. Stata Journal. 2002;2(1):22–44.

13. Hougaard P. Frailty models for survival data. Lifetime Data Anal. 1995 Sep 1;1(3):255–73.

14. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography. 1979;16(3):439–54.

15. Barabadi A, Barabady J, Markeset T. Maintainability analysis considering time-dependent and time-independent covariates. Reliability Engineering & System Safety. 2011;96(1):210–7.

16. Abbring JH, Van Den Berg GJ. The unobserved heterogeneity distribution in duration analysis. Biometrika. 2007;94(1):87–99.

17. Leszczyc PTP, Bass FM. Determining the effects of observed and unobserved heterogeneity on consumer brand choice. Applied stochastic models and data analysis. 1998;14(2):95–115.

18. Barabadi A, Gudmestad OT, Barabady J. RAMS data collection under Arctic conditions. Reliability Engineering & System Safety. 2015;135:92–9.

19. Gao X, Barabady J, Markeset T. An approach for prediction of petroleum production facility performance considering Arctic influence factors. Reliability Engineering & System Safety. 2010;95(8):837–46.

20. Garmabaki AHS, Ahmadi A, Mahmood YA, Barabadi A. Reliability Modelling of Multiple Repairable Units. Quality and Reliability Engineering International. 2016 Nov 1;32(7):2329–43.

21. Asfaw ZG, Lindqvist BH. Unobserved heterogeneity in the power law nonhomogeneous Poisson process. Reliability Engineering & System Safety. 2015;134:59–65.

22. Li M, Liu J. Bayesian hazard modeling based on lifetime data with latent heterogeneity. Reliability Engineering & System Safety. 2016 Jan 1;145:183–9.

23. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika. 1978;65(1):141–51.

24. Asadzadeh S, Aghaie A, Shahriari H, Niaki STA. The application of proportional hazards and frailty models to multistage processes surveillance. Int J Adv Manuf Technol. 2014 Sep 1;74(1):461–70.

25. Slimacek V, Lindqvist BH. Nonhomogeneous Poisson process with nonparametric frailty. Reliability Engineering & System Safety. 2016 May 1;149:14–23.

26. Asha G, Raja AV, Ravishanker N. Reliability modelling incorporating load share and frailty. Applied Stochastic Models in Business and Industry. 2018;34(2):206–23.

27. Xu M, Li X. Negative dependence in frailty models. Journal of Statistical Planning and Inference. 2008;138(5):1433–41.

28. Misra N, Gupta N, Gupta RD. Stochastic comparisons of multivariate frailty models. Journal of Statistical Planning and Inference. 2009 Jun 1;139(6):2084–90.

29. Slimacek V, Lindqvist BH. Reliability of wind turbines modeled by a Poisson process with covariates, unobserved heterogeneity and seasonality. Wind Energy. 2016;19(11):1991–2002.

30. Finkelstein M. Shocks in homogeneous and heterogeneous populations. Reliability Engineering & System Safety. 2007;92(5):569–74.

31. Barker K, Baroud H. Proportional hazards models of infrastructure system recovery. Reliability Engineering & System Safety. 2014 Apr 1;124:201–6.

32. Liu B, Liang Z, Parlikad AK, Xie M, Kuo W. Condition-based maintenance for systems with aging and cumulative damage based on proportional hazards model. Reliability Engineering & System Safety. 2017 Dec 1;168:200–9.

33. Slimacek V, Lindqvist BH. Nonhomogeneous Poisson process with nonparametric frailty and covariates. Reliability Engineering & System Safety. 2017 Nov 1;167:75–83.

34. Cha JH, Finkelstein M. Some notes on unobserved parameters (frailties) in reliability modeling. Reliability Engineering & System Safety. 2014 Mar 1;123:99–103.

35. Barabadi A, Garmabaki A, Yuan F, Lu J. Maintainability analysis of equipment using point process models. In IEEE; 2015. p. 797–801.

36. Ma Z, Krings AW. Multivariate survival analysis (I): shared frailty approaches to reliability and dependence modeling. In IEEE; 2008. p. 1–21.

37. Barabadi A, Barabady J, Markeset T. Application of reliability models with covariates in spare part prediction and optimization–a case study. Reliability Engineering & System Safety. 2014;123:1–7.

38. Lagakos SW. The graphical evaluation of explanatory variables in proportional hazard regression models. Biometrika. 1981 Apr 1;68(1):93–8.

39. Arjas E. A graphical method for assessing goodness of fit in Cox's proportional hazards model. Journal of the American Statistical Association. 1988;83(401):204–12.

40. Gray RJ. Some Diagnostic Methods for Cox Regression Models Through Hazard Smoothing. Biometrics. 1990;46(1):93–102.

41. Cox DR. A note on the graphical analysis of survival data. Biometrika. 1979 Apr 1;66(1):188–90.

42. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika. 1982 Apr 1;69(1):239–41.

43. Pettitt AN, Daud IB. Investigating Time Dependence in Cox's Proportional Hazards Model. Journal of the Royal Statistical Society Series C (Applied Statistics). 1990;39(3):313–29.

44. Park S, Hendry DJ. Reassessing Schoenfeld Residual Tests of Proportional Hazards in Political Science Event History Analyses. American Journal of Political Science. 2015;59(4):1072–87.

45. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. Wiley-Interscience; 2011.

46. Schoenfeld D. Chi-squared goodness-of-fit tests for the proportional hazards regression model. Biometrika. 1980 Jan 1;67(1):145–53.

47. O'Quigley J, Pessione F. Score Tests for Homogeneity of Regression Effect in the Proportional Hazards Model. Biometrics. 1989;45(1):135–44.

48. McKeague IW, Utikal KJ. Goodness-of-Fit Tests for Additive Hazards and Proportional Hazards Models. Scandinavian Journal of Statistics. 1991;18(3):177–95.

49. Gill R, Schumacher M. A simple test of the proportional hazards assumption. Biometrika. 1987 Jun 1;74(2):289–300.

50. Neumann GR. A Generalized Moments Specification Test of the Proportional Hazards Model AU - Horowitz, Joel L. Journal of the American Statistical Association. 1992 Mar 1;87(417):234–40.

51. Kimber AC. A Weibull-based score test for heterogeneity. Lifetime Data Anal. 1996 Mar 1;2(1):63–71.

52. Gray RJ. Tests for Variation over Groups in Survival Data. Journal of the American Statistical Association [Internet]. 2012 Feb 27 [cited 2019 Jan 30]; Available from: https://www.tandfonline.com/doi/abs/10.1080/01621459.1995.10476502

53. Commenges D, Andersen PK. Score test of homogeneity for survival data. Lifetime Data Anal. 1995 Jun 1;1(2):145–56.

54. Mittlböck M, Heinzl H. A simulation study comparing properties of heterogeneity measures in meta-analyses. Statistics in Medicine. 2006;25(24):4321–33.

55. Ioannidis JPA. Interpretation of tests of heterogeneity and bias in meta-analysis. Journal of Evaluation in Clinical Practice. 2008;14(5):951–7.

56. Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? Psychological methods. 2006;11(2):193.

57. Brewer MJ, Butler A, Cooksley SL. The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. Methods in Ecology and Evolution. 2016;7(6):679–92.

58. Yurekli K, Kurunc A. Performances of stochastic approaches in generating low streamflow data for drought analysis. Journal of Spatial Hydrology. 2005;5(1).

59. Kvaløy JT, Lindqvist BH. TTT-based tests for trend in repairable systems data. Reliability Engineering & System Safety. 1998;60(1):13–28.

60. Lee ET, Wang J. Statistical Methods for Survival Data Analysis. John Wiley & Sons; 2003. 528 p.

61. Coppini D, Bowtell P, Weng C, Young P, Sönksen P. Showing neuropathy is related to increased mortality in diabetic patients—a survival analysis using an accelerated failure time model. Journal of clinical epidemiology. 2000;53(5):519–23.

Paper C

Zaki, Rezgar; Barabadi, Abbas; Garmabaki, Amir Hossein Soleiman; Nouri, Ali. A mixture frailty model for maintainability analysis of mechanical components: A case study. International Journal of Systems Assurance Engineering and Management 2019; Volum 10 (6)

.

Int J Syst Assur Eng Manag
https://doi.org/10.1007/s13198-019-00917-3

ORIGINAL ARTICLE

A mixture frailty model for maintainability analysis of mechanical components: a case study

Rezgar Zaki[1] · Abbas Barabadi[1] · Ali Nouri Qarahasanlou[2] · A. H. S. Garmabaki[3]

Received: 29 April 2019 / Revised: 2 October 2019

Abstract Knowing the maintainability of a component or a system means that repair resource allocations, such as spare part procurement and maintenance training, can be planned and optimized more effectively. Repair data are often collected from multiple and distributed units in different operational conditions, which can introduce heterogeneity into the data. Part of such heterogeneity can be explained and isolated by the observable covariates, whose values and the way that they can affect the item's maintainability are known. However, some factors which may affect maintainability are typically unknown (unobserved covariates), leading to unobserved heterogeneity. Nevertheless, many researchers have ignored the effect of observed and un-observed covariates, and this may lead to erroneous model selection, as well as wrong conclusions and decisions. Moreover, many authors have simplified their analysis by considering a complex system as a single item. In these studies, the assumption is that all repair data represent an identical repair process for the item. In practice, mechanical systems are composed of multiple parts, with various failure mechanisms, which need different repair processes (repair modes) to return to the operational phase; classical distribution, such as lognormal, which is only a function of time, may not be able to model such complexity. The paper utilizes the mixture frailty model (MFM) in the presence of some specific observed or unobserved covariates to predict maintainability more precisely. MFMs can model the effect of observed and unobserved covariates, as well as identifying different repair processes in the repair dataset. The application of the proposed model is demonstrated by a case study.

Keywords Mixture Weibull · Failure model · Repair process · Covariates · Repair time · Maintainability · Frailty model

1 Introduction

In today's society, we are strongly dependent on the optimal functioning of complex technical systems, such as communication networks, railways, power plant control systems, aircraft, mining, the oil and gas industry, etc. As these systems fail, they should be repaired as soon as possible in a safe manner, to reduce the consequences of the failure, including production loss, safety and health effects. To achieve such a goal, the system should be designed for maintainability.

Maintainability is a design factor that decides the degree to which a product allows safe, quick and easy replacement of its component parts. Design for maintainability refers to designing the system to find the optimum balance between capital cost and ongoing maintenance cost (Barabadi et al. 2010; Tortorella 2015; Tsarouhas 2015; Garmabaki et al. 2016a; Gharahasanlou et al. 2017; Kumar et al. 2017; Aggarwal et al. 2017). Design for maintainability needs to consider human ergonomics, logistics management, design layout, the level of experience and training of maintenance

✉ Abbas Barabadi
abbas.b.abadi@uit.no

[1] Department of Technology and Safety, UiT The Arctic University of Norway, Tromsø, Norway

[2] Faculty of Mining Engineering, Petroleum and Geophysics, Shahrood University of Technology, Shahrood, Iran

[3] Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden

Published online: 07 November 2019

personnel and so on (Knezevic 1993; Naseri and Barabady 2016). Ease of access, standardization of equipment—both internally and between companies—skill levels to maintain equipment, organization culture, service delivery infrastructure, etc. are some of the main elements of maintainability. Maintainability performance is defined as "the ability of an item under given conditions of use, to be retained in, or restored to, a state in which it can perform a required function, when maintenance is performed under given conditions and using stated procedures and resources" (Rausand and Høyland 2004). 'Item' here refers to a system, component, or subsystem (Barabadi and Markeset 2011; Furuly et al. 2013). Existing studies regarding the maintainability analysis of historical data have ignored the effect of observed and unobserved covariates (risk factors). Moreover, many authors have simplified their analysis by considering a complex system as a single item, then modeling the repair data using a classical distribution, mostly lognormal (Tsarouhas et al. 2009; Hoseinie et al. 2011; Tsarouhas and Arvanitoyannis 2012; Wang 2014; Sellitto and Brusius Jr. 2017; Tsarouhas 2018). However, the relationship between the elements of maintainability is complex and mediated by many influence factors such as ambient temperature, human factors, and dissimilarity in personality or skill level between maintenance crews, etc. As a main part of maintainability, logistics and spare parts constitute a complex activity that is time- and location-dependent (Wijaya and Lundberg 2012; Ahmadzadeh and Lundberg 2014; Barabadi et al. 2015, 2016; Gudmestad and Markeset 2015; Ayele et al. 2016). Hence, a single distribution, such as lognormal, which is only a function of time, is not able to capture such complexities.

Recently, some attempts have been made to relate maintainability to both historical repair data and operational conditions as observed covariates. Gao et al. (2010) developed the proportional repair model (PRM), based on the proportional hazard model (PHM), which is the most widely used in reliability analysis, when considering the effect of operational conditions (Kumar and Klefsjö 1994; Rosen and Tanner 1999; Gao et al. 2010; Van Horenbeek et al. 2010; Barabadi et al. 2011). An important alternative to the PHM is the accelerated failure time model (AFT). The AFT model accounts for the effects of the covariates directly on survival times, instead of the hazard rate as in the PHM (Patel et al. 2006; Barabadi et al. 2011; Garmabaki et al. 2016c; Ayele et al. 2018).

The PRM is a product of the baseline repair rate and a functional term incorporating the effects of time-independent observed covariates. PRMs are only able to model the effect of time-independent observed covariates. In the case of time-dependent covariates, the assumption of proportionality is violated, and the PRM cannot be built. To deal with non-proportionality, Barabadi and Markeset (2011) used a stratification approach to model the effect of time-independent covariates. In the stratification approach, the data are categorized based on different levels of time-dependent observed covariates. However, their studies did not consider the effect of unobserved covariates. Unobserved covariates are covariates whose effects on the repair process are typically unknown or whose associated levels during repair time are not available in the repair database (Gimenez et al. 2018). Ignoring the effect of observed and unobserved covariates would lead to significant differences in the estimation of the effects of covariates (Vaupel et al. 1979; Kumar and Klefsjö 1994; Hougaard 1995; Ayele et al. 2016). Observed and unobserved covariates result in observed or unobserved heterogeneity among repair data (Asfaw and Lindqvist 2015). A systematic literature review revealed no articles dealing with the modeling of unobserved covariate effect on the maintainability of items.

Moreover, in these studies, the assumption is that all repair data represent an identical repair process for the item. In reality, mechanical systems are composed of multiple parts, with various failure mechanisms, which need different repair processes (repair modes) to return them to the operational phase. For instance, a gearbox failure may result from individual failures in the gears, bearings, or shafts and include fatigue cracks, teeth breakage, wear, etc. These failure modes may have completely different repair processes and resources. In most of the available databases, the repair data are mixed together under the title, "Repair data for gearbox" (Barabadi et al. 2015; Gharahasanlou et al. 2017). In dealing with such datasets, as mentioned, analysts simplify their analysis by considering such complex datasets as homogeneous, with repair data being represented by an identical repair process. These studies have viewed the historical data as a black box, with no information regarding the repair process and its operational conditions.

According to the discussion, two issues should be considered when modeling the maintainability of an item:

(1) The selected model should be able to capture the effect of observed and unobserved covariates on the time of the maintenance, and
(2) As in many cases the repair data are a mix of different repair processes (repair modes), the applied model should be able to isolate different repair modes.

When each repair process is regarded as an independent repair mode with a repair distribution in the presence of some specific observed or un-observed risk factors, then a mixture frailty model (MFM) can be separately constructed to effectively predict maintainability.

An MFM is an extension of the PRM, where unobserved and observed covariates have a multiplicative effect on the repair rate. In MFM, the repair rate of an item is the product of a baseline hazard rate multiplied by two positive functions: an observed and an unobserved covariate function (frailty function). In addition, the MFM has the ability to model different repair modes. The frailty term was introduced by Clayton et al. (1978) and Vaupel et al. (1979) in survival analysis in medical science, based on the PHM. Several researchers later used the frailty concept to model the effect of unobserved covariates on the reliability of an item, to describe the influence of unobserved covariates. For example, Slimacek and Lindqvist (2016) used frailty to model the effect of unobservable differences between turbines, as unobserved covariates, on the reliability of wind turbines, using a Poisson process. Giorgio et al. (2014) modeled the failure pattern of a powertrain system in the presence of observed and unobserved heterogeneity, via a joint probability distribution on power law process parameters. Finkelstein (2007) used the frailty model to study the reliability of a system subject to shocks, which occur in accordance with a non-homogeneous Poisson process. He showed that that reliability analysis for a heterogeneous case could differ dramatically from that for a homogeneous setting.

The rest of the paper is organized as follows. In Sect. 2, the MFM is explained; thereafter, in Sect. 3, the application of the proposed model is illustrated by a case study. Finally, Sect. 4 provides the conclusions.

2 Mixture frailty model (MFM) for maintainability analysis

The repair rate of an item is the rate at which a repair action is performed. It is expressed in terms of the number of repair actions performed and successfully completed per unit of time, by considering time-dependent and time-independent observed and unobserved covariates. It can be expressed as follows:

$$\mu\left(t, z_i, z_j(t)|A\right) = A\mu_0(t)\psi(z, z(t); P; \delta) \tag{1}$$

where z_i and $z_j(t)$ are time-independent and time-dependent observed covariates $P; \delta$ are column vectors, consisting of the regression parameters for identified time-independent and time-dependent observed covariates; and A is a random positive quantity, representing the cumulative effect of one or more unobserved covariates.

Here, the repair rate, $\mu\left(t, z_i, z_j(t); A\right)$, consists of three multiplicative factors: (1) the baseline repair rate $\mu_0(t)$, dependent on time alone, which is modeled using appropriate distributions; (2) a positive multiplicative factor, $\psi(z, z(t); P; \delta)$, to describe the function of time-independent and time-dependent observed covariates; and (3) a positive multiplicative factor, A, representing the effect of unobserved covariates. The observed and unobserved covariates can affect the repair rate, so that the actual repair rate, $\mu\left(t, z_i, z_j(t); A\right)$, is either greater (e.g. in the case of poor maintenance) or less (e.g. with better training for operators and maintenance crew) than the baseline repair rate. Here, those items with $A>1$ are said to be less frail, for reasons left unexplained by the observed covariates, and will have an increased repair rate. Those items for which $A < 1$ are frailer; hence, given a certain observed covariate pattern, they tend to reduce the repair time. In general, the exponential functional form and gamma distribution (with the mean equal to one and variance of θ) are the most commonly used functions for modeling observed and unobserved covariates, respectively (Cha and Finkelstein 2014; Asfaw and Lindqvist 2015; Garmabaki et al. 2016b; Slimacek and Lindqvist 2017). Under these assumptions, the maintainability function can be written as:

$$\begin{aligned}
M\left(t, z_i, z_j(t)|A\right) &= 1 - \exp\left[-\int_0^t \mu(u|A)du\right] \\
&= 1 - \exp\left[-A\int_0^t \frac{m(u)}{1-M(u)}du\right] \\
&= 1 - \left\{1 - M\left(t, z_i, z_j(t)\right)\right\}^A
\end{aligned} \tag{2}$$

Because A is unobservable, it must be integrated out of $M\left(t, z_i, z_j(t)|A\right)$ to obtain the unconditional maintainability function. When A is distributed as gamma with mean one and variance θ:

$$g(A) = \frac{A^{\frac{1}{\theta}-1}e^{-\frac{A}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)\theta^{\frac{1}{\theta}}} \tag{3}$$

Then maintainability becomes:

$$\begin{aligned}
M_\theta(t) &= 1 - \int_0^\infty \left\{1 - M\left(t, z_i, z_j(t)\right)\right\}^A \cdot \frac{A^{\frac{1}{\theta}-1}e^{-\frac{A}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)\theta^{\frac{1}{\theta}}}dA \\
&= 1 - \left[1 - \theta\ln\left\{1 - M\left(t, z_i, z_j(t)\right)\right\}\right]^{-\frac{1}{\theta}}
\end{aligned} \tag{4}$$

If the observed covariate follows the exponential function in the presence of W time-independent observed covariates and M time-dependent observed covariates, $M\left(t, z_i, z_j(t)\right)$ can be written as:

$$M\left(t, z_i, z_j(t)\right) = 1 - [1 - M_0(t)]^{\exp\left[\sum_{i=1}^{W} p_i z_i + \sum_{j=1}^{M} \delta_j z_j(t)\right]} \tag{5}$$

where $M_0(t)$ is the baseline maintainability function dependent only on the time, as follows:

$$M_0(t) = 1 - \exp\left[-\int_0^t \mu_0(t')dt'\right] \tag{6}$$

As mentioned, in reality, the historical repair data are a mix of different repair processes (repair modes); hence, the applied model should be able to isolate different repair modes. If each repair process is regarded as an independent process with an individual, repair distribution that the presence of some specific covariates, then mixture distribution can be used to model the maintainability baseline. Suppose a repair dataset of specific items consists of N repair processes, which require different maintenance tasks and repair actions comprised of several subsidiary tasks of unequal frequency and time duration. Under these conditions, the mixture baseline maintainability function, ($M_{0m}(\mathrm{t})$), can be defined by mixing the $M_0(t)$ of the several repair processes as:

$$\begin{aligned} M_{0m}(\mathrm{t}) &= \sum_{\mathrm{k}=1}^{\mathrm{N}} \gamma_{\mathrm{k}} \cdot M_{0k} \\ &= \sum_{\mathrm{k}=1}^{\mathrm{N}} \gamma_{\mathrm{k}} \cdot \left(1 - \exp\left[-\int_0^t \mu_{k0}(t')dt'\right]\right) \end{aligned} \tag{7}$$

where M_{0k} is the baseline maintainability of the kth repair process and Υ_{k} is the proportion of the repair tasks belonging to the kth repair process. If the basic principle of the probability dominance, which states that the summation of all of the proportion of Υ_{k} has to be one therefore $\sum_{k=1}^{N} \gamma_k = 1$ should hold.

The baseline maintainability function, if the repair rate for all repair processes follows 2-parameter Weibull distribution, is given by:

$$M_{0m}(\mathrm{t}) = \sum_{\mathrm{k}=1}^{\mathrm{N}} \gamma_{\mathrm{k}} \cdot \left(1 - e^{-\left(\frac{t}{\eta_{\mathrm{k}}}\right)^{\beta_{\mathrm{k}}}}\right) \tag{8}$$

where β_{k} and η_{k} are the shape parameter and scale parameter of Weibull distribution for the kth repair process. Likelihood function can be used to estimate the parameters in Eq. (8). The estimation of the maximum likelihood for the given log-likelihood function is demanding. Therefore, some type of iterative algorithms can be employed to approximately estimate the parameters for the mixed distribution.

3 Case study

Figure 1 shows a black diagram for a production line in the Sungun Copper Mine in Iran. In this production line, seven Komatsu HD 325-6 dump trucks work, defined as DT.1–DT.7. Here, the repair data for all dump trucks are collected through daily repair and operation reports. As these reports are not designed for maintainability analysis, the repair processes are mostly neither well recorded nor detailed. However, based on a discussion with experts, the collected repair dataset is a mixture of different repair processes. In addition, for each repair time, associated

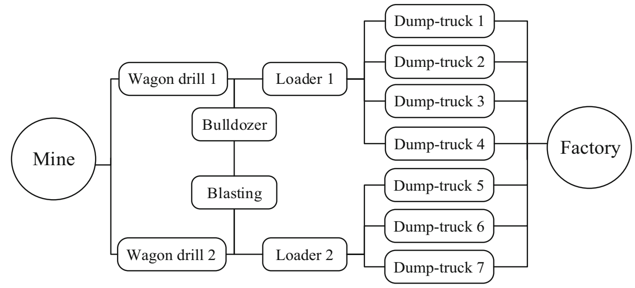


Fig. 1 Block diagram for a production line at the Sungun Copper Mine

observed covariates have been collected through discussion with experts at the mine. Table 1 shows the identified observed covariates and their associated levels. The levels for each covariate are identified, based on the different operational conditions that trucks will experience during their mission time. The maintainability covariates include working shift, weather condition, precipitation, temperature, and number of involved maintenance crews. The shift generally represents a diverse maintenance crew, whose different skills and expertise may affect the maintainability performance of the trucks. In addition, some maintenance tasks can take a long time to complete. In such scenarios, several maintenance crews will work to repair the trucks over a number of shifts. Under these conditions, repair crews need effective communication. Ineffective communication will significantly reduce maintainability, as some jobs need to be repeated or must be double-checked. Hence, the number of maintenance crews working on a truck is considered a maintainability covariate. Moreover, as most of the maintenance is performed outdoors, precipitation is considered a covariate.

Using the MFM and this assumption that all covariates are time-independent, the maintainability for each

Table 1 Identified maintainability covariates and their associated levels

Maintainability covariates	Covariate level	Assigned code
Working shift (z_1)	Morning	1
	Afternoon	2
	Night	3
Weather condition (z_2)	Sunny & clear	1
	Semi cloudy	2
	Overcast	3
	Dense fog	4
Precipitation (z_3)	Continuous covariate	
Temperature (z_4)	Continuous covariate	
Involved maintenance crew (z_5)	One	1
	More than one	2

component can be estimated. The results of the analysis are shown in Table 2. Here, the p value of 5% is considered the upper limit to check the significance of observed and unobserved covariates. This table shows the likelihood ratio (LR), which is used to check whether the unobserved covariate(s) has (have) a significant effect on maintainability. In this case, LR can be written as follows (Garmabaki et al. 2016b):

$$LR = 2(\ln L(\beta_k,\ \eta_k,\ p_k, \gamma_k, \theta) - \ln(\beta_{0k},\ \eta_{0k},\ p_{0k}, \gamma_{0k}, 0)) \tag{9}$$

Here $\beta_{0k}, \eta_{0k}, p_{0k}$ and γ_{0k} are estimated parameters under the null hypothesis, where $\theta = 0$, which means the unobserved covariate(s) has (have) no significant effect on the maintainability of the trucks. As Table 2 shows, for example for truck DT.7, the baseline maintainability is Mixture 2 Weibull distribution, the first population is 43%, with $\beta_1 = 4.189$ and $\eta_1 = 6.249$, while the second population is around 57%, with $\beta_2 = 0.982$ and $\eta_2 = 21.166$. Using these parameters, the baseline mean time to repairs (MTTR) will be equal to 5.7 and 21.3 for Population No. 1 and No. 2, respectively. Moreover, involved maintenance crew (z_5) and precipitation (z_3) have a significant effect on the maintainability of DT.7, with their regression coefficients being equal to − 1.872 and − 1.187, respectively. Moreover, the p value associated with the LR of DT.7 is equal to 0.000, which hints that unobserved covariates have a significant effect on the maintainability of DT.7, with $\theta = 9.22$.

In the next step, to compare the results of analysis with the traditional model, the data were analyzed by classical distributions. Here, we nominated five distributions, including 3P-Weibull, 2P-Exponential, 1P-Exponential, 2P-Weibull, and Normal distributions. Thereafter, using goodness of fit test, the best fit distribution for each truck was identified.

Table 2 Estimated parameters of the maintainability performance of selected items

Truck	Observed covariates			Unobserved covariates			Baseline model		Baseline MTTR	γ (%)
	Covariates	p_i	p value	LR	p value	θ				
DT.1	Working shift (z_1)	− 1.370	0.001	16.3	0	3.24	Mixture 3 Weibull	$\beta_1 = 2.234$; $\eta_1 = 2.289$	2.03	19
	Involved maintenance crew (z_5)	− 1.179	0.001					$\beta_2 = 2.845$; $\eta_2 = 6.7054$	5.97	47
								$\beta_3 = 0.718$; $\eta_3 = 26.7674$	33.1	35
DT.2	Working shift (z_1)	− 0.751	0.005	39.7	0	3.82	Weibull-3P	$\beta_1 = 1.054$; $\eta_1 = 6.810$; $\gamma_1 = 0.815$	6.95	100
	Involved maintenance crew (z_5)	− 1.351	0.000							
	Precipitation (z_3)	− 0.090	0.035							
DT.3	Working shift (z_1)	− 3.267	0.000	57.1	0	7.59	Weibull-3P	$\beta_1 = 0.946$; $\eta_1 = 7.886$; $\gamma_1 = 0.925$	9.01	100
	Involved maintenance crew (z_5)	0.356	0.033							
	Precipitation (z_3)	− 0.237	0.001							
DT.4	Working shift (z_1)	− 1.214	0.000	64.2		3.08	Mixture 3 Weibull	$\beta_1 = 3.750$; $\eta_1 = 3.980$	3.6	32
	Involved maintenance crew (z_5)	− 1.059	0.001					$\beta_2 = 18.160$; $\eta_2 = 6.624$	6.43	21
								$\beta_2 = 1.122$; $\eta_2 = 10.435$	10.0	47
DT.5	Involved maintenance crew (z_5)	− 5.077	0.000	64.0	0	11.94	Mixture 4 Weibull	$\beta_1 = 6.399$; $\eta_1 = 1.752$	1.63	11
								$\beta_2 = 3.023$; $\eta_2 = 3.402$	3.03	38
	Weather condition (z_2)	0.408	0.004					$\beta_3 = 20.828$; $\eta_3 = 6.548$	6.38	33
								$\beta_4 = 0.683$; $\eta_4 = 72.208$	93.6	18
DT.6	Working shift (z_1)	− 0.485	0.094	26.5	0	2.33	Mixture 4 Weibull	$\beta_1 = 3.465$;$\eta_1 = 1.793$	1.65	21
	Involved maintenance crew (z_5)	− 2.505	0.000							
	Weather condition (z_2)	0.387	0.022					$\beta_2 = 4.965$; $\eta_2 = 3.809$	3.5	40
	Precipitation (z_3)	− 0.248	0.028					$\beta_3 = 22.331$; $\eta_3 = 6.622$	6.5	23
								$\beta_4 = 0.901$; $\eta_4 = 28.634$	30.1	16
DT.7	Precipitation (z_3)	− 1.827	0.000	41.2	0	9.22	Mixture 2 Weibull	$\beta_1 = 4.189$; $\eta_1 = 6.249$	5.7	43
	Involved maintenance crew (z_5)	− 1.187	0.000					$\beta_2 = 0.982$; $\eta_2 = 21.166$	21.3	57

For example, Table 3 shows the results of analysis for DT.7 and the goodness of fit test, using three different models. The *AvGOF* column contains the average values from the Kolmogorov–Smirnov (*GOF*) test, the *AvPLOT* column contains the average values from the correlation coefficient (*PLOT*) test, and the *LKV* column contains the average values from the Likelihood Value (*LKV*) test. As the goodness of fit tests in Table 3 (*AvGOF*, *AvPLOT* and *LKV*) show the 3P-Weibull with $\beta = 1.046$; $\eta = 12.836$ and $\gamma = 0.877$ is the best fit distribution for DT.7. Hence, the maintainability of truck DT.7 can be written as:

$$M(t) = 1 - \exp\left(\frac{t - 0.877}{12.836}\right)^{1.046} \tag{10}$$

Using Eq. (10), the MTTR of DT.7 will be equal to 13.5 h.

The repair rates for DT.7 in both models (MFM and classical distribution) are compared in Fig. 2. There is a significant difference between the repair rates of DT.7 using these two models. In other words, observed and unobserved covariates have a significant effect on the repair rates of trucks; ignoring this factor may mislead a further decision on the operation and maintenance strategy. For example, according to the MFM, the repair rate of DT.7 after 10 h will be equal to 0.12, while using the classical approaches, and equal to 0.2 using MFM. Moreover, using MFM for maintainability analysis, we will obtain more information regarding the influencing factors. This information will help managers and decision-makers establish a more effective maintenance plan. For example, we know that the involved maintenance crew will decrease DT.7's repair rate by 90%. Hence, in critical situations, increasing the number of involved maintenance crews can significantly increase the availability of DT.7.

4 Conclusion

The existing studies regarding the analysis of historical repair data have mostly ignored the effect of observed and unobserved covariates. As observed and unobserved covariates result in heterogeneity in repair data, the selected model should be able to capture the effects of both types of covariates.

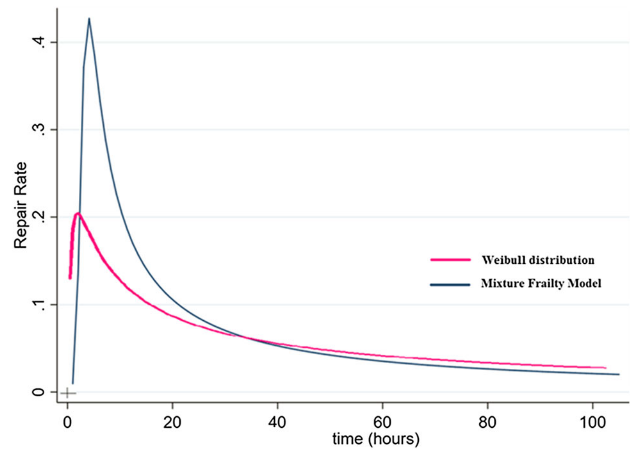


Fig. 2 Repair rate of DT.7, using MFM approach and 3P Weibull distribution

In this paper, the application of MFM for maintainability analysis has been discussed. MFM has the ability to model the effect of observed and unobserved covariates on maintainability. Moreover, it can capture different repair processes in a single database, by the use of a convex combination of their associated distributions.

In the second part of the paper, the application of the developed model is illustrated by investigating the effect of observed and unobserved covariates on the maintainability of trucks at a copper mine. The results of analysis show that most identified observed covariates and unobserved covariate(s) have a significant effect on the maintainability of trucks. The results suggest that, in most cases, the baseline maintainability of trucks contains a mixture of different distributions. Finally, comparing the results of analysis using the MFM approach and classical distribution shows that ignoring the effects of observed and unobserved covariates can lead to significant deviation in the maintainability estimation. Such deviations may significantly affect any future operation and maintenance planning of the production process.

Table 3 The result of GOF analysis for DT.7

Distribution	Goodness of fit test			Parameters
	AvGOF	AvPLOT	LKV	
3P-Weibull	99.9827	6.1446	− 318.8260	Beta = 1.046; Eta = 12.836; Gamma = 0.877
2P-Exponential	99.9974	10.1770	− 343.8620	Lambda = 5.7E − 02; Gamma = -4.097
1P-Exponential	99.9986	9.1357	− 323.1657	Lambda = 0.058
2P-Weibull	99.9994	79.7068	− 349.1245	Beta = 1.418; Eta = 13.232
Normal	99.9999	12.9616	− 384.6652	Mean = 14.898; Std = 15.220

References

Aggarwal A, Kumar S, Singh V (2017) Performance modeling of the serial processes in refining system of a sugar plant using RAMD analysis. Int J Syst Assur Eng Manag 8:1910–1922. https://doi.org/10.1007/s13198-016-0496-1

Ahmadzadeh F, Lundberg J (2014) Remaining useful life estimation: review. Int J Syst Assur Eng Manag 5:461–474. https://doi.org/10.1007/s13198-013-0195-0

Asfaw ZG, Lindqvist BH (2015) Unobserved heterogeneity in the power law nonhomogeneous Poisson process. Reliab Eng Syst Saf 134:59–65

Ayele YZ, Barabadi A, Barabady J (2016) Dynamic spare parts transportation model for Arctic production facility. Int J Syst Assur Eng Manag 7:84–98. https://doi.org/10.1007/s13198-015-0379-x

Ayele Y, Barabadi A, Fuqing Y (2018) Mixture lognormal cox regression repair model for prediction of the repair time. In: IEEE, pp 1155–1159

Barabadi A, Markeset T (2011) Reliability and maintainability performance under Arctic conditions. Int J Syst Assur Eng Manag 2:205–217

Barabadi A, Barabady J, Markeset T (2010) Application of accelerated failure model for the oil and gas industry in Arctic region. In: IEEE, pp 2244–2248

Barabadi A, Barabady J, Markeset T (2011) Maintainability analysis considering time-dependent and time-independent covariates. Reliab Eng Syst Saf 96:210–217

Barabadi A, Gudmestad OT, Barabady J (2015) RAMS data collection under Arctic conditions. Reliab Eng Syst Saf 135:92–99. https://doi.org/10.1016/j.ress.2014.11.008

Barabadi A, Garmabaki A, Zaki R (2016) Designing for performability: an icing risk index for Arctic offshore. Cold Reg Sci Technol 124:77–86

Cha JH, Finkelstein M (2014) Some notes on unobserved parameters (frailties) in reliability modeling. Reliab Eng Syst Saf 123:99–103. https://doi.org/10.1016/j.ress.2013.10.008

Clayton DG (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65:141–151

Finkelstein M (2007) Shocks in homogeneous and heterogeneous populations. Reliab Eng Syst Saf 92:569–574

Furuly S, Barabadi A, Barabady J (2013) Reliability analysis of mining equipment considering operational environments: a case study. Int J Perform Eng 9:287–294

Gao X, Barabady J, Markeset T (2010) An approach for prediction of petroleum production facility performance considering Arctic influence factors. Reliab Eng Syst Saf 95:837–846

Garmabaki A, Ahmadi A, Ahmadi M (2016a) Maintenance optimization using multi-attribute utility theory. In: Kumar U, Ahmadi A, Verma AK, Varde P (eds) Current trends in reliability, availability, maintainability and safety. Springer, Berlin, pp 13–25

Garmabaki A, Ahmadi A, Block J et al (2016b) A reliability decision framework for multiple repairable units. Reliab Eng Syst Saf 150:78–88. https://doi.org/10.1016/j.ress.2016.01.020

Garmabaki AHS, Ahmadi A, Mahmood YA, Barabadi A (2016c) Reliability modelling of multiple repairable units. Qual Reliab Eng Int 32:2329–2343. https://doi.org/10.1002/qre.1938

Gharahasanlou AN, Ataei M, Khalokakaie R et al (2017) Risk based maintenance strategy: a quantitative approach based on time-to-failure model. Int J Syst Assur Eng Manag. https://doi.org/10.1007/s13198-017-0607-7

Gimenez O, Cam E, Gaillard J-M (2018) Individual heterogeneity and capture–recapture models: what, why and how? Oikos 127:664–686. https://doi.org/10.1111/oik.04532

Giorgio M, Guida M, Pulcini G (2014) Repairable system analysis in presence of covariates and random effects. Reliab Eng Syst Saf 131:271–281. https://doi.org/10.1016/j.ress.2014.04.009

Gudmestad OT, Markeset T (2015) Oil and gas operations under extreme conditions in the cold north. Int J Comput Methods Exp Meas 3:7–12

Hoseinie SH, Ataei M, Khalokakaie R, Kumar U (2011) Reliability and maintainability analysis of electrical system of drum shearers. J Coal Sci Eng (China) 17:192–197. https://doi.org/10.1007/s12404-011-0216-z

Hougaard P (1995) Frailty models for survival data. Lifetime Data Anal 1:255–273. https://doi.org/10.1007/BF00985760

Knezevic J (1993) Reliability, maintainability, and supportability: a probabilistic approach. McGraw-Hill, New York City

Kumar D, Klefsjö B (1994) Proportional hazards model: a review. Reliab Eng Syst Saf 44:177–188

Kumar L, Kumar M, Rath S (2017) Maintainability prediction of web service using support vector machine with various kernel methods. Int J Syst Assur Eng Manag 8:205–222. https://doi.org/10.1007/s13198-016-0415-5

Naseri M, Barabady J (2016) On RAM performance of production facilities operating under the Barents Sea harsh environmental conditions. Int J Syst Assur Eng Manag 7:273–298. https://doi.org/10.1007/s13198-016-0463-x

Patel K, Kay R, Rowell L (2006) Comparing proportional hazards and accelerated failure time models: an application in influenza. Pharm Stat 5:213–224. https://doi.org/10.1002/pst.213

Rausand M, Høyland A (2004) System reliability theory: models, statistical methods, and applications. Wiley, New York

Rosen O, Tanner M (1999) Mixtures of proportional hazards regression models. Stat Med 18:1119–1131. https://doi.org/10.1002/(SICI)1097-0258(19990515)18:9%3c1119:AID-SIM116%3e3.0.CO;2-V

Sellitto MA, Brusius W Jr (2017) Maintenance strategy based on reliability analytical models for three parallel mechanical transformation machines. IEEE Latin Am Trans 15:988–993

Slimacek V, Lindqvist BH (2016) Reliability of wind turbines modeled by a Poisson process with covariates, unobserved heterogeneity and seasonality. Wind Energy 19:1991–2002. https://doi.org/10.1002/we.1964

Slimacek V, Lindqvist BH (2017) Nonhomogeneous Poisson process with nonparametric frailty and covariates. Reliab Eng Syst Saf 167:75–83. https://doi.org/10.1016/j.ress.2017.05.026

Tortorella M (2015) Reliability, maintainability, and supportability: best practices for systems engineers. Wiley, New York

Tsarouhas PH (2015) Maintainability analysis in the yogurt industry. Int J Syst Assur Eng Manag 6:119–128. https://doi.org/10.1007/s13198-014-0254-1

Tsarouhas P (2018) Reliability, availability and maintainability (RAM) analysis for wine packaging production line. Int J Qual Reliab Manag 35:821–842

Tsarouhas PH, Arvanitoyannis IS (2012) Reliability and maintainability analysis to improve the operation of the limoncello production line. Int J Food Sci Technol 47:1669–1675

Tsarouhas PH, Arvanitoyannis IS, Ampatzis ZD (2009) A case study of investigating reliability and maintainability in a Greek juice bottling medium size enterprise (MSE). J Food Eng 95:479–488

Van Horenbeek A, Pintelon L, Muchiri P (2010) Maintenance optimization models and criteria. Int J Syst Assur Eng Manag 1:189–200. https://doi.org/10.1007/s13198-011-0045-x

Vaupel JW, Manton KG, Stallard E (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16:439–454

Wang J (2014) Proceedings of the first symposium on aviation maintenance and management-volume II. Springer, Berlin

Wijaya AR, Lundberg J (2012) The effect of the operator, the mine room and their interaction on the measured vibration level of a scaling machine. Int J Syst Assur Eng Manag 3:145–152. https://doi.org/10.1007/s13198-012-0096-7