UNIVERSITÀ DEGLI STUDI FIRENZE

PhD in
Industrial Engineering

CYCLE XXXII

COORDINATOR Prof. Maurizio De Lucia

Dynamic Risk-based Asset Integrity Modelling of Engineering Processes

Academic Discipline (SSD) ING-IND/17

Doctoral Candidate
Dr. Ahmad BahooToroody

(signature)

Supervisor
Prof. Filippo De Carlo

(signature)

Coordinator
Prof. Maurizio De Lucia

(signature)

Years 2016/2020

[Page intentionally left blank]

[Page intentionally left blank]

Abstract

Poorly maintenance scheduling and the resulting downtime are costly. Deliver maximum performance while minimising costs and risks over the whole life of engineering systems required a developmental transition from traditional maintenance strategies to smart predictive maintenance. This PhD study aimed at maximising the value realised from complex engineering assets and systems. To this end, statistical modelling and machine learning were established to problems in intelligent maintenance operations, characterized by data-driven innovations.

The outcome of this study are far better avoidance of failures, significantly reducing whole-life costs, better running of assets, increasing the resilience of components while no inappropriate risk has to be taken.

This thesis contains six chapters. The first and final chapters are devoted to **Introduction** (including the general structure of the thesis) and **Conclusion**, while the remainder chapters are presented in the following order:

- **Chapter 2:** To quantify the uncertainty associated with failure modelling a new framework was developed. The stochastic trend existed in the inter-arrival time of failures has to be incorporated if making an informed decision over the availability and maintainability of system is of importance. Accordingly, and given that the time between failures are dependently distributed or independently, the Hierarchical Bayesian Model and Maximum Likelihood Estimation was applied to model the reliability of the process and highlight the magnitude of the deviation value in different failure modelling approaches. A case study of Natural Gas Regulating and Metering Stations in Italy has been considered to illustrate the application of proposed framework. The results highlight that relaxing the renewal process assumption and taking the time dependency of the observed data into account will result in more precise failure models.

- **Chapter 3:** Following to the research outcome of second chapter, a dynamic time dependent reliability assessment was proposed. This chapter aimed at proposing a probabilistic model to predict the complexity of the non-stationary behavior in monitoring data. To this end, Bayesian inference with hierarchical structure was employed given minimal repair assumption. The raw data collected from observations in condition

monitoring of gas pipelines consists of non-stationary trends, short term cycle and noise. As the noise has a complex time-dependent auto-correlation structure, Empirical Mode Decompositions (EMD) was adopted to extract the disturbing noises from the time series. The advantages of the methodology were demonstrated through a case study of a Natural Gas Regulating and Metering Station operating in Italy. Based on pressure data acquired from the case study, the model is able to predict overpressure thus directly avoiding unnecessary maintenance and safety consequences.

- **Chapter 4:** In this chapter an online reliability assessment was developed in order to signifying the impact of risk factors on safety indicators when consideration is given to uncertainty quantification. For this purpose, the generalized linear model (GLM) is applied to offer the explanatory model as a regression function for risk factors and safety indicators. Hierarchical Bayesian approach (HBA) is then inferred for the calculations of regression function including interpretation of the intercept and coefficient factors. With Markov Chain Monte Carlo simulation from likelihood function and prior distribution, the HBA is capable of capturing the aforementioned fluctuations and uncertainties in the process of obtaining the posterior values of the intercept and coefficient factors. To illustrate the capabilities of the developed framework, an autonomous operation of Natural Gas Regulating and Metering Station in Italy has been considered as case study. The resulted model provides designers, risk managers and operators a framework for risk mitigation planning within the energy supply processes, whilst also assessing the online reliability.

- **Chapter 5:** A generic framework of Dynamic Bayesian Network (DBN) based Markovian deterioration model was proposed to predict the health condition of the system and to clarify the behaviour of failure probability, considering the exogenous undisciplinable perturbations. The DBN is then extended to an ID for decision making regarding the optimum maintenance interval as well as the maintenance type. This will assist the operators and risk and safety managers for a more robust risk analysis and maintenance decision making to improving the lifetime reliability and availability of the industrial operations. An example of Natural Gas Regulating and Metering Stations (NGRMS) is given to show how the application of developed ID in risk-based maintenance optimization can be applied. To demonstrate the applicability of the methodology, three cases of seasonal observations of specific PV (pressure) are considered.

This thesis helps policymakers and asset managers move from a cost-based to a value-based approach to asset management, from traditional maintenance policies to intelligent maintenance.

Keywords: Reliability Engineering, Informed decision making, complex engineering systems, Bayesian inference, smart predictive maintenance.

[Page intentionally left blank]

[Page intentionally left blank]

Declaration and Statements

Declaration of Originality

I declare that this is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been duly acknowledged in the text and a list of references if given.

Signature: Date: 31/10/2019

Authority of Access

This thesis may be made available for loan and limited copying in accordance with the Copyright Act 1968.

Statement Regarding Published Work Contained in the Thesis

The publishers of the papers comprising Chapters 2 to 6, inclusive, hold the copyright for that content, and access to the material should be sought from the respective journals. The remaining non published content of the thesis may be made available for loan and limited copying and communication in accordance with the above Statement of Access and the Copyright Act 1968.

Statement of Ethical Conduct

To develop the methodologies and tools data collection from the seafarers around the globe was required in this PhD research. Therefore, a human research ethics approval was obtained from the University of Tasmania's human research ethics committee (Ethics Ref No: H0015701).

Acknowledgement

I would like to express my special appreciation and thanks to my advisor Professor Filippo De Carlo, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been invaluable. Without your guidance and constant feedback this PhD would not have been achievable.

A special thanks to my family. Words cannot express how grateful I am to my father and mother. Your prayer for me was what sustained me thus far. I am very thankful to my Brothers, Farshad and Mohammad, and my beloved fiancé, Roxana, for all of the sacrifices that you've made on my behalf.

I gratefully acknowledge the funding received towards my PhD from the DIEF/UNIFI (Department of Industrial Engineering, University of Florence) PhD fellowship. Thanks to Professor Mario Tucci for his encouragement and supervisory role and also Doctor Rouzbeh Abbassi for his valuable input. I am also grateful to the funding received through the ESTRA Ltd. to undertake my PhD.

I greatly appreciate the support received through the collaborative work undertaken with the Delft University of Technology (TU Delft) the Netherlands, during my visit program – thank you to Professor Pieter Van Gelder, Doctor Mohammad Mahdi Abaei. My thanks also go out to the support I received from the collaborative work I undertook with the RAMS Group at Department of Mechanical and Industrial Engineering of Norwegian University of Science and Technology (NTNU), Trondheim, Norway. I am especially grateful to Doctor Nicola Paltrinieri for believing in my research and for the financial support I received through the visit program. I am also very grateful to all those at the IBIS Laboratory in Department of Industrial Engineering, who were always so helpful and provided me with their assistance throughout my PhD program.

Dedication

Dedicated to the kindest teacher ever, my beloved mother **Afsaneh Shahrashoub,** the gentlest colonel, my amiable uncle **Reza BahooToroody** and my best friend’s grandfather **Amin-Allah Rafiei** whom all passed away during this truly life-changing experience.

“What doesn't kill you, makes you stronger”

Friedrich Nietzsche

“What won’t make you stronger, would kill you”

Unknown

Thesis by Journal Articles

The following four published journal articles constitute the content of this thesis.

- Chapter 2 **A. BahooToroody**, M. M. Abaei, E. Arzaghi, G. Song, F. De Carlo, N. Paltrinieri, R. Abbassi. *Analogical Reasoning in Reliability Challenges of Repairable System Using Hierarchical Bayesian and Maximum Likelihood Estimation* (Submitted, Under review)

- Chapter 3: **A. BahooToroody,** M.M. Abaei, F. BahooToroody, F. De Carlo, R. Abbassi, S. Khalaj. *A Condition Monitoring Based Signal Filtering Approach for Dynamic Time Dependent Safety Assessment of Natural Gas Distribution Process.* Journal of Process Safety and Environmental Protection. 2019, Volume 123, 335-343. doi.org/10.1016/j.psep.2019.01.016

- Chapter 4: **A. BahooToroody,** M. M. Abaei, E. Arzaghi, G. Song, F. De Carlo, N. Paltrinieri, R. Abbassi. *A Bayesian Regression Based Condition Monitoring Approach for Effective Maintenance Analysis of Random Process*. (Submitted, Under review)

- Chapter 5: **A. BahooToroody,** M.M. Abaei, E. Arzaghi, F. BahooToroody, F. De Carlo, R. Abbassi. *Multi-level optimization of maintenance plan for natural gas system exposed to deterioration process.* Journal of hazardous materials, 2019, Volume 362, 412-423. doi.org/10.1016/j.jhazmat.2018.09.044

Co-Authorship for Journal Articles

Journal articles co-authorship are as follows:

Conceived idea and designed the case Study: BahooToroody, De Carlo

Performed the case studies: BahooToroody

Analysed the data: BahooToroody

Wrote the manuscript: BahooToroody

Manuscript evaluation and submission: BahooToroody, De Carlo, Abaei, Abbassi, Arzaghi, Paltrinieri, Van Gelder

Dr. Filippo De Carlo

Department of Industrial Engineering

University of Florence

Florence, Toscana, Italy

Signature: Date: 31/10/2019

Mr Ahmad BahooToroody

Department of Industrial Engineering

University of Florence

Florence, Toscana, Italy

Signature: Date: 31/10/2019

List of Symbols and Abbreviations

$f(t)$	probability density function
$F(t)$	cumulative distribution function
t	time (sec)
MLE	Maximum Likelihood Estimation
$\bar{T}$	mean of the inter-arrival times
PR	Perfect Repair
T_f	successive failure times
n	total number of failures
$MTTF$	Mean Time To Failure
$E[N(t)]$	expected number of failures
$\lambda(t)$	rate of failure limit
HBM	Hierarchical Bayesian Modelling
P	pressure (kPa)
BN	Bayesian Network
DBN	Dynamic Bayesian Network
MR	Minimal Repair
x	time series data
c_k	k-th IMF
r	residue function
$c_{T,k}$	k-th IMF of true signal
$c_{N,k}$	k-th IMF of noise data
$T_i{}'$	observed exceedances time of the limit
T	mean period
m	number of raw data points
P_k	number of peaks
E	energy density
$E(NE)$	expected number of exceedances
PV	Process Variable

K	State of process variables
CF	Cost of Failure
CR	Cost of Replacement
CR'	Cost of Repair
CPT	Conditional Probability Table
Ω	Sensor uncertainty
Ɛ	Perturbations
K′	state of process variable before and after maintenance
O	Observation
M	Decision alternative
UF	Utility of Failure (Cost associated with failure)
UM	Utility of Maintenance (cost associated with maintenance)
PoD	Probability of Detection
λ_D	expected value of smallest detectable perturbation
D	Actual perturbation
h	average ROCOF
i_f	inflation rate of failure (percentage)
i_r	inflation rate of replacement (percentage)
$i_{r'}$	inflation rate of repair (percentage)
α_i	perturbation parameter
β_i	perturbation parameter
α, A	shape parameter
β, B	scale parameter
P_0	maximum probability of detection
LS	Limit state function
G	Failure function
C	critical PV interval
PR	actual PV interval

Content

List of Figures

List of Tables

[Page intentionally left blank]

1. Introduction

1.1 Background

Asset management is not limited to managing the maintenance and securing a reliable safety level for the system or process as usually thought, but also includes the kinds of people, knowledge and skills an organization needs. The asset management system is, therefore, defined as the ability to effectively extract value from its assets throughout the whole life of engineering systems. Consequently, the ongoing researches on this topic are classified into three strands; first to measure the performance of given assets. This cluster known as strategic asset management. To address this issue, the correlations and associations of all elements, within the assets of given engineering process, have to be treated (Hanski et al., 2019; Márquez et al., 2020; Too et al., 2006). The second division focused more on technical aspects, including the type of data required to be collected from the operation, how to inferred the information from collected data, how to make an informed decision based on available information to optimize the maintenance policy or to assess the required budget to spend on the maintenance investments (Aremu et al., 2018; Dewey et al., 2019; Liang et al., 2019; Liu et al., 2020; Rasmekomen et al., 2016). Finally, the third strand devoted to the quality of information. This is how to manage the data to support the asset management (Heaton et al, 2019; Raouf et al., 2006; Srinivasan et al., 2013; Zhang et al., 2007).

Accordingly, given aforementioned strands, this PhD thesis is devoted to investigate the risk-based asset integrity model using advanced mathematical models. The methodologies and tools developed in this PhD research can be applied to engineering operations on any of devices and structures.

1.2 Research Objectives and Research Questions

The primary objective of this PhD research project is to improve the current state of scientific knowledge regarding the smart maintenance in engineering process. This is particularly to be addressed through the following objectives:

- to develop a comparison model for enabling industry on indicating the possible differentiation in failure assessment of random process under the assumption constraint.
- to develop a framework for predicting failure mode given the noise associated with the operational data.
- to address the involved and, most of the time, unconsidered risk to make a prediction of safety conditions through the autonomous operation in future.
- to develop a decision making framework for maintenance time schedule optimization using Markov degradation model.

In order to achieve the research objectives in this PhD research, the proposed work is broken down into a number of specific research questions:

- How to highlight the magnitude of the deviation value in different failure modelling approaches?
- How to identify and filter out the noise with its time-dependent auto-correlated structure from the operational data?
- How to address the involved and, most of the time, unconsidered risk to make a prediction of safety conditions through the autonomous operation in future?
- How to optimize the maintenance time schedule given Markov degradation model?

1.3 Scope and limitations

The focus of this PhD research is to develop new methodologies for intelligent predictive maintenance of engineering operations. The presented frameworks are proposed based on the condition monitoring data of the autonomous operation to predict the reliability of the operation. The scope of this research are:

- firstly, to highlight the magnitude of the deviation value in different failure modelling approaches. Hierarchical Bayesian modelling (HBM) and Maximum Likelihood

Estimation (MLE) approaches are applied to investigate the effect of utilizing observed data on inter-arrival failure time modelling. The results highlight that relaxing the renewal process assumption and taking the time dependency of the observed data into account will result in more precise failure models.

- secondly, to propose a probabilistic model to predict the system reliability given the complexity of the non-stationary behavior in monitoring data. To this end, an integration of Empirical Mode Decomposition (EMD) and Hierarchical Bayesian Model (HBM) is developed. Based on observation data acquired from the case study, the model is able to predict failures when consideration is given to uncertainty quantification.
- thirdly, to present a probabilistic framework for risk mitigation planning within the energy supply processes, whilst also assessing the online reliability. For this purpose, the canonical link function has been employed to describe the interactions and association between risk factors and safety indicators. Bayesian inference with hierarchical structure has been inferred for the calculations of regression function including interpretation of the intercept and coefficient factors.
- finally, to develop a decision model for intelligent predictive maintenance schedule. Given the exogenous undisciplinable perturbations associated with time series data, a DBN based deterioration model was proposed to extend the remaining useful lifetime of the operations. Discrete time case is considered through measuring or observing the PVs. Decision configurations and utility nodes are defined through the model to represent maintenance activities and their associated costs.

The methodologies and tools developed in this study are able to be adopted to any type of engineering process from oil and gas to renewable and offshore operations. This PhD research can quantify the uncertainties associated with prediction of failures or defects, thus opening the door to risk remediation planning and smart predictive maintenance.

1.4 Organization of the thesis

This thesis is written in manuscript format (paper-based). A summary of the thesis outline is provided in the section below. To a large extent these chapters are independent and can be read individually.

Chapter 2: On Reliability Challenges of Repairable Systems Using Hierarchical Bayesian Inference and Maximum Likelihood Estimation

Chapter two explores the magnitude of the deviation value in different failure modelling approaches. The resulted methodology is capable of presenting a comparison model for enabling industry on indicating the possible differentiation in failure assessment of random process under the assumption constraint. This study can help asset managers to optimize the reliability assessment of repairable systems based on available data.

Chapter 3: A Condition Monitoring Based Signal Filtering Approach for Dynamic Time Dependent Safety Assessment of Natural Gas Distribution Process

This chapter presents a dynamic failure model based on noisy monitoring data. The developed model can successfully predict failure mode thus directly avoiding unnecessary maintenance and safety consequences which is because of I) analyzing the observational data given their non-stationary and nonlinear nature in time series, II) Filtering the noise associated with the raw data considering the time-dependent auto-correlation structure of noise, III) Observing the correlation and variability of data source over the time.

Chapter 4: A Bayesian Regression Based Condition Monitoring Approach for Effective Reliability Prediction of Random Processes in Autonomous Energy Supply Operation

Chapter four proposes an online reliability assessment for an ongoing engineering process. Assuming that the intercept and coefficients are uncertain, a generalized linear model (GLM) were employed to offer the explanatory model as a regression function to describe the interactions and association between risk factors and safety indicators. The proposed model can provide a risk remediation plan based on developed online reliability.

Chapter 5: Multi-Level Optimization of Maintenance Plan for Natural Gas System Exposed to Deterioration Process.

Through the fifth chapter, a risk-based optimization methodology for a maintenance schedule considering Process Variables (PVs), within the framework of asset integrity assessment was proposed. Accordingly, Dynamic Bayesian Network, Damage Modelling and sensitivity analysis was integrated to clarify the behavior of failure probability, considering the exogenous undisciplinable perturbations. The proposed methodology could either analyse the failure based on precursor data of PVs or obtain the optimum maintenance schedule based on actual condition of the systems.

.

2. On Reliability Challenges of Repairable Systems Using Hierarchical Bayesian Inference and Maximum Likelihood Estimation

Abstract

Failure modelling and reliability assessment of repairable systems has been receiving a great deal of attention due to its pivotal role in risk and safety management of process industries. Meanwhile, the level of uncertainty that comes with characterizing the parameters of reliability models require a sound parameter estimator tool. For the purpose of comparison and cross-verification, this paper aims at identifying the most efficient and minimal variance parameter estimator. Hierarchical Bayesian modelling (HBM) and Maximum Likelihood Estimation (MLE) approaches are applied to investigate the effect of utilizing observed data on inter-arrival failure time modelling. A case study of Natural Gas Regulating and Metering Stations in Italy has been considered to illustrate the application of proposed framework. The results highlight that relaxing the renewal process assumption and taking the time dependency of the observed data into account will result in more precise failure models. The outcomes of this study can help asset managers to find the optimum approach to reliability assessment of repairable systems.

Keywords: Repairable system, Failure modelling, Time dependency, Hierarchical Bayesian Analysis, Maximum Likelihood Estimation

2.1 Introduction

Failure time modelling of repairable components has attracted a great deal of attention owing to the high level of risk associated with the failure events occurring within process industries. Using statistical inference, different probability distributions are adopted to model the rate of occurrence of failures (ROCOF). These probability distributions are characterized by one or more parameters. The parameter estimation process may be implemented based on different assumptions regarding maintenance strategies including Perfect Repair (PR) or Minimal Repair (MR).

PR represents an ideal model in which the time between successive failures of a given system are independent and identically distributed (iid) random variables. Although PR is recognized as the most applied assumption by a number of researches (Nandi et al., (2005), Quy et al., (2006), Toroody et al., (2016a), Quy et al., (2008), Louit et al., (2009) Toroody et al., (2016b) and Barabadi et al., (2014)), neither of these have accounted for the system to be "as bad as old" after repair. In the present paper, it is illustrated that analyzing failure times given PR (also known as renewal process) often yields improper results and subjects to a significant level of uncertainty.

A comprehensive reliability analysis must include a time dependent study, if the system is degrading or improving. Therefore, an ongoing effort on reliability assessment based on MR are carried out (Majeske 2007, Slimacek and Lindqvist 2016, Antonov and Chepurko 2017, Peng, Shen et al. 2018, Sheu et al., 2018). Li et al. (2017) used two recurrent-event change-point models arisen from a non-homogeneous Poisson process (NHPP) to find the time of change in driving risk. In another recent study, Pesinis and Tee (2017) presented a model for reliability analysis of failure data based on NHPP incorporated with a robust structural reliability model. Furthermore, an extensive review of PR, MR and probabilistic knowledge elicitation made with a wide range of engineering applications is presented by Crow (1975), Asher and Feingold (1984) and most recently by Ross (2014) and Modarres et al., (2016).

In the present paper, two mathematically robust and efficient approaches are implemented to represent the state of system after repair. Maximum Likelihood Estimation (MLE) and Hierarchical Bayesian Modelling (HBM) are established based on actual data in order to predict the likelihood of studied failures, given PR and MR assumptions. The capability of HBM in modelling the variability of non-stationary data and the correlation between nonlinear data via open source Markov Chain Monte Carlo (MCMC) sampling software packages, i.e., OpenBUGS (Spiegelhalter et al. 2007), have resulted in its widespread use in engineering applications, e.g. probabilistic risk assessment and condition monitoring (Behmanesh et al., 2015, Chitsazan et al., 2015, Yu et al., 2017, Mishra et al., 2018). Recently, Abaei et al. (2018) developed an HBM for safety assessment of vessels crossing shallow-waters based on time-domain hydrodynamic simulations. There is also a great deal of methods developed based on MLE that show the applicability of this method in risk and reliability assessment of complex engineering systems, examples of which are structural degradation modeling, risk-based maintenance planning, geotechnical risk assessment, etc. (Straub 2009, Arzaghi et al. 2017, Abaei et al. 2018, Leoni et al. 2018, BahooToroody et al. 2019). Nielsen and Sørensen (2017) estimated the remaining useful lifetime (RUL) of a wind turbine to calibrate a Markovian deterioration model based on MLE approach.

Different assumptions (e.g. MR, PR) and tools (e.g. MLE, HBM) lead to distinct results which are discussed here for the purpose of comparison and cross-verification. Accordingly, this paper aimed at presenting a comparison model for enabling industry on indicating the possible differentiation in failure assessment of random process under the assumption constraint. Consequently, the magnitude of the deviation value in different failure modelling approaches is highlighted. The developed framework in this study opens the door for the use of engineering researchers in risk analysis and reduction plan throughout the industries.

The remainder of present paper is structured as follows: the procedure of models specification and overview is presented first, followed by a characterization of the application of the developed methodology in a Natural Gas Regulating and Metering Stations (NGRMS) in Florence, Italy. In Section 2.10, the results and discussions including a comparison of the investigated methods are presented, while Section 2.11 provides the concluding remarks of this research.

2.2 Assumptions

The outcomes of a maintenance plan, which is the condition of repaired systems, can be modelled stochastically throughout the operation. The division of repair categories is made based on a number of factors including whether the failure interarrival times are dependent over the asset operational time or not. The differentiation formed based on such factor is explained through the following assumptions:

2.2.1 Perfect Repair

The renewal process belongs to the class of stochastic point processes where inter-arrival times are assumed to be *iid* random variables. In this category, any repair originated by a failure in the system is assumed to be perfect and subsequently the system is said to be "as good as new". The expected number of failures, $E[N(t)]$, in time, t, is defined as renewal function given by Equation 2.1:

$$E[N(t)] = m(t), \qquad m(t) = F(t) + \int F\left(t - T_f\right) dm\left(T_f\right) \tag{2.1}$$

where T_f is successive failure times, $N(t)$ is the number of failure and $F(t)$ is the cumulative distribution function (CDF) of T. Measuring the changes of variables in both sides of Equation 1 with respect to the change of time results in Equation 2.2:

$$m'(t) = \lambda(t), \qquad \lambda(t) = f(t) + \int_0^t f\left(t - T_f\right)\lambda\left(T_f\right) dT_f \tag{2.2}$$

where $f(t)$ is the corresponding probability density function (PDF) of successive inter-arrival times, T_f. One of the most celebrated renewal processes including *iid* assumption is the Homogeneous Poisson Process (HPP) (Louit et al., 2009, Barabadi et al., 2014, Hajati et al., 2015) which is recognized by the method presented in this paper.

2.2.2 Minimal Repair

Based on the HPP assumption, the failure rate will be independent of time. However, in reality the system condition in i^{th} time-step is dependent on its condition in time-step t_{i-1}. Relaxing the *iid* assumption leads to the Nonhomogeneous Poisson Process (NHPP) in which the system retains the "as bad as old" condition following a relatively instant repair action. Implementing MR, the observation process can be carried out either failure-truncated or stopped in a fix time. The calculation methods are similar in both mentioned approaches and here the failure-truncated case will be adopted as recommended by Kelly and Smith (2009). The expected number of failures through the specific time interval, $[t_n, t_{n+1}]$, $E[N(t)]$, is given using Equation 2.3.

$$E[N(t)] = \int_{t_n}^{t_{n+1}} \lambda(t)dt \tag{2.3}$$

where an appropriate functional form for ROCOF, $\lambda(t)$, must be determined to represent the expected number of failures, accordingly. For this purpose, power-law, log-linear and linear models are suggested in the literature (Kelly and Smith 2009, El-Gheriani et al., 2017). Power-law model is one of the most common forms of ROCOF in reliability assessment (Abaei et al., 2018) as it can predict the nonlinearity of the stochastic process with reasonable precision. The relationship for power law is given by Equation (2.4).

$$\lambda(t) = \frac{\alpha}{\beta}\left(\frac{T}{\beta}\right)^{\alpha-1} \tag{2.4}$$

According to Arzaghi et al. (2018), the inter-arrival of times between successive failures, T, in the power-law process follows a Weibull distribution, $f(t,\beta,\alpha)$, with shape parameter, α, and scale parameter, β, given by:

$$f(t,\beta,\alpha) = \frac{\alpha}{\beta}\left(\frac{T}{\beta}\right)^{\alpha-1} \exp[-(T/\beta)^{\alpha}] \tag{2.5}$$

2.3 Parameter estimation methods

Observed data, manipulated information, and gathered knowledge are three consecutive steps of making inference. The effectiveness of a specific model must be examined, that is, how well the model fits the collected data. This question is answered through the process of parameter estimation. In case of reliability analysis, not only the assumptions but also the methods of parameter estimations are of high importance affecting the accuracy level of final results. Once the assumption is specified, and the data is observed, the mathematical method for estimating the parameter of interest should be established. In the present study, MLE and HBM as the most popular choices of model fitting in reliability assessment are utilized, as recommended by previous researchers (Mahadevan and Rebba 2005, Neil et al., 2008, Rebba and Mahadevan 2008, Peng et al., 2013). A brief introduction to these methods can be found in the following sections.

2.3.1 Maximum Likelihood Estimation (MLE)

Assuming that vector $g = (g_1, \dots, g_n)$ is a random sample of an available data source, MLE is performed to predict the most likely data source that would yield the random sample, g. For this purpose, it is necessary to identify both the appropriate distribution of data source and its corresponding parameters. Consider that $x = (x_1, \dots, x_n)$ is a vector specified within the parameter space, therefore the PDF of the data vector, g, would be achieved by Equation 6, given by:

$$f(x_1, x_2, \dots, x_n | g) = f_1(x_1|g) f_2(x_2|g) \dots f_n(x_n|g) \tag{2.6}$$

2.3.2 Hierarchical Bayesian Model (HBM)

A summary of the process for performing inference using data and a probabilistic model is presented in Figure 2.1. As shown in this figure, the raw data are the collected values from a process. Evaluation of the data results in information and knowledge is obtained by gathering information. The process of making conclusion based on what once knows is referred to as inference. There is a need for models for obtaining information based on raw data. The models available for this purpose can be categorized into deterministic or probabilistic approaches (Kelly and Smith 2009). In this regard, probabilistic models are able to represent the uncertainty associated with available data where a HBM approach will assist in achieving the posterior distribution of the parameters of interest. HBM is carried out based on the Baye's theorem, given by Equation 2.7 (El-Gherian et al., 2017):

$$\pi_1(\theta|x) = \frac{f(x|\theta)\pi_0(\theta)}{\int_\theta f(x|\theta)\pi_0(\theta)d\theta} \tag{2.7}$$

where θ denotes the unknown parameter of interest, $\pi_1(\theta|x)$ is the posterior distribution, and $f(x|\theta)$ is the likelihood function. HBM utilizes multistage prior distributions for the parameter of interest indicated by $\pi_0(\theta)$ (Abaei et al., 2018) as follow:

$$\pi_0(\theta) = \int_\emptyset \pi_1(\theta|\varphi)\, \pi_2(\varphi) d\varphi \tag{2.8}$$

where, $\pi_1(\theta|\varphi)$ is the first-stage prior as the population variability in $\theta; \varphi$ denotes a vector of hyper-parameters, (e.g. $\varphi = (\alpha, \beta)$) , while α and β are the shape and scale parameters of a

Weibull distribution, respectively. The uncertainty in φ is represented by $\pi_2(\varphi)$ as the hyper-prior distribution. The prior distribution, $\pi_0(\theta)$ is specified using generic data collected from different sources (numerical simulations, experiments or collected from industrial operations) to estimate the posterior distribution (Abaei et al., 2018).

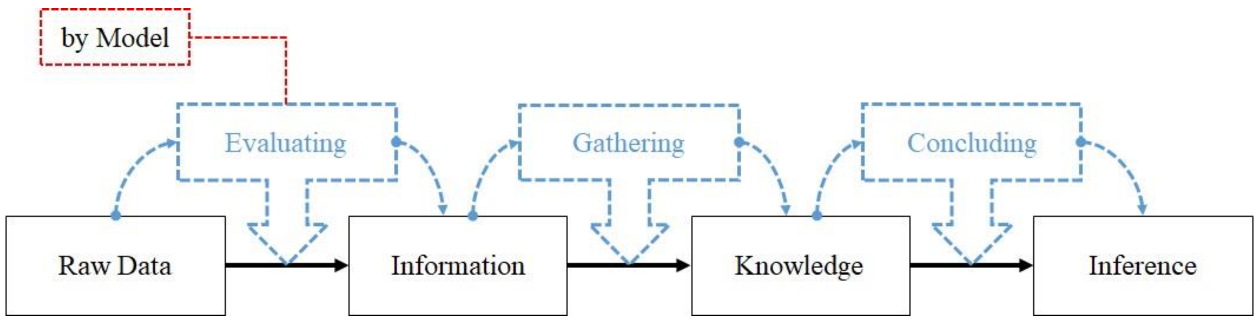


Figure 2.1 An overview of inference process and its key elements.

2.4 Model specification

In order to predict the condition of a system after it has undergone a repair, the precise and powerful mathematical approaches are established. A well-known parameter estimator (MLE) and the recent advances in Bayesian statistical methods (HBM) are accounted for revealing the gap between PR and MR. Based on the numbers of parameter required for characterization of distribution of failure time, Weibull and Exponential distributions are used to perform the analysis. Particularly, with a NHPP assumption, the time to failure cannot be characterized by an exponential distribution where $\alpha = 1$. So, a two-parameter Weibull distribution is required (Dar et al., 2015, Pesinis and Tee 2017). This is while for a HPP, Exponential distributions can be employed, as recommended by Hajati, Langenbruch et al. (2015) and Kumar and Chakraborti (2015).

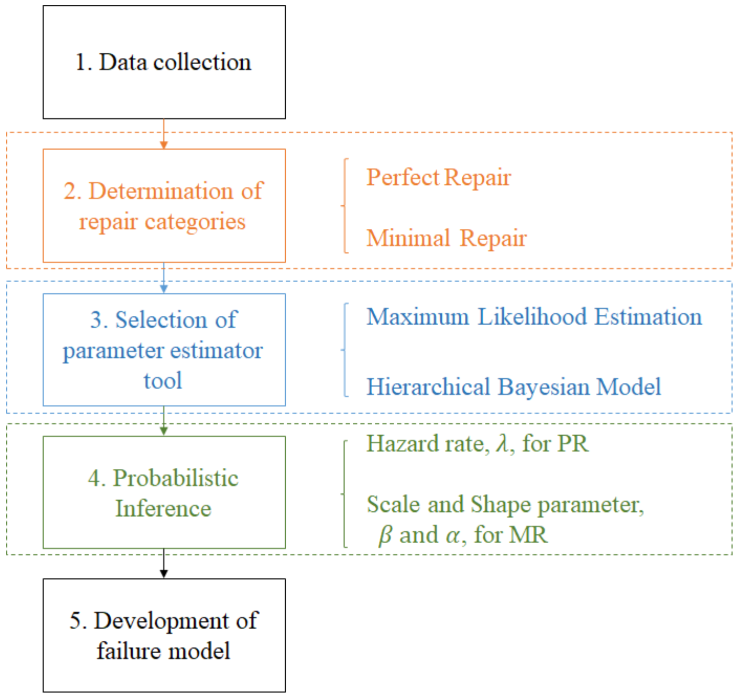


Figure 2.2 Developed framework for failure modelling based on different repair categories and parameter estimator tools.

2.5 Homogenous Poisson Process

2.5.1 Maximum Likelihood Estimation (MLE)

As discussed in Section 2.2.1, given a perfect repair condition, the probability distributions of failure inter-arrival times, denoted by T, are expressed by Equations. (2.9) and (2.10):

$$F(t) = 1 - e^{-\lambda t} \tag{2.9}$$

$$f(t) = \lambda\, e^{-\lambda t} \tag{2.10}$$

where λ is defined as the rate parameter with a likelihood function expressed by Equation (2.11):

$$L(\lambda) = \prod_{i=1}^{n} \lambda\, e^{-\lambda T_i} = \lambda^n e^{\left(-\lambda \sum_{i=1}^{n} T_i\right)} = \lambda^n\, e^{(-\lambda n \bar{T})} \tag{2.11}$$

where $\bar{T}$ is the mean of the inter-arrival times and *n* is the total number of failures observed. Thus, the maximum likelihood of rate parameter λ, is given by Equation (2.12) (Ross 2014):

$$\hat{\lambda} = \frac{1}{\bar{\bar{T}}} = \frac{n}{\sum_i T_i} \tag{2.12}$$

It is anticipated that by using this approach to uncertainty modelling, the obtained exponential distribution would be a representative of the failure inter-arrival times during the future operations of studied system.

2.5.2 Hierarchical Bayesian Model (HBM)

For a system with failure events that follow a Poisson distribution, the number of failures, x, can be modelled by Equation (2.13):

$$f(x|\lambda) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}\ , \quad x = 0,1, \ldots \tag{2.13}$$

where t is the exposure time and λ is the intensity of the Poisson distribution. For a HPP, a gamma distribution can be utilized to describe the variability of λ among the observed failure times. Therefore, given the hyper-parameters α and β, the first-stage prior distribution can be achieved by the Gamma distribution, as expressed by Equation (2.14) (Siu and Kelly 1998):

$$\pi_1(\lambda|\alpha,\beta) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \tag{2.14}$$

Diffusive Gamma distribution is applied independently to model the prior distribution of hyper-parameters, as suggested by El-Gheriani, Khan et al. (2017).

It should be noted that although the hyper-parameters are considered as independent random variables prior to any observations, they become dependent as soon as observations are introduced. According to Kelly and Smith (2009), this dependency is accounted for by the joint posterior distribution.

Once the model, including the prior distribution and likelihood functions, are developed, MCMC simulations are carried to predict the posterior distribution of the Gamma parameters, α , β. This results in the estimation of Exponential distribution with a rate parameter of λ, and its associated uncertainty.

2.6 Non-Homogenous Poisson Process

2.6.1 Maximum Likelihood Estimation (MLE)

The method of estimating the probability distribution of failure inter-arrival times, based on a MR assumption, is explained earlier in section 2.2.2. In order to obtain the ML of Weibull parameters, the recommended likelihood function is given by Equation (2.15) (Asher and Feingold 1984):

$$L = \prod_{i=1}^{n} f(T_i) \tag{2.15}$$

where T_i is the time at which the i^{th} failure has occurred and n is the total number of failures. The ML of shape and scale parameters α, β are given by Equations (2.16) and (2.17) (Crow 1975):

$$\hat{\alpha} = \frac{T_n}{n^{1/\beta}} \tag{2.16}$$

$$\hat{\beta} = \frac{n-1}{\sum_{i=1}^{n-1} \ln\left(T_n/T_i\right)} \tag{2.17}$$

where T_n is the time at which last failure, n, has occurred.

2.6.2 Hierarchical Bayesian Model (HBM)

In order to reflect on the dependency of the inter-arrival times, T_i, a conditional probability must be established. This probability for the time interval $[t_{i-1}, t_i]$ can be expressed by Equation 2.18 (El-Gheriani et al., 2017).

$$f(t_i|t_{i-1}) = f(t_i|T_i > t_{i-1}) = \frac{f(t_i)}{\Pr(T_i > t_{i-1})} \tag{2.18}$$

Consequently, the Weibull distribution and the corresponding likelihood function are given by Equations. (2.19) and (2.20), respectively.

$$f(t_i|t_{i-1}) = \frac{\alpha}{\beta^{\alpha}}(t_i)^{\alpha-1}e^{\left[\left(\frac{t_{i-1}}{\beta}\right)^{\alpha}-\left(\frac{t_i}{\beta}\right)^{\alpha}\right]} \tag{2.19}$$

where $i = 2, \dots, n$.

$$f(T_1, T_2, \dots, T_n|\alpha, \beta) = f(T_1)\prod_{i-2}^{n} f(t_i|t_{i-1}) \tag{2.20}$$

where $T_1$and T_n are the times of first and *n*th failure events. Furthermore, the uncertainty of parameters α and β are modelled by HBM representing the variability of failure inter-occurrence times. Similar to the HPP case, these parameters will become inter-dependent once observations are made. The likelihood function is not pre-programmed into MCMC sampling software packages, OpenBUGS. Based on the suggestions provided by Kelly and Smith (2009), the likelihood function, φ, which is a vector of n array can be assigned to the model. This function $\varphi = log(likelihood)$ as defined by Equation 2.21 adopts samples of α and β from the prior distribution in Equation 22 (Abaei, Arzaghi et al. 2018).

$$\varphi = \log(\alpha) - \alpha \times \log(\beta) + (\alpha - 1)\log(T_i) - (T_n/\beta)^{\alpha}/n \tag{2.21}$$

$$\begin{cases} \alpha \sim Gamma(0.0001{,}0.0001) \\ \beta \sim Gamma(0.0001{,}0.0001) \end{cases} \tag{2.22}$$

where T_n and T_i are the last and i^{th} observation of the failure times in the simulation, and n is the vector size. As similar to HPP, an independent diffuse is assumed for the prior distribution of hyper-parameters (El-Gheriani et al., 2017). The updated posterior distribution of the hyper-parameters obtained from the MCMC sampling using the observed data are inserted into the Weibull distribution, $f(t, \beta, \alpha)$, in order to estimate the PDF of failure under MR assumption.

2.7 Application of methodology

In order to demonstrate the application of the developed method and establish a comparison among the employed models, a practical example from the degradation process of Natural Gas Regulating and Metering Stations (NGRMS) operating in Italy is considered as the case study.

2.7.1 Scenario development

NGRMS is installed in a distribution network and supplied with natural gas flow through a (number of) transmission pipeline(s). Pressure reduction and gas flow measurements are the fundamental duties of these facilities that consist of five main sections including the inlet, filter, metering, regulator and outlet. In order to prevent any interruptions in the process caused by failures events, the redundant line is set up. A schematic of NGRMS is illustrated in Figure 2.3.

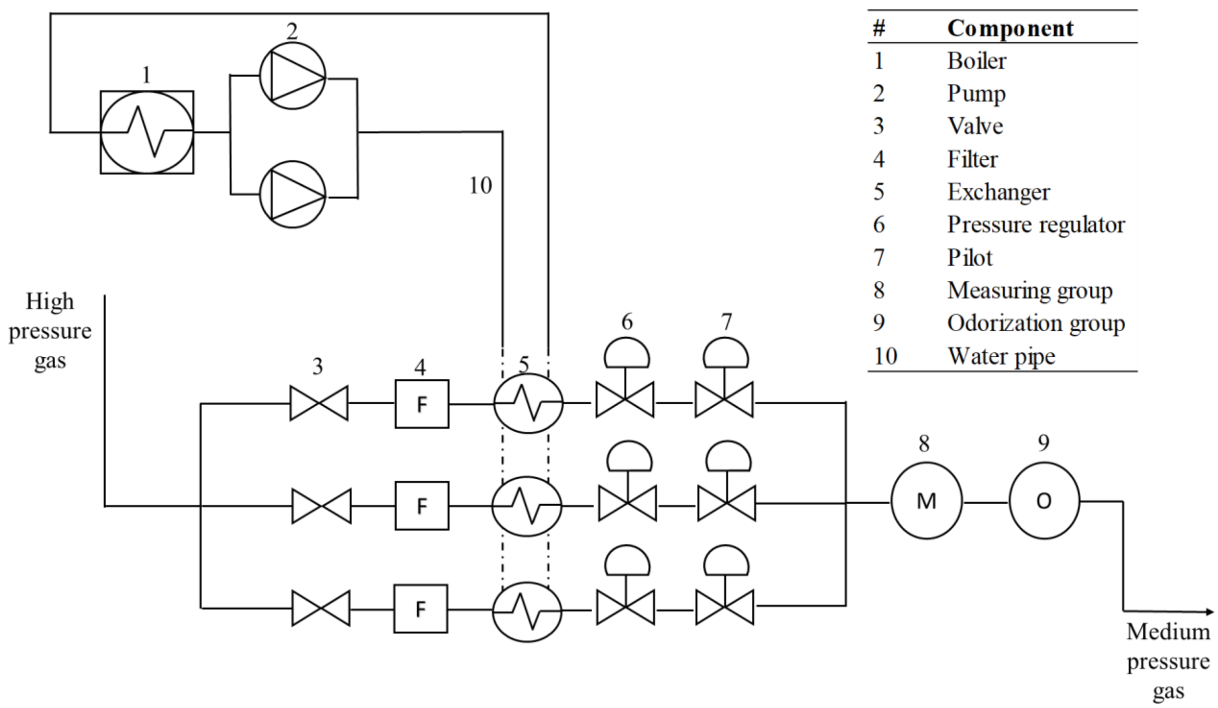
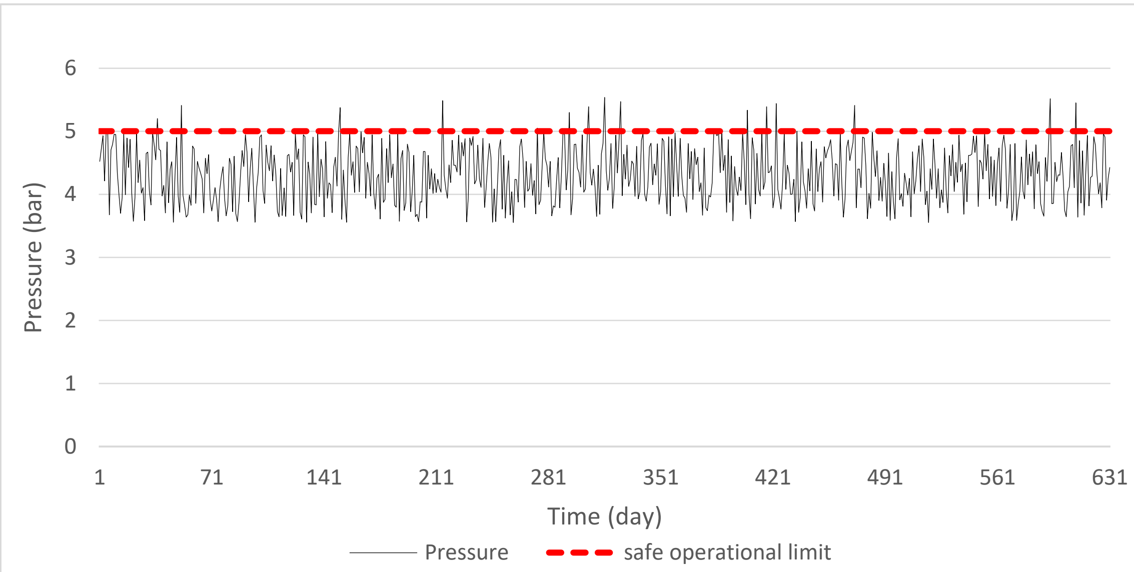
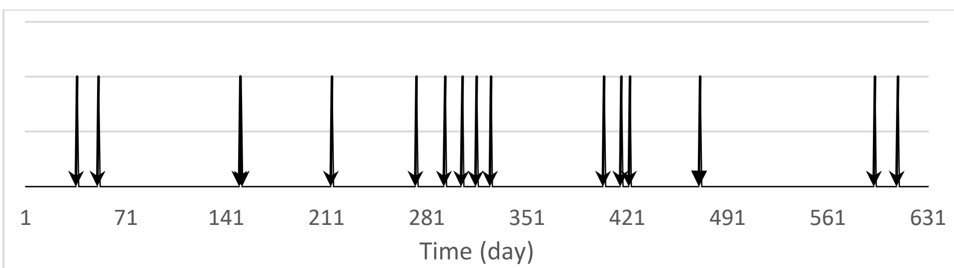


Figure 2.3 A schematic of Natural Gas Regulating and Metering Stations (Gonzalez-Bustamante et al., 2007).

A range of process variables characterize the health condition of the process in NGRMS, e.g. pressure, temperature and vibration. In this study pressure is considered as the variable of interest for the analysis of the degradation process. The failure of system is defined as an event where the value of pressure exceeds the desired safety limit of the operation. The recorded time series of operational, are illustrated in Figure 2.4. This figure also shows the observed failure times of system. It is worthwhile to mention that the random noise in operational pressure data is filtered from nonstationary and nonlinear raw data by adopting Empirical Mode Decomposition (EMD) method. The explanation of EMD is beyond the scope of this paper and readers are referred to the following researches for detailed discussions on this topic (Huang et al., 1998, Wu and Huang 2004, Wu et al., 2007, Li and Pandey 2017, BahooToroody et al., 2019).



(a)



(b)

Figure 2.4 (a) Time series of pressure data collected from NGRMS (b) Time to failure for given pressure values.

As depicted in Figure 2.4, pressure values are recorded during a 631-days operation where 15 failure events have occurred. These data have been utilized in the analysis process.

2.8 Homogenous Poisson Process modelling

2.8.1 Maximum Likelihood Estimation (MLE)

The presented method in section 2.5.1 is applied on the pressure data based on an HPP assumption. The ML of rate parameter is estimated as $\lambda = 0.0246$ (per day) resulting in an exponential PDF of failure inter-arrival time illustrated in Figure 2.5. The specifications of this exponential distribution including its Mean Time To Failure (MTTF) are provided later in the summary of statistical analyses (see Table 2.2).

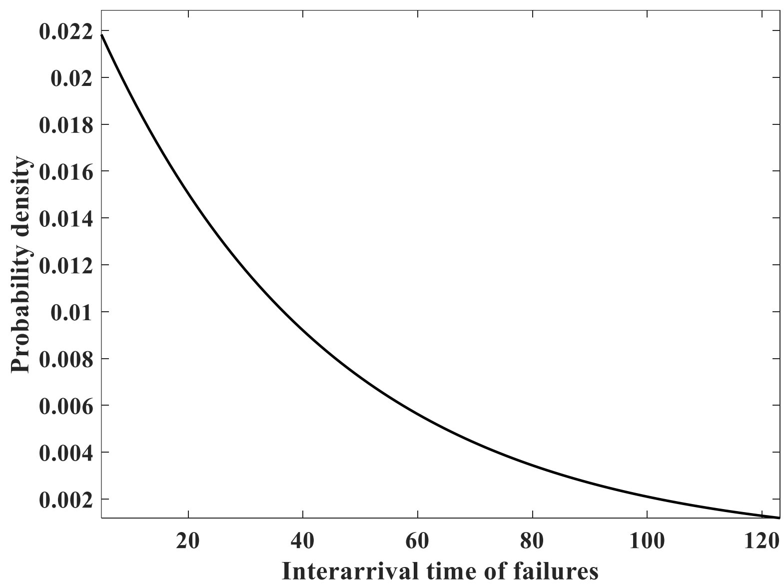


Figure 2.5 assigned Exponential probability distributions fitted on operational data given a perfect repair assumption.

2.8.2 Hierarchical Bayesian Model (HBM)

The failure rates of considered operation in Table 2.1, have been extracted from the pressure time-series. Three chains with over-dispersed initial values of α and β were used to ensure the convergence of simulations. In order to calculate the parameters of interest, the HBM was performed in OpenBUGS with 1000 burn-in iterations, followed by 300,000 iterations through each chain. The caterpillar plot of credible intervals of failure rate, λ, for all 12 interval is illustrated in Figure 2.6. The mean value of posterior predictive distribution of λ is 0.0243 per day (see Table 2.2). This average value of failure rate also incorporates an estimate of the interval-to-interval variability.

Table 2.1 Failure (pressure exceedance) rate data during NGRMS operation.

Region	No. of Failures	Exposure time (day)
1	2	52
2	0	57
3	1	50
4	0	53
5	1	56
6	5	55
7	0	54
8	3	45
9	1	50
10	0	54
11	0	53
12	2	49

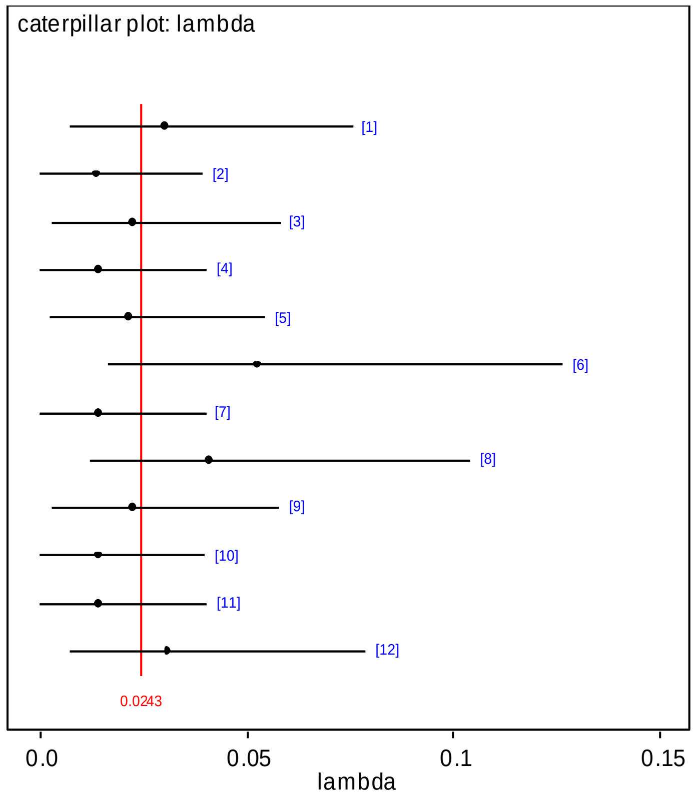


Figure 2.6 The posterior mean and 95% credible interval for pressure exceedance from safety limits over time given a PR assumption. Note: black dots are posterior means for each interval, the red line is average of posterior means.

2.9 Non-Homogenous Poisson Process modelling

2.9.1 Maximum Likelihood Estimation (MLE)

Given a NHPP, the failure inter-occurrence times, T, in the power-law process generate a Weibull distribution, $f(t, \beta, \alpha)$, with shape parameter, α, and scale parameter, β, which can be estimated using Equations (15-17) through the application of MLE. The maximum likely α and β are computed as 52.83 and 1.1069, respectively (for more details see Table 2.2). Figure 2.7 shows the resultant Weibull distributions of MLE on the available data.

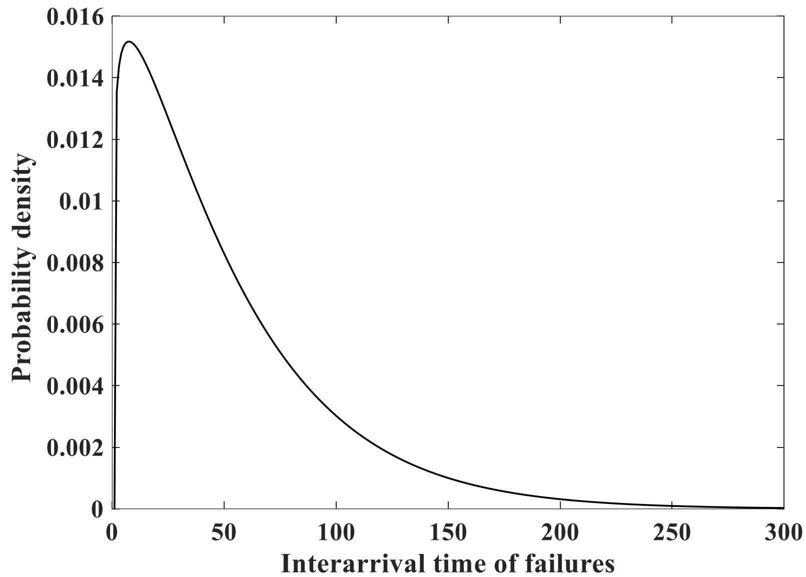


Figure 2.7 Obtained Weibull distribution of failure inter-arrival times considering a MR assumption.

2.9.2 Hierarchical Bayesian Model (HBM)

In order to estimate the likelihood function and posterior probability of Weibull parameters, α and β, the recorded failures were entered into the HBM. Similar to the HPP application, using MCMC simulations, three chains from separate points were assigned with 1000 burn-in iterations, followed by 300,000 iterations at each chain in order to ensure the convergence of the simulation and accurately predict the posterior probabilities of the parameters of interest. Figure 2.8 shows the predicted posterior distribution of α and β. The dynamic trace of Weibull parameters is depicted in Figure 2.9 confirming the convergence. A summary of the estimated marginal posterior distributions for α, β as well as their corresponding MTTFs are listed in Table 2.2.

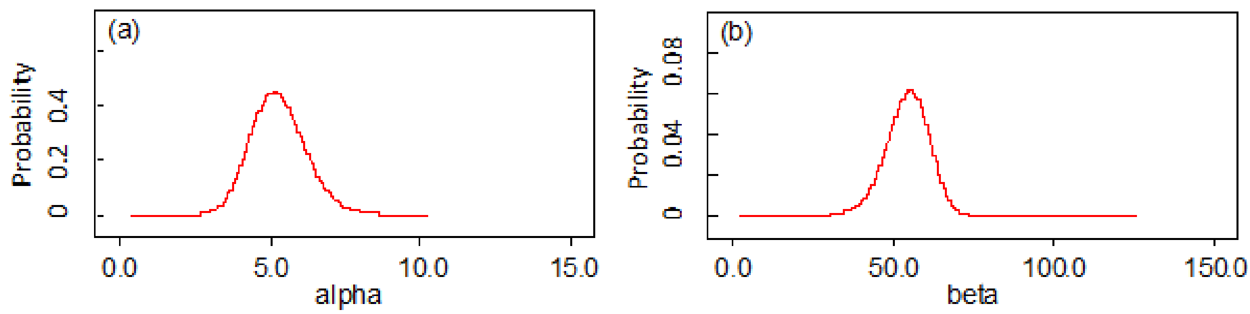


Figure 2.8 Posterior distributions of Weibull (a) shape parameter (b) and scale parameter.

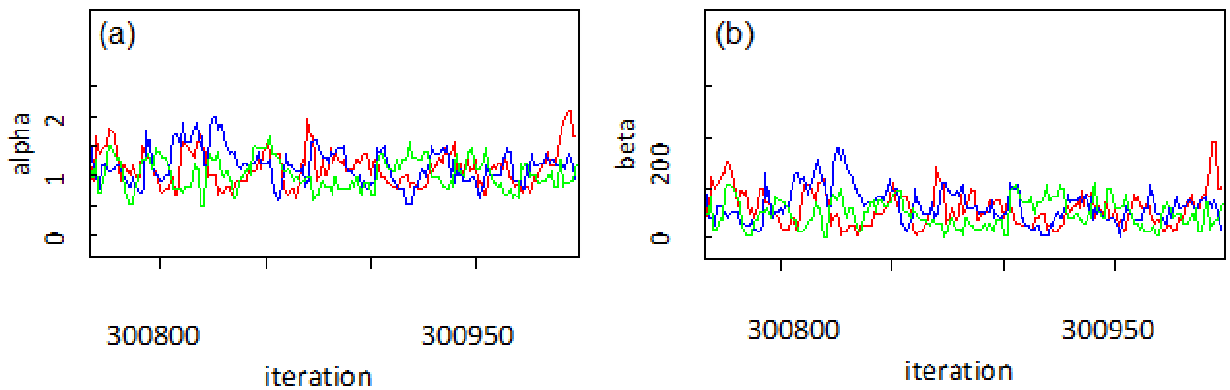


Figure 2.9 Dynamic trace of Weibull shape parameters (a) and scale parameter (b) in MCMC simulation.

2.10 Results and discussion

2.10.1 comparison

According to the presented models, a range of comparisons are drawn to illustrate the deviation of uncertainty quantifications throughout the characterized failure functions. To this end, Cumulative Distribution Functions (CDFs) of each failure modelling approach were developed, as illustrated in Figure 2.10. The estimated MTTF for each approach is also shown in their CDFs in this Figure. Table 2.2 summarizes the details of obtained results. An assumption-based approach covers the comparison between the two assumptions made regarding the distribution of inter-arrival times of failure events during the studied operation i.e. those with the iid assumption, represented by a HPP and those without this assumption which are modelled as a NHPP. The comparison of HBM and MLE reveals a difference in the estimated posterior distribution of parameters, which is attributed to the impact of correlation between the observed data. The results of these analyses are discussed in more details in the following sections.

2.10.1.1 Assumption-based comparison

The MTTF values, estimated by the MLE approach, are 40.67 and 50.87 days for PR and MR, respectively. The significant difference between the obtained MTTF is due to the fact that the PR assumption neglects the dependency amongst failure interarrival while MR accounts for this. As shown in Figure 2.10, a similar difference level is observable between the MTTF values estimated by using a HBM approach. The MTTF of HPP and NHPP are found to be 41.15 and 55.57 days, respectively. These results confirm that discounting the time dependency of failure events may lead to between 25% to 35% difference in MTTF, regardless of the modelling approach (HBM or MLE).

2.10.1.2 Parameter estimator-based comparison

Regarding the HBM, to allow the results to be compared with MLE, independent and diffuse priories were adopted for α and β. A gamma distribution prior was used for both shape and scale parameters, as suggested by Kelly and Smith (2009).

Based on PR, the MLE yields a failure rate of $\lambda_{MLE} = 0.0246$ with a 95% confidence interval of (0.0137,0.0385) while the posterior mean of this parameter is estimated by HBM as $\lambda_{HBM} = 0.0243$, having a 95% confidence interval of (0.0044,0.0631). In the light of estimated value for parameter of interest, lambda, given HPP, it is interpreted that the source to source variability of data carrying out by MCMC simulation in the HBM is less reflected. Subsequently, the MLE and HBM of MTTF are estimated at 40.67 and 41.15, respectively, suggesting a minimal difference (see Figure 2.10).

For a MR assumption, HBM yields a posterior mean for the shape parameter $\alpha = 1.107$ with a 95% credible interval of (0.605, 1.757) while the MLE model resulted in $\alpha = 1.106$, highlighting a good agreement between the employed approaches. Bayesian inference β, yields a posterior mean of 57.71. However, the MLE, which disregards the correlation between observed data, results in a smaller shape parameter of 52.83 (8% deviation). Finally, this deviation in the posterior mean of β in HBM and MLE is appeared again in the MTTF (see Table 2.2).

Table 2.2 Summary of statistical analyses results of different failure modelling approaches.

Assumption	Distribution	Estimation method	Parameter value	MTTF	SD*
Perfect repair	Exponential	MLE	λ= 0.0246	40.666	1.6538e+03
Perfect repair	Exponential	HBM	λ= 0.0243	41.152	1.6935e+03
Minimal repair	Weibull	MLE	α= 1.106 β= 52.83	50.872	2.1179e+03
Minimal repair	Weibull	HBM	α= 1.107 β= 57.71	55.569	2.5270e+03

*Standard Deviation

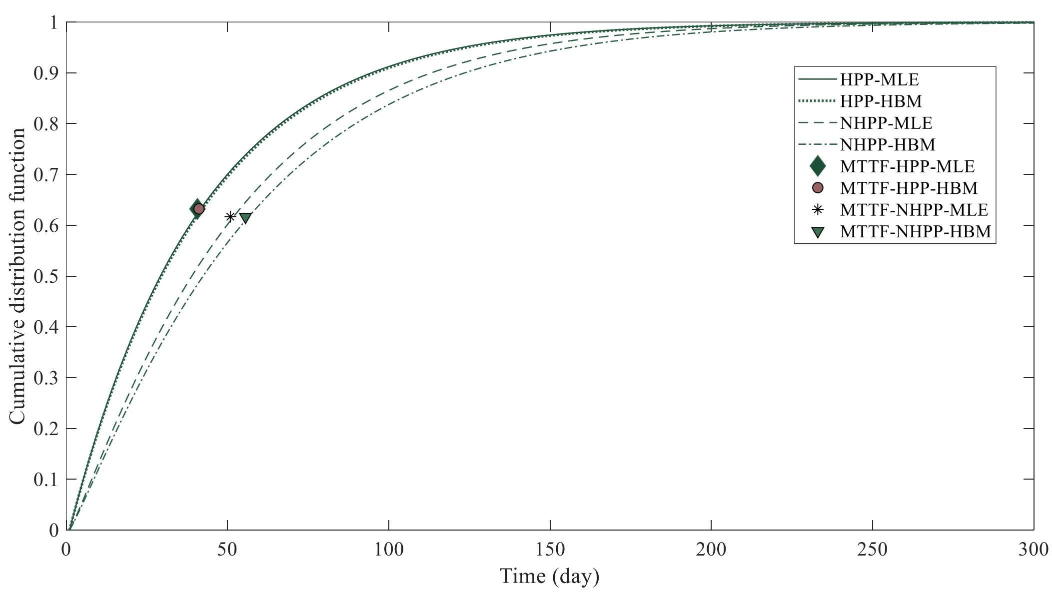


Figure 2.10 Cumulative distribution function and corresponding MTTF values for different repair categories estimated by MLE and HBM methods.

2.10.2 Discussion: The unbiased and minimal variance category of failure modelling approaches

The results listed in Table 2.2 reveal that the value of the Weibull shape parameter, estimated using both parameter estimation approaches (HBM, MLE), are higher than 1, confirming that the number of failure events are dependent upon time. This is in contrary to the PR assumption, where the failure rate is assumed to be constant with time. That is, a MR assumption is appearing to be more credible for the failure modelling of NGRMS. In order to categorized the efficient approach among the presented models given MR assumption, the probability plot for Weibull distribution was developed, as depicted in Figure 2.11. As shown in the figure, the ML of shape and scale parameters of Weibull distribution include higher uncertainty than their estimation through the HBM approach. Therefore, the HB model given an MR assumption seems to be the most reliable approach amongst the reviewed methods. That is, the Bayesian method can efficiently take advantage of the available data to predict the parameters of failure model hence providing an opportunity for improvements of asset management plans.

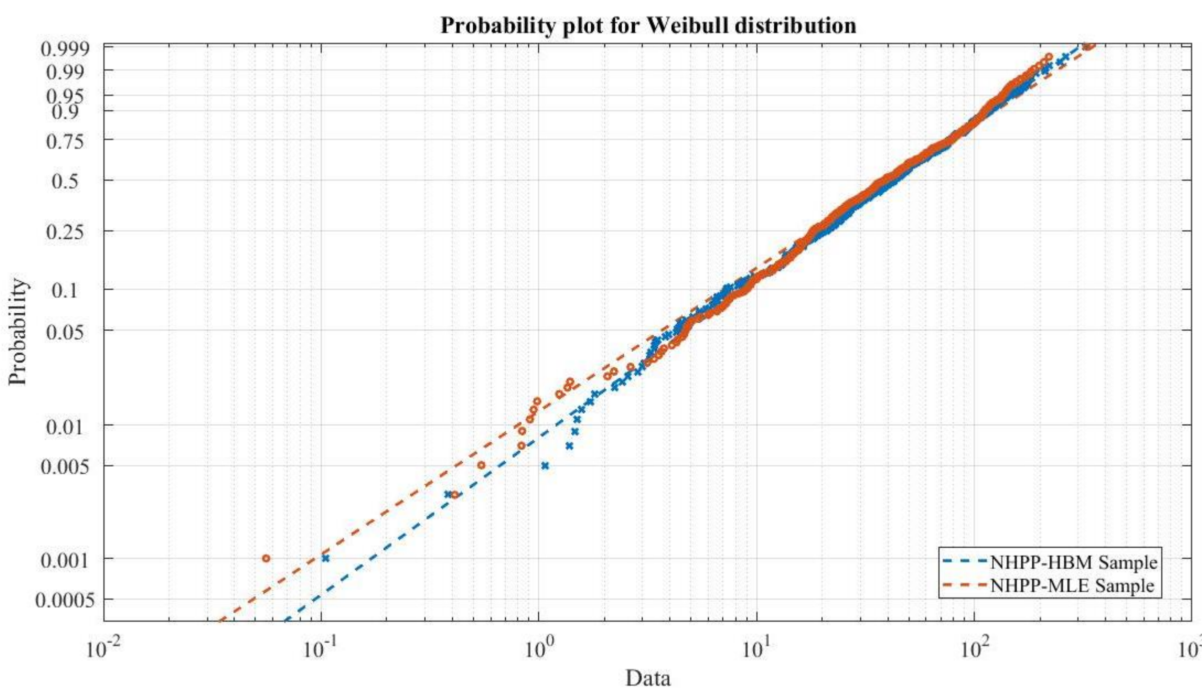


Figure 2.11 Weibull probability plot for time-dependent failure modelling approaches.

2.11 Conclusion

A major challenge in failure modelling of repairable systems is choosing applicable tools and making valid assumptions. This will also help in reducing the uncertainty associated with the obtained results. The differences between application of two mostly utilized assumptions in failure modelling, MR and PR, have been addressed in this paper. This was carried out in a case study of natural gas regulation and measurement plant by MLE and Bayesian inference method. The final results highlighted that relaxing the renewal process assumption (constant failure rate) and taking the time dependency between the observed failure times into account, results in a more precision of failure modelling where the shape parameter value of Weibull distribution in both parameter estimation approaches (HBM, MLE) are higher than 1, confirming that the number of failure events are dependent upon time. On the other hand, HBM is able to model the correlation between the failure data through an MCMC simulation, leading to less uncertainty in MTTF calculations. This is approved through the developed probability plot for Weibull distribution where the estimated shape and scale parameters of HB model has better precisions than ML estimation. The results also suggest that a minimal repair assumption for an HBM failure analysis estimates longer MTTF which avoid the conduct of premature maintenance or compromise operational safety. As a further investigation, it is recommended to model generalized perfect repair assumption with hierarchical Bayesian inference.

3. A Condition Monitoring Based Signal Filtering Approach for Dynamic Time Dependent Safety Assessment of Natural Gas Distribution Process

Abstract

Condition monitoring of natural gas distribution networks is a fundamental prerequisite for evaluating safety of the operation during the lifetime of the system. Due to the high level of uncertainty in the observed data, predicting the operational reliability of the networks is complicated. Moreover, there is a fluctuation in most of the monitoring data in different time scales, as most of the derived data tend to be of non-stationary nature and are complex to model or forecast. Therefore, a more realistic data driven approach for developing a reliability framework needs to be considered. This paper aims at proposing a probabilistic model to predict the complexity of the non-stationary behaviour in monitoring data. It also aims at developing a novel framework for the time dependent reliability assessment of a natural gas distribution system using condition-monitoring data. To this end a methodology by integrating Empirical Mode Decomposition (EMD) and Hierarchical Bayesian Model (HBM) is developed. The advantages of the methodology are demonstrated through a case study of a Natural Gas Regulating and Metering Station operating in Italy. Based on pressure data acquired from the case study, the model is able to predict overpressure thus directly avoiding unnecessary maintenance and safety consequences.

Keywords: Condition monitoring, Time dependency assumption, Empirical Mode Decomposition (EMD), Hierarchical Bayesian Model (HBM), Noise

3.1 Introduction

Natural gas operational facilities and distribution networks are associated with potential hazards, which pose a threat not only to the workers, but also to the people living around the facilities. Natural gas is considered as the cleanest burning fossil fuel which supplies more than 20% of energy consumption to the European Union (Montiel et al., 1996). A great deal of ongoing effort is made to increase the operational reliability of the natural gas networks by considering real-time condition monitoring of their subsystems. For conducting realistic real time monitoring of the networks, identifying the convenient Process Variables (PVs) is necessary to obtain a safe operation.

Different damages such as cumulative damage, fault damage zone, etc. occur gradually or suddenly (Deloux et al., 2009). Sudden damage is defined as an undetected steady deterioration trend in the system. Therefore, owing to the rise in the components' wear, regular observation of conditions must take place to detect the gradual deterioration (Fouladirad et al., 2008) and achieve suitable predictive maintenance decisions according to the deterioration trend (Yam et al., 2001). Considering the previous researches on Condition Monitoring (CM) in engineering systems, different statistical approaches have been developed to investigate the performance of systems under the associated uncertainties (Collacott 2012, Hameed et al., 2009, Liu et al., 2013, Chetouani 2014, Nandi et al., 2005) The CM maintenance paradigm has received significant development in recent decades, although a longstanding gap continues to exist as there is still lack of a unified model to capture the effect of noise on the raw data, and also the associated uncertainty with time-variant parameters.

The classical models of time series data (Box et al., 2015), such as Auto Regressive Moving Average (ARMA) models, regression methods (e.g. Least Squares Regression (LSR)), and statistical process control (SPC) methods have been widely used by different researchers. For instance, Carden and Brownjohn (Carden et al., 2008) presented a sound physical basis for forming ARMA models of structural response data. Nair et al. (2006) applied ARMA to analytical and experimental outcomes of the American Society of Civil Engineering (ASCE) benchmark structure to detect and locate damaged signals. Pham and Yang (2010) proposed the hybrid model of ARMA and generalized autoregressive conditional heteroscedasticity (GARCH) to predict the machine state based on vibration signal. Kruger and Dimitriadis (2008) developed fault diagnosis scheme to extract the fault signature by applying local Partial Least Square (PLS) model. Wang et al. (2003) presented an application of recursive partial least squares (RPLS) algorithms together with adaptive confidence limits to reduce the number of false alarms. Zhou et al. (2006) integrates the statistical process control technology and the Haar wavelet transform for cycle-based waveform signal to detect a process degradation and to estimate the magnitude of mean shifts.

These approaches focused on evaluating the health condition of considered system over time. However, neither of these models are practical for data with non-stationary and nonlinear nature. ARMA models are applicable to stationary time series data without identifying any long term trends. In the application of statistical regression method, the predicted trend is predetermined since the form of the data should be specified prior to performing a prediction. LSR and SPC are subject to the same drawback and need a pre-specified form that representing the trend. In the case of SPC especially, the level of noise is assumed to be small and data are given to be distributed normally.

The raw data collected from observations in condition monitoring of gas pipelines consists of non-stationary trends, short term cycle and noise. As the noise has a complex time-dependent auto-correlation structure, Empirical Mode Decompositions (EMD) is recommended in recent studies as a suitable statistical method to extract the disturbing noises from the time series. EMD has been successfully applied in different fields, from engineering to climate science and tourism challenges, in order to predict the long-term trend of the observed data with a minimum trace of noise. The model is particularly useful for degradation modeling and for gaining useful knowledge for future decision making strategies (Bin et al., 2012, Lei et al., 2013, Liu et al., 2006, Žvokelj et al., 2010)

Li and Pandey (2017) presented EMD as a statistical algorithm method for condition monitoring, able to isolate the noise and diagnose the ongoing degradation process by recognizing the long-term trend. Although in previous studies, the noise is extracted from the non-stationary data, the long-term trend is considered as the complement of fluctuation trend which can be either mean trend or a constant. This provides uncertainty in the long-term trend, since the correlation and variability of data is not observed over time. Therefore, an appropriate probability model is necessary to consider the time dependency of the data. HBM is a probabilistic tool which incorporates the information on various types of uncertainties over time. HBM is considering the nonlinear nature of the observed data via Markov Chain Monte Carlo (MCMC) sampling. It has been widely used in different fields including probabilistic risk assessment (PRA) and CM (Averill et al., 2018, Kelly et al., 2009, Yang et al., 2013, Weidl et al., 2005, Yu et al., 2017, Abaei et al., 2018a). There is also a great deal of research on PRA and CM applying Bayesian inference (Abaei et al., 2018b, Abbassi et al., 2016, Xin et al., 2017, Straub et al., 2010, Toroody et al., 2016, Arzaghi et al., 2018, Leoni et al., 2018). Likewise, several engineering challenges have been struggled by the extensions of Bayesian approach, i.e., Dynamic Bayesian Network (Abaei et al., 2017, Arzaghi et al., 2017, BahooToroody et al., 2019, Luque et al., 2013).

This paper attempts to develop methodology by integrating the EMD and HBM in a systematic framework for dealing with the noise and the uncertainty associated with variability of monitoring data. Given the non-stationary and nonlinear nature of the captured data, the condition of PVs is monitored to predict the likelihood of the variables exceeding the safe operational limit. To this end, nonhomogeneous Poisson process (NHPP) is adopted to model the number of times that PVs pass the safety threshold. A Natural Gas Regulating and Metering Station (NGRMS) operating in Italy is selected as a case study to indicate the advantages of the developed methodology.

3.2 Empirical Mode Decomposition

The raw data is an amalgamation of true signal and noise. Noise is introduced to the data by either data gathering instruments such as sensors or system conditions due to concurrent phenomena. Considering the time-dependent auto-correlation structure of noise, filtering processes are complex and in some cases impossible (Wu et al., 2004) As an effective filtering method, EMD is introduced based on Hilbert-Huang Transform (Huang et al., 1998) According to EMD method, the time series data is decomposed into a set of functions, known as Intrinsic Mode Functions (IMFs) and a trend known as residual term given by Equation 3.1 (Li et al., 2017).

$$x(t) = \sum_{k=1}^{n} c_k(t) + r(t) \tag{3.1}$$

where $c_k(t)$ is k-th IMF, n as the number of sifted IMFs and $r(t)$ is the residue which indicates a long term trend in the process. An extensive review of the sifting process to decompose a time series, including wide range of applications in condition monitoring, is provided by (Huang et al.,.1998) Each IMF can be either random noise or true signal. Thus, the time series is finally decomposed according to Equation 3.2.

$$x(t) = \sum_{k=1}^{i} c_{T,k}(t) + \sum_{k=1}^{j} c_{N,k}(t) + r(t) \tag{3.2}$$

where $c_{T,k}(t)$ and $c_{N,k}(t)$ are a k-th IMF of true and noise data, respectively and $r(t)$ is the residue similarly. As suggested by Wu et al. (2007) and Li et al. (2017), it is necessary to distinguish and filter out the noise from the raw data by conducting a Statistical Significance Test (SST) on the recorded time series of the process. The idea behind the SST is based on evaluating the energy density and the mean period of determined IMFs (Wu et al., 2004).

3.3 Hierarchical Bayesian Modelling

Observed data, manipulated information, and gathered knowledge are three consecutive steps of making inference throughout a model. Models have two fundamental types; aleatory and deterministic. Aleatory models are uncertain since they are imprecisely known. Herein, HBM as one of the most advanced of Bayesian statistical methods, can be applied using open source MCMC software packages such as OpenBUGS (Lunn et al., 2009) to describe aleatoric uncertainty. Subsequently, the associated uncertainty with variability of the observations existing among the data source is to be properly represented by the resulting posterior distribution (Kelly et al., 2009) given by Equation 3.3.

$$\pi_1(\theta|x) = \frac{f(x|\theta)\pi_0(\theta)}{\int_\theta f(x|\theta)\pi_0(\theta)d\theta} \tag{3.3}$$

where θ is the unknown parameter of interest, $f(x|\theta)$ is the likelihood function, and $\pi_1(\theta|x)$ is the posterior distribution. Hierarchical Bayes utilizes multistage prior distribution for the parameter of interest indicated $\pi_0(\theta)$ (Kelly et al., 2009) as follow:

$$\pi_0(\theta) = \int_{\emptyset} \pi_1(\theta|\varphi)\, \pi_2(\varphi) d\varphi \tag{3.4}$$

where, $\pi_1(\theta|x)$ denotes the first-stage prior as the population variability in θ; φ is a vector of hyper-parameters, e.g., $\varphi = (\alpha, \beta)$, while α and β are the shape and scale parameters of a Weibull distribution respectively. The uncertainty in φ is represented by $\pi_2(\theta)$ as the hyper-prior distribution. An informative prior distribution, $\pi_0(\theta)$, is developed using generic data collected from different sources (numerical simulations, experiments or collected from different industrial sectors) to estimate the posterior distribution (Abaei et al., 2018b)

3.4 Methodology: Time Dependent Reliability Assessment

To predict the existing health condition of the gas distribution, an integrated approach is proposed to conform EMD and HBM in a unified framework. Although EMD is ideally appropriate for analyzing data of non-stationary and nonlinear nature, it still cannot resolve the most complex cases, e.g., nonlinear process in which the noise also has the same time-scale as the signal (Wu et al., 2004). Accordingly, to model the remaining uncertainty in noise extracted time series data, HBM is adopted. The developed methodology eliminates random noise emerging from monitored data and reduces the uncertainty involved in the engineering process. The developed methodology includes two different parts, as presented in Figure 3.1. A detailed description of these parts is provided in the following sections. The developed framework in this study could be used in different engineering contexts to compute the failure rates. It would also be a base for further development on risk-based maintenance scheduling optimization.

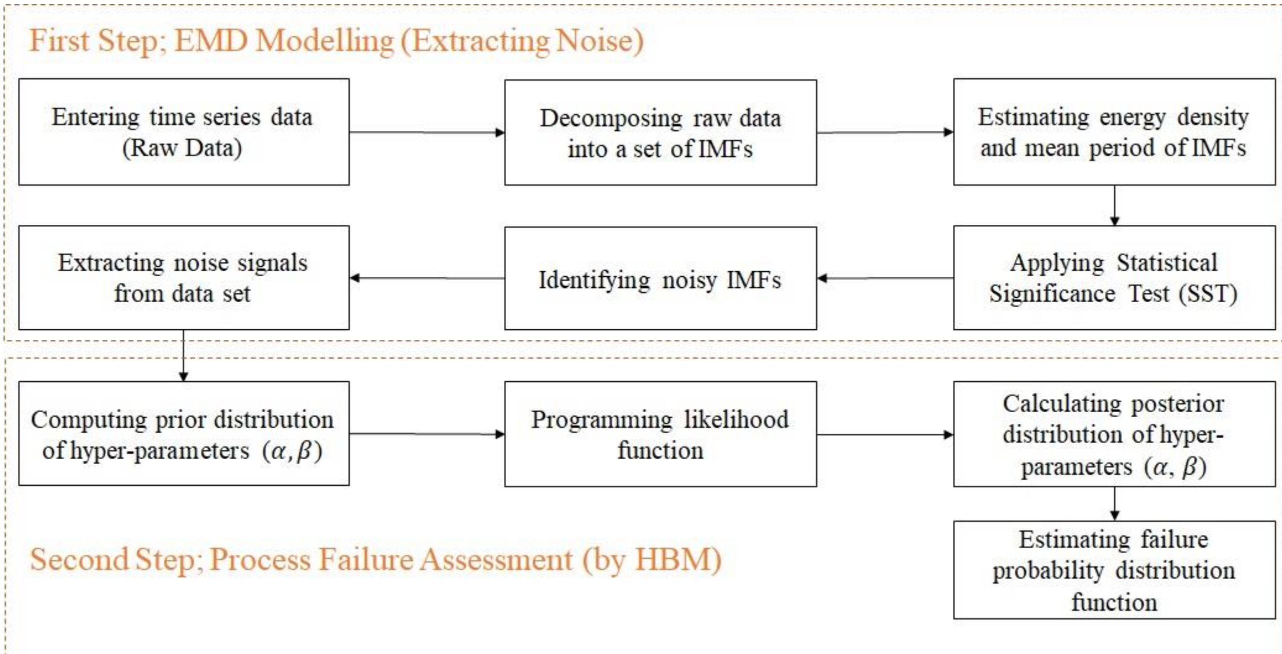


Figure 3.1 Developed methodology for time dependent reliability assessment of gas distribution networks.

3.4.1 EMD Modelling

Appropriate signal processing techniques must be applied for acquiring and processing the raw data to estimate the states of the system. The acquired raw data is generated either from engineering PVs, such as pressure, or environmental conditions such as temperature. The EMD is adopted in developed methodology to extract the random noise from nonstationary and nonlinear raw data. To this end, the considered time series data is decomposed into its IMFs and a long term trend. Considering mean period and energy density of each IMF, the SST will be carried out correspondingly. Mean period of the kth IMF, T_k , is given by Equation 3.5.

$$T_k = \frac{m}{P_k} \tag{3.5}$$

where m is the number of raw data points and P_k indicates number of peaks in the kth IMF, c_k . The general properties of the energy density are considered as function of period for the data (Wu et al., 2004) and is given by Equation 3.6.

$$E_k = \frac{1}{m} \sum_{j}^{m} |c_k(j)|^2 \tag{3.6}$$

where E_k is the energy density of the kth IMF. Similar to the mean period, m denotes the number of data points and $c_1(j), \dots, c_n(j), j = 1, \dots, m$ are n IMFs. In order to identify the noisy IMFs, the SST is applied based on mean and variance of the IMFs which are represented by mean period and energy density, respectively. According to previous conducted studies, there are two substantial beliefs for selecting the first IMF as the main source of the noise in the process (Abaei et al., 2017, Wu et al., 2004, Wu et al., 2007). One belief is that the first IMF has the highest order of fluctuations, and the second one is the mean period (T_1) and energy density (E_1) are not much affected by the sampling uncertainty. Wu et al. (2007) proposed a hypothesis test for any kth IMF in which the Null Hypothesis is that an IMF, $c_k, k = 2, \dots, n$, is a noisy IMF and the test statistics is ($\ln E_k + \ln T_k$). The confidence interval of this hypothesis is defined by Equation 3.7.

$$\ln \left(\frac{1}{3} E_1\right) + \ln T_1 < \ln E_k + \ln T_k < \ln (3 E_1) + \ln T_1 \tag{3.7}$$

Given this SST, noise signals and true signals are those IMFs for which the Null Hypothesis is rejected and accepted, respectively. After identifying and removing the noisy IMFs, the combination of remaining IMFs and the trend function will result in the noise separated signal. This signal will be then used as the input for the second part of the methodology for investigating the exceedance times and frequencies.

3.4.2 Process Failure Assessment

An operational limit should be taken into account to preserve the process in a safe condition during the operation. For this purpose, the times that the operation enters the unsafe zone is recorded considering the noise separated data. These observations are given as the input for predicting the likelihood of safety threshold exceedances. To model the uncertainty in a process, it is more realistic to indicate the correlation of monitoring data in a time series. Unlike a renewal process, that presumes that the inter-arrival times of an observation data are independently and identically distributed (*iid*), Nonhomogeneous Poisson Process (NHPP) is based on the assumption that i^{th} time-step (t_i) is dependent on the value in previous time-step, t_{i-1}. Therefore, in this study, the exceedance rate of safety limit, $\lambda(t)$, in a time series is modeled by NHPP. Consequently, the expected number of exceedances through the specific time interval, $[t_n, t_{n+1}]$ in the process, $E(NE)$, is given by Equation 3.8.

$$E(NE) = \int_{t_n}^{t_{n+1}} \lambda(t)dt \tag{3.8}$$

where an appropriate function for $\lambda(t)$ must be determined to represent the rate of exceedance limit accordingly. To develop an appropriate function for the exceedance rate of safety limit, power-law, log-linear and linear models are recommended by previous researchers (Kelly et al., 2009, Chang 2001). In order to predict the nonlinearity of random process more precisely in comparison with linear modelling, the power-law is taken into account for this study (see Equation 3.9) as suggested by different researchers (Kelly et al., 2009, Abaei et al., 2018a).

$$\lambda(t) = \frac{\alpha}{\beta}(\frac{t}{\beta})^{\alpha-1} \tag{3.9}$$

Constant model (Equation 3.10) and linear model (Equation 3.11) are conjugate to power-law function; considering power-law function, the constant model can be included by $\alpha = 1$.

$$\lambda(t) = \frac{1}{\beta} \tag{3.10}$$

and by $\alpha = 2$, the linear model can be produced which is typically addressed by:

$$\lambda(t) = \alpha + \beta t \tag{3.11}$$

The time to first exceedance of safety limit based on the power-low process generates a Weibull distribution, (t, β, α) , with shape parameter, α, and scale parameter, β, given by Equation 3.12 (Kelly et al., 2009).

$$f(t,\beta,\alpha)=\frac{\alpha}{\beta}(\frac{t_1}{\beta})^{\alpha-1}\exp[-(t_1/\beta)^{\alpha}] \tag{3.12}$$

The observation process will be performed by conducting a failure-truncated approach as recommended by (Kelly et al., 2009). In order to reflect the dependency of times in which the PV exceeds the safety limit, a conditional probability must be defined for each desired time interval $[t_{i-1}, t_i]$ by using Equation 3.13 (Ross et al., 2004).

$$f(t_i|t_{i-1})=f(t_i|T_i{}'>t_{i-1})=\frac{f(t_i)}{\Pr(T_i{}'>t_{i-1})} \tag{3.13}$$

where $T_i{}'$ is the observed exceedances time of the safety limit for a considered PV. Consequently, the truncated Weibull distribution based on the power law function, Equation 3.14 would be achieved.

$$f(t_i|t_{i-1})=\frac{\alpha}{\beta^{\alpha}}(t_i)^{\alpha-1}\exp[-\left(\frac{t_i}{\beta}\right)^{\alpha}+\left(\frac{t_{i-1}}{\beta}\right)^{\alpha}] \tag{3.14}$$

where, $i=\ 2,\ldots,n$, and subsequently the likelihood function is given by using Equation 3.15.

$$f(t_1,t_2,\ldots,t_n|\alpha,\beta)=f(t_1)\prod_{i-2}^{n}f(t_i|t_{i-1}) \tag{3.15}$$

α and β are modelled by HBM to represent the population variability of exceedances time of the safety limits in an operation. It is worth noting that α and β as the hyper-parameters are independent, prior to the observation of the data. Once, an operational variable is observed, these parameters would be dependent. Using Openbugs, the marginal posterior distributions as well as statistics of hyper-parameters are executed by MCMC sampling from their joint distribution. The likelihood function provided by using Equation 3.15 is not pre-programmed into Openbugs. While, as suggested by (Kelly et al., 2009), it is possible to create a vector of n array which is assigned to a generic distribution with parameter, φ. By defining $\varphi=log(likelihood)$ given by Equation 3.16 and considering samples of α and β from the prior distribution in Equation 3.17, Openbugs can update the parameters in the likelihood function (phi) (Abaei et al., 2018a)

$$\varphi = \log(\alpha) - \alpha \times \log(\beta) + (\alpha - 1)\log(t_i) - (t_n/\beta)^{\alpha}/n \quad (3.16)$$

where t_n and t_i are the last and i^{th} observation of the exceedances event in the simulation, respectively, and n is the vector size. Diffusive Gamma distribution is applied independently for the prior distribution of hyper-parameters, α and β, as suggested by (El-Gheriani et al., 2017).

$$\begin{cases} \alpha \sim Gamma(0.0001, 0.0001) \\ \beta \sim Gamma(0.0001, 0.0001) \end{cases} \quad (3.17)$$

MCMC sampling for $i = 1, \dots, n$ leads to estimating the updated posterior distribution of hyper-parameters (α, β).

Based on the resulted Weibull distribution, $f(t, \beta, \alpha)$, the failure probability distribution function for each PV will be predicted. This process could be repeated for each PV to control the operation and enhance the operational reliability of a system.

3.5 Application to a Case Study

Application of the proposed framework is explained by using a practical example of stochastic deterioration process of Natural Gas Regulating and Metering Stations (NGRMS) operating in Italy. Matlab and Openbugs are the available tools applied in this study for execution of the proposed method. The following sections provide a detailed discussion on application of each part of the proposed methodology to the case study.

3.5.1 Scenario development

NGRMS are set up in a distribution system and are fed by transmission pipelines. They are designed in five main sections; inlet, filter, metering, regulator, and outlet. The basic functions of a NGRMS are to reduce the pressure and to measure the gas flow by regulators and metering devices, respectively. NGRMS are designed with redundant parts to ensure that if one part fails, the entire system will not stop. Figure 3.2 illustrates a typical NGRMS scheme.

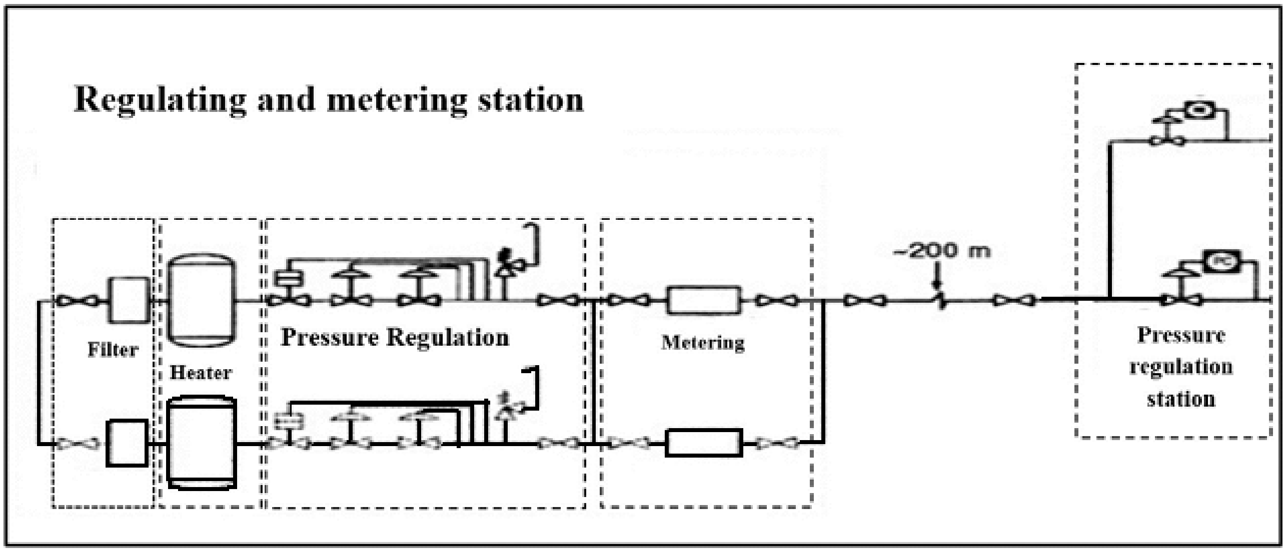


Figure 3.2 Typical plan of Natural Gas Regulating and Metering Station (Gonzalez et al., 2007).

In this study, the pressure is a selected PV to extract the noise and to analyze the potential deterioration process by means of predicting the performance of the network over time.

3.6 EMD modelling of pressure data

The condition monitoring data collected from NGRMS are depicted in Figure 3.3. This historical data consists of pressure values gathered over 108 weeks and would be used as raw data to extract the noise.

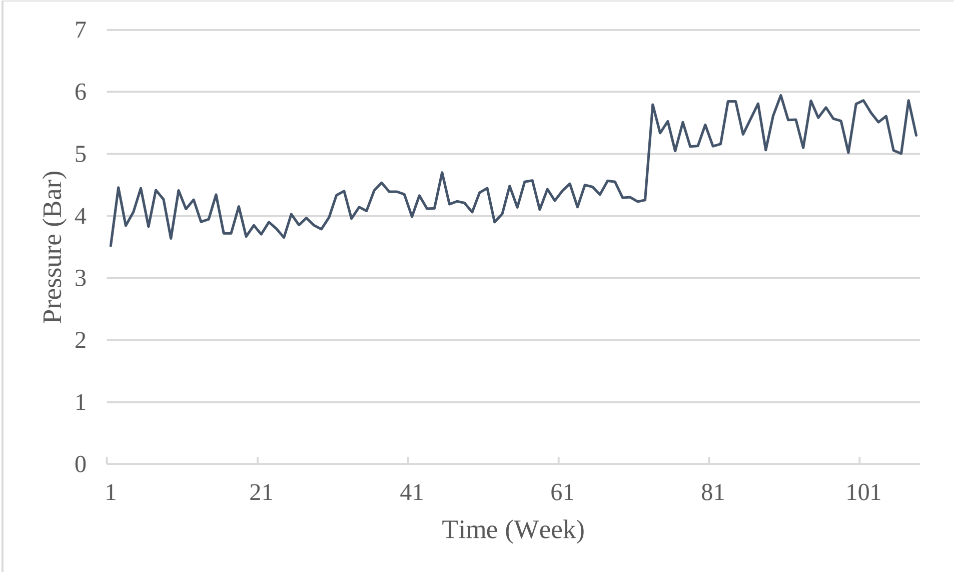


Figure 3.3 Pressure data collected from the NGRMS.

Applying the EMD method on the empirical time series data led to 6 IMFs and a residual trend (see Figure 3.4). In order to discriminate noisy IMFs from actual signals, statistical significance test (SST) was taken based on energy density and mean period. Table 3.1 gives the energy density and the mean period of each IMF. Based on the application of the SST, the first two IMFs, $c_1$and c_2 were recognized as noise signals (see Figure 3.5). So the four remaining IMFs, $c_3 - c_6$, are actual signals.

The long term trend function (see Figure3.4. (g)) illustrates that the pressure values are increasing over time. However, since correlation modelling between data are not considered, this trend is not reliable.

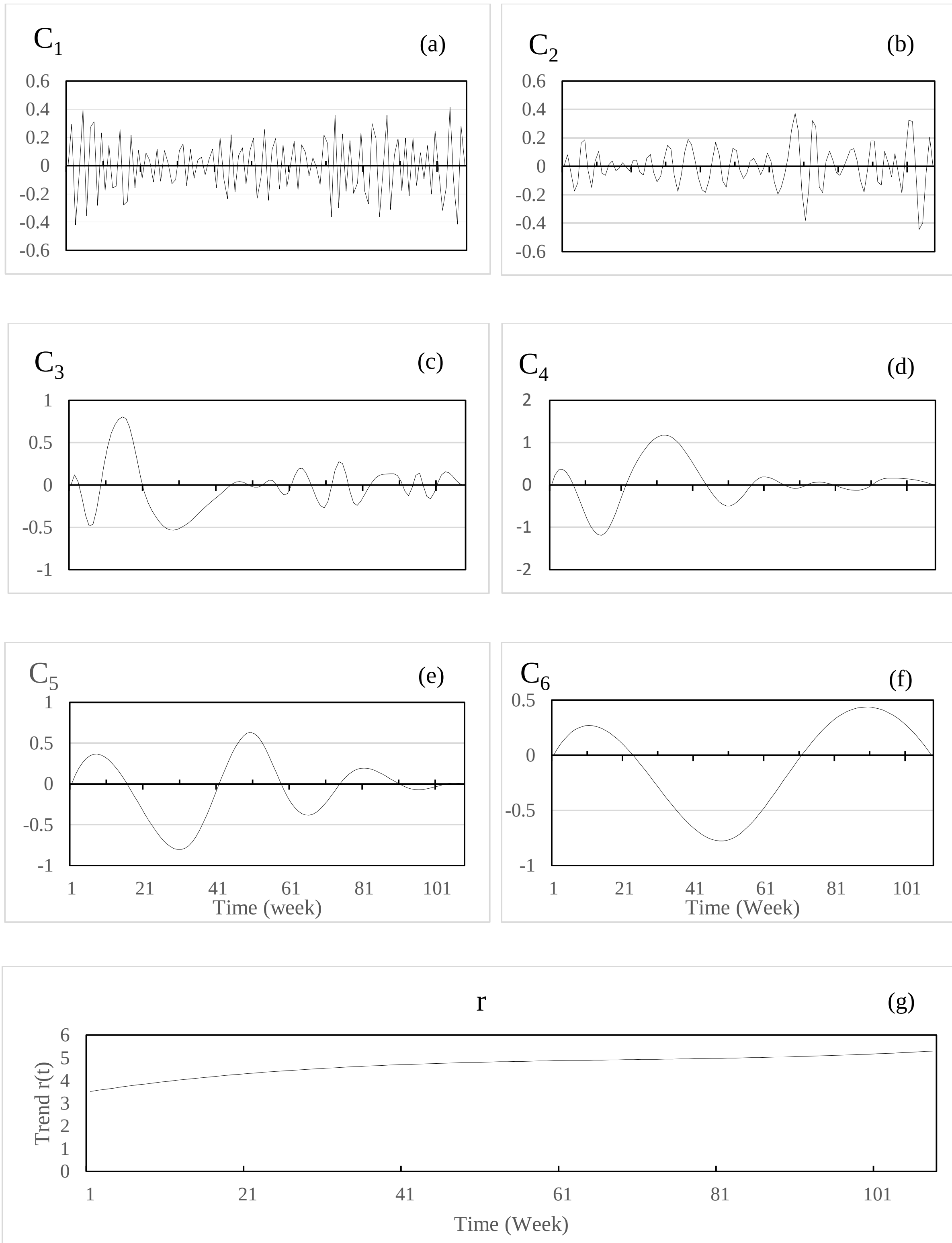


Figure 3.4 Estimated IMFs (C_1, C_2, C_3, C_4, C_5, C_6) and the residue function (r) of pressure in time series.

Table 3.1 Energy density and mean period of IMFs for pressure data.

kth IMF	E_k [(mgKOH/g)^2]	T_k (week)
1	0.0398	2.7
2	0.0203	4.909
3	0.0800	12
4	0.2578	21.6
5	0.1277	27
6	0.1659	54

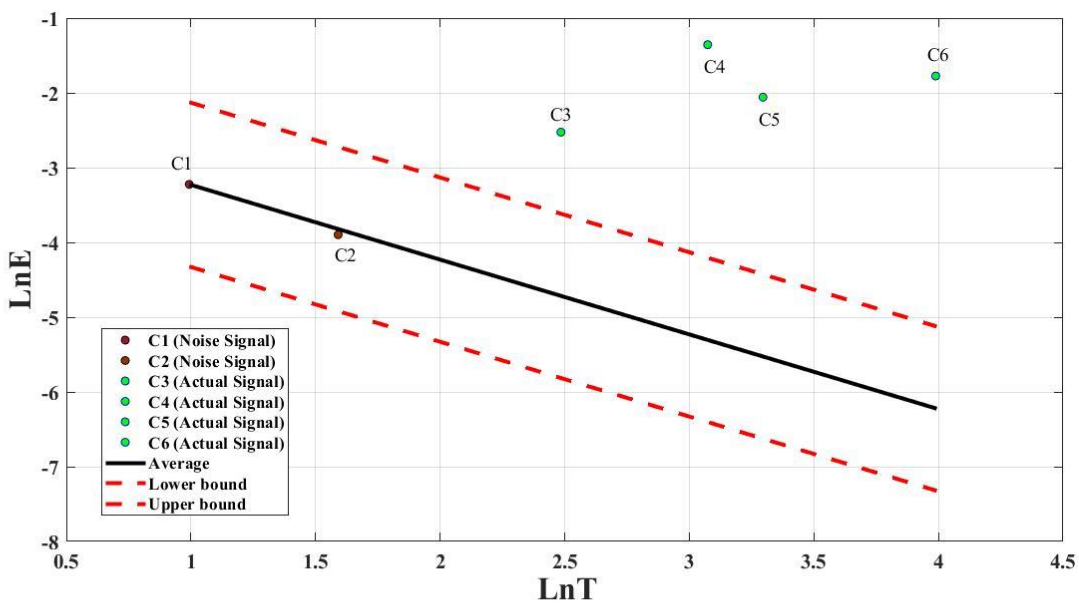


Figure 3.5 SST adopted to the decomposed IMFs of pressure data to identify the noise signals.

The total noise signals extracted from the raw data are depicted in Figure 3.6. A comparison between mean and standard deviations of noise signals and actual signals (presented in Table 3.2) proved that EMD is remarkably effective for extracting the noise. Figure 3.7 shows this comparison by true signal (defined as a summation of noise signals and actual signals) and noise separated signal in time series.

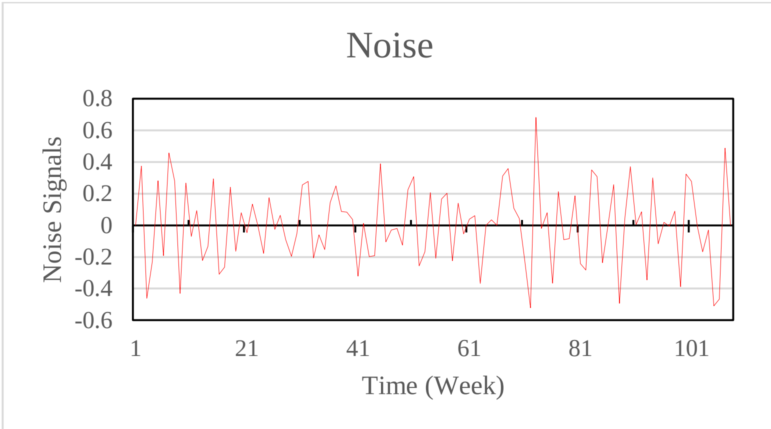


Figure 3.6 The superimposed noisy signals extracted by EMD.

Table 3.2 Statistical Summary of the noise and true signals derived from EMD.

Signal	**Mean**	**Standard deviation**
Noise separated signal	-0.0875	0.4478
True signal	-0.0855	0.5003

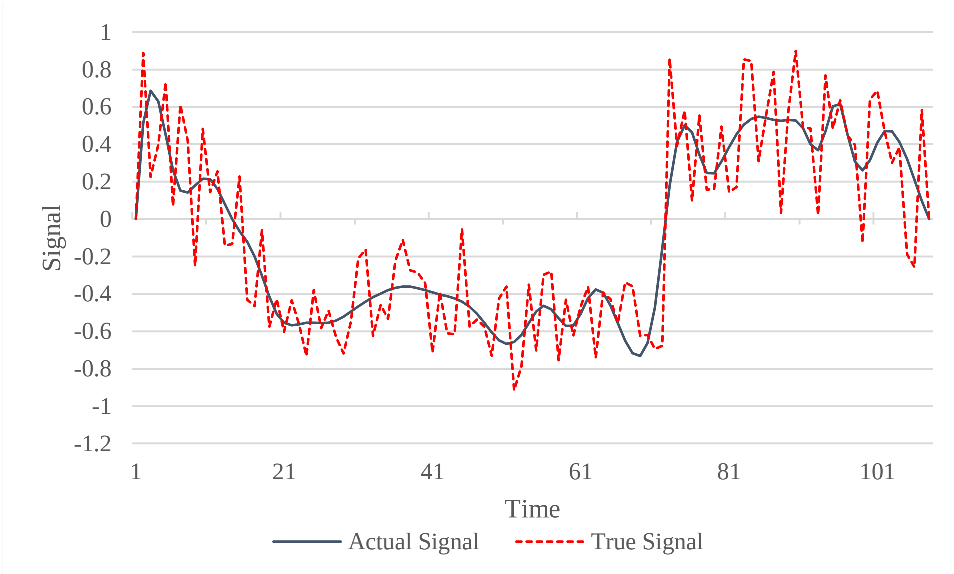


Figure 3.7 Comparison of True Signal (actual signal and noise signal) and actual signal (noise separated signal).

In order to estimate the Noise Separated Signal (NSS), the noisy IMFs were removed from the monitoring data and then the remaining IMFs (C3 to C6) and long term trend function (r) are subsequently superimposed. Figure 3.8 represents the NSS graph along with raw monitoring data and trend function. The NSS is the final filtered signal which is considered as the input for the failure assessment of the process in the second part of the framework.

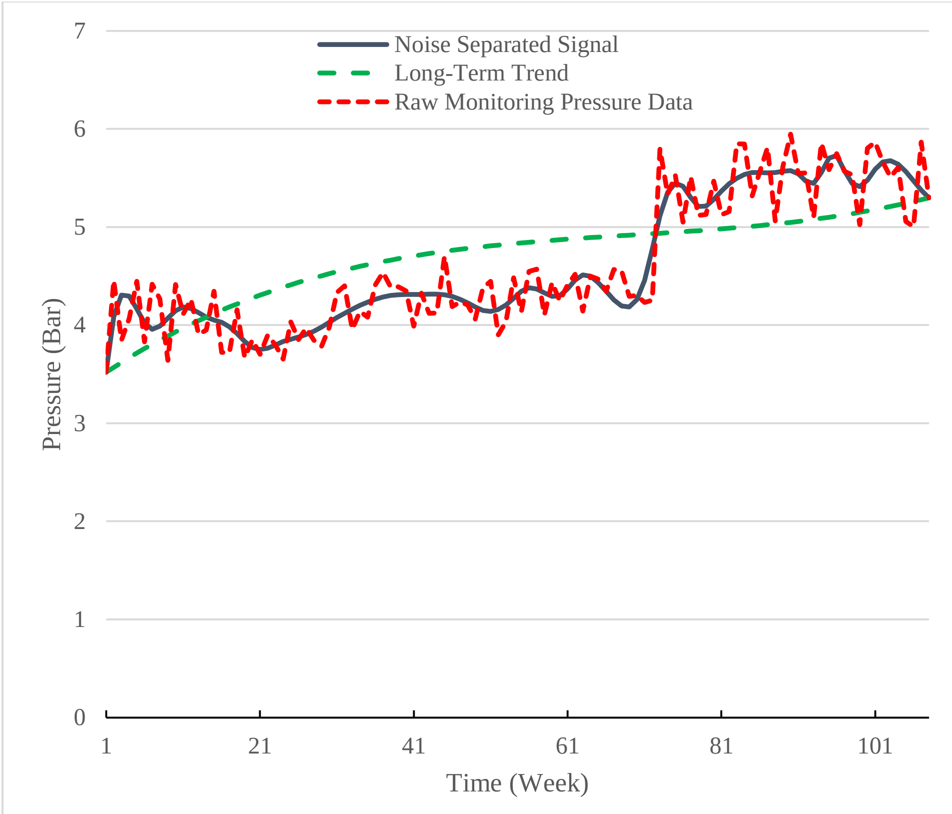


Figure 3.8 Pressure time series data after noise separation.

3.7 Failure assessment: Hierarchical Bayesian Model

In order to evaluate performance of the operation over time, a safety limit was assigned according to the Italian National Gas distribution regulations (UNI 2012). Based on the characteristics of the selected NGRMS, an output pressure bound of [4,5] bar, was determined as the safety limit. Considering the noise separated pressure data, the number of times that the pressure exceeded safety limit were recorded. These observations were then entered into the HBM in order to estimate the likelihood function and calculate the posterior probability of the Weibull parameters, α and β. To this end, three chains were used by the MCMC simulation to check the convergence and to predict the posterior distribution of the parameters (α , β). Each chain started from a separate point with 300,000 iterations, so a total number of 900,000 iterations was established. Figure 3.9 illustrates the iteration history of the shape parameter, α. The results for the estimated posterior probability of α, as well as the correlation between Weibull parameters (α , β) are plotted in

Figure 3.10. It is worth to mention that three colors in Figure 3.9 and Figure 3.10 (b) are representing the treatment of mentioned chains through the MCMC simulation. Furthermore, the summary of estimated marginal posterior distribution for α , β and the expected value for the First Time To Exceed (FTTE) of safe limit are listed in Table 3.3. FTTE is interpreted as the first sign of gradual degradation in the process.

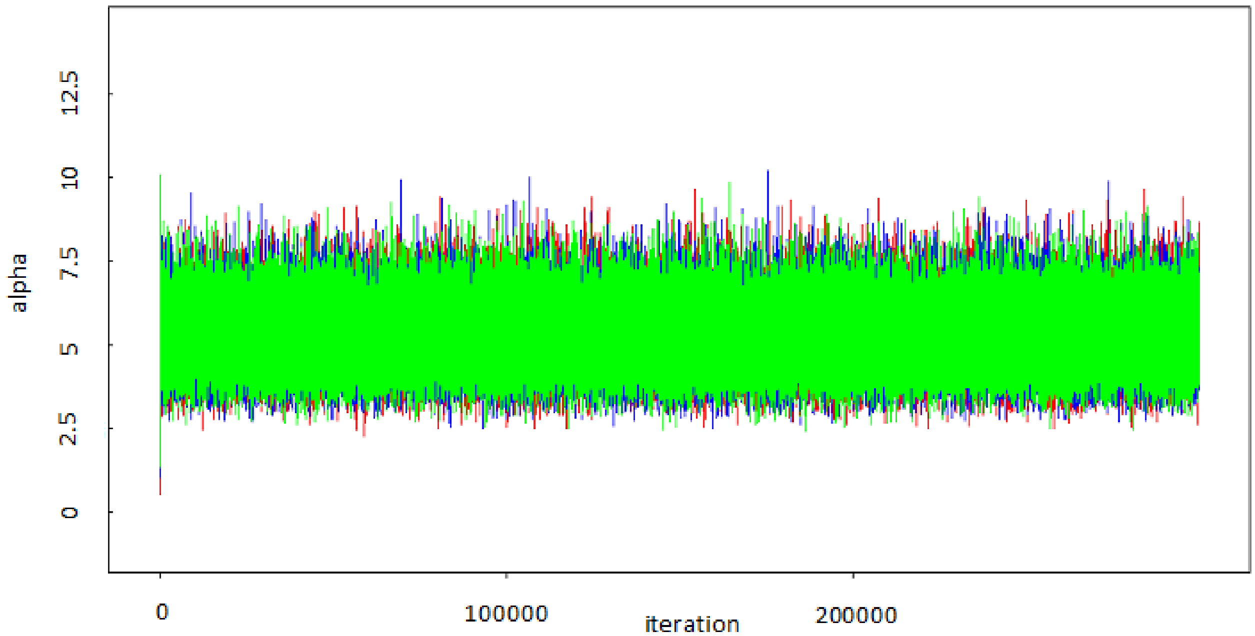


Figure 3.9 Iteration history of three chains for posterior estimation of shape parameter.

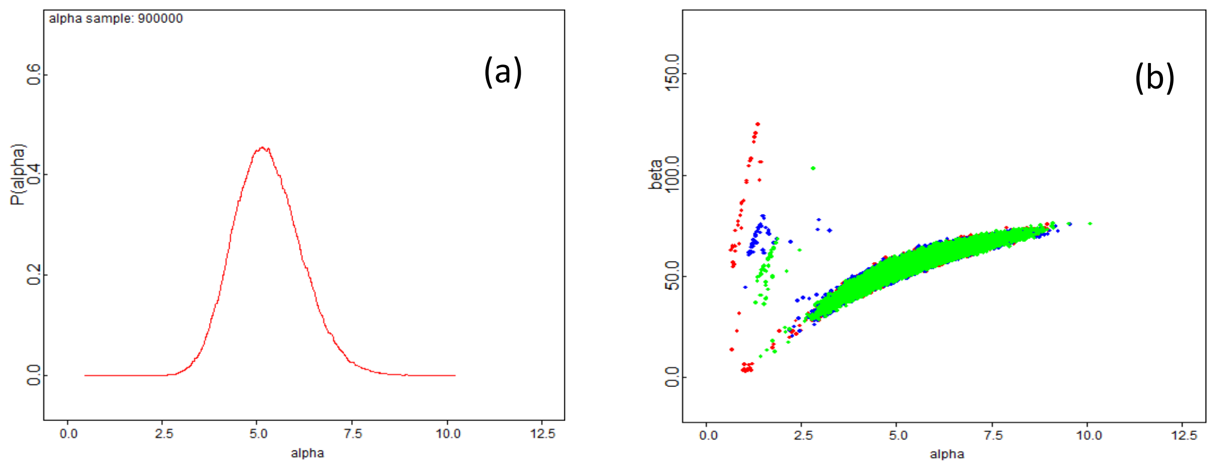


Figure 3.10 Posterior distribution of Weibull shape parameter, α (a) and the correlation between α and β, (b).

The posterior mean value of α is 5.29 with a 95% credible interval of (3.68, 7.17). The shape parameter value is higher than one, inferring the number of times entering into unsafe limits for pressure, are increasing with time. The statistics for the posterior β, were evaluated as a mean value of 54.36 with a 95% credible interval of (40.59, 66.19). Thereby, probability of pressure exceedance of safe limit in NGRMS operation is computed according to the estimated uncertainties; and the cumulative density function (CDF) illustrated in Figure 3.11. The lower and upper percentile of FTTE accounted for were 36.62 % and 61.996 %, respectively. The expected value of FTTE was estimated at 50.07 week which means the first exceedance is expected to occur in the 50th week.

Owing to relaxing the renewal process assumption (constant failure rate) and taking the time dependency of the observed data into account, the proposed framework can model the pressure exceedance from the safety limit more precisely.

Table 3.3 Statistical summary of the Weibull parameters.

	α			β			FTTE		
	mean	2.5 percentile	97.5 percentile	mean	2.5 percentile	97.5 percentile	mean	2.5 percentile	97.5 percentile
HBM	5.293	3.683	7.176	54.36	40.59	66.19	50.075	36.621	61.996

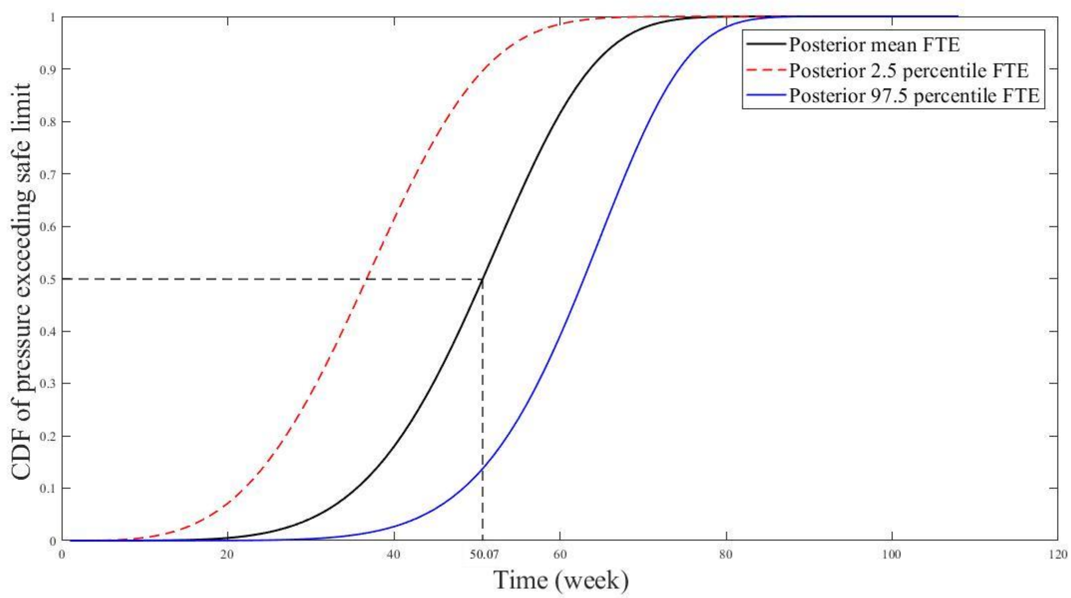


Figure 3.11 CDF of pressure exceedance of the recognized safety limit for the selected NGRMS.

3.8 Conclusion

The uncertainties associated with the deterioration of natural gas distribution networks require a sound condition monitoring methodology for reliability assessment. This paper presented a methodology for time dependent reliability assessment of engineering operations by considering a strategy for noise reduction in monitoring demanding parameters. For this purpose, condition monitoring of an NGRMS subject to degradation was selected to simply explain the application of the developed methodology. Pressure was considered as the PV for observing and modelling the associated uncertainty throughout the process. The considered data had nonlinear and non-stationary nature, so it could not be analyzed by a standard method, e.g., SPC or LSR. Subsequently, in order to remove the noise from the raw data in the observation process, EMD was selected as the statistical tool to filter out the data. During the sifting process, the raw data in the time series were decomposed into a set of IMFs, while the noisy IMFs were identified by conducting the SST approach. Later, a Bayesian predictive tool was employed to model the associated uncertainties influenced on the process over the operational time. The results show that the expected time for exceeding the safe limit is 50 weeks with a credible interval of (38, 64) weeks for the 2.5 and 97.5 percentile of estimated distribution, respectively. The predicted exceedance distributions facilitate the exploration of the onset of deterioration. The developed methodology is capable of being considered as a predictive tool for estimating lifetime condition of an engineering process, and regarded as a platform for future decision making analysis to improve asset integrity management of an industrial operation.

4. A Bayesian Regression Based Condition Monitoring Approach for Effective Reliability Prediction of Random Processes in Autonomous Energy Supply Operation

Abstract

The probabilistic analysis on condition monitoring data has been widely established through the energy supply process to specify the optimum risk remediation program. In such studies, the fluctuations and uncertainties of the operational data including the variability between source of data and the correlation of observations, have to be incorporated if the efficiency is of importance. This study presents a novel probabilistic methodology based on observation data for signifying the impact of risk factors on safety indicators when consideration is given to uncertainty quantification. It provides designers, risk managers and operators a framework for risk mitigation planning within the energy supply processes, whilst also assessing the online reliability. These calculations address the involved and, most of the time, unconsidered risk to make a prediction of safety conditions of the operation in future. To this end, the generalized linear model (GLM) is applied to offer the explanatory model as a regression function for risk factors and safety indicators. Hierarchical Bayesian approach (HBA) is then inferred for the calculations of regression function including interpretation of the intercept and coefficient factors. With Markov Chain Monte Carlo simulation from likelihood function and prior distribution, the HBA is capable of capturing the aforementioned fluctuations and uncertainties in the process of obtaining the

posterior values of the intercept and coefficient factors. To illustrate the capabilities of the developed framework, an autonomous operation of Natural Gas Regulating and Metering Station in Italy has been considered as case study.

Keywords: Condition monitoring, Canonical link function, Bayesian regression, Reliability assessment, Energy supply

4.1 Introduction

Adverse safety consequences of natural gas distribution systems can persuade policymakers, asset managers and stakeholders to apply well established predictive risk management techniques as well as novel probabilistic approaches in parallel with new indicators of safety performance. Within this fact, there are two possibilities for the probabilistic design over components' failure times in a repairable system; first is to rely on the accident and incident reporting system. Next is to count on (near) real time censored operational data. The uncertainty associated with the variations in operational conditions makes the establishment of incident precursor data unreliable. Moreover, since any occurred event in the process is reflected by observation data, the failure analysis based on censored operational data has attracted more attention in process engineering than taking historical data into account (Fast and Palme 2010, Tahan et al. 2017, Artigao et al. 2018, Valdés et al. 2018).

Condition monitoring analysis consists of two key steps; data acquisition (Bechhoefer and Taylor 2011) and data processing (Heng et al. 2009). The process of communicating raw data from operation by means of various sensors such as micro-sensors (Mitchell et al. 1999), ultrasonic sensors (Giurgiutiu et al. 2002), acoustic emission sensors (Loutas et al. 2009), etc. or wireless technologies like Bluetooth (Benghanem 2009), to data storage and handling like enterprise resource planning systems (ERP) (Moore and Starr 2006), computerized maintenance management systems (CMMS) (Kans 2008), etc., is defined as data acquisition. The data processing step is captured within data cleaning (Li and Pandey 2017) and data analysis (Nikula et al. 2016). Three different categories are specified (Jardine et al. 2006) based on the type of collected data for data analysis as: value type (Allgood and Upadhyaya 2000, Sinha 2002), waveform type (Back et al. 2016, Guk et al. 2018) and multi-dimension type (Antonino-Daviu et al. 2017, Zhang et al. 2018).

Different tools have been applied to develop the data analysis in condition monitoring of safety indicators, like autoregressive moving average (ARMA), support vector machine (SVM), principal component analysis (PCA), fast Fourier transform (FFT), wavelet transform, etc. examples include: Baptista et al. (2018) proposed an integration of ARMA and PCA to predict the failure time of the aircraft engine in order to optimize the maintenance time schedule. Miao et al. (2017) developed a one-against-one SVM to model the degraded machine using observation data of the accelerated life time test. Gowid et al. (2015) applied FFT in the acoustic emission field, using operational data to diagnose the fault of high speed centrifugal equipment. Han et al. (2016) established wavelet transform as a part of fault diagnosis of a bearing system to detect informative weak signals buried under random noises.

Across these algorithms the development of random-effect models (REM) and specially Markov Chain Monte Carlo (MCMC) lead to a vast and quick extension of the Bayesian inference applications in different research areas (Kelly and Smith 2009) like econometrics (Kastner 2018), psychology (Wagenmakers et al. 2018), medicine (Elkin et al. 2018), climate and geophysical science (Hermans et al. 2018), engineering (Paltrinieri and Khan 2016), etc. Related literature set sound examples of HBA for engineering process (Abbassi et al. 2016, Toroody et al. 2016, Abaei et al. 2018, Arzaghi et al. 2018, BahooToroody et al. 2019b). Zarei et al. (2018) employed BN to exploit the dynamic feature of this tool in a framework integrated with intuitionistic Fuzzy set theory to model a hybrid dynamic human factor analysis. In a very recent published study, Leoni et al. (2019) applied BN as a heart of the Risk Based Maintenance (RBM) model to estimate the maintenance time for the components of a Natural Gas Regulating and Measuring Station (NGRMS). Choi et al. (2018) established HBA to work out the ground thermal conductivity and borehole thermal resistance parameters given the associated uncertainty in order to design a valid ground-source heat pump. In present study, Bayesian regression approach is employed to analyze the behavior of process variables (PVs) and its impact on their exceedance rate from the safe operational limit throughout the engineering processes. For this purpose, a systematic framework is developed to estimate the reliability of very high values of PVs like pressure, temperature, etc., through the NGRMS process. These calculations address the involved and, most of the time, unconsidered risk to make a prediction of safety conditions of the operation in future. A NGRMS in Florence, Italy is considered to verify the presented model on real data of the pressure values.

The aggregated framework of this study can aid the engineering researchers in estimating the likelihood of safety incidents as one of the most important practical implications in all industries.

4.2 Methodology: hierarchical Bayesian regression

The proposed methodology attempts to specify a dynamic model signifying the impact of risk factors on the treatment of safety indicators, e.g., failure probability function. A variety of mathematical and statistical approaches have been established to define a function with observational data as inputs and failure prediction as outputs (Tinsley and Brown 2000, Roffel and Betlem 2007). The complexity of process, stochastic characteristic of operational data and nonlinear behavior of the industrial process make the specification of function more sophisticated. Employing an appropriate model is still highly important that can assist the researcher not only in minimizing the associated uncertainties but also for a more robust risk analysis and improving the lifetime reliability and availability of the industrial operations. Consequently, in present study, this function is implemented based on Generalized Linear Model (GLM) as a mathematically robust and efficient model (Zeger and Karim 1991, Breslow and Clayton 1993, Guisan et al. 2002) considering both of the variability of non-stationary data and the correlation between nonlinear data with the application of hierarchy levels through the Bayesian inference via MCMC simulation based on Gibbs sampling with open source software packages, i.e., OpenBUGS (Spiegelhalter et al. 2007). The promising framework is introduced in the ensuing sub-sections.

4.3 Canonical link function

A safety indicator function like hazard rate function can be an explicit function of condition monitoring data which leads to relaxing the constant rate assumption. Accordingly, the stochastic behavior of observable quantities is reflected by this function. To this end, the GLM implementation is sketched out incorporating the key steps of this model, as shown in Figure 4.1.

The combination of explanatory variables is characterized by systematic components. An example of the systematic components is:

$$\alpha_0 + \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{r} \alpha_i C_j^k \tag{4.1}$$

where α_0 is the intercept, α_i are the coefficients, C_j are the covariates i.e., process variables, k is an integer exponent, i and j are the number of coefficients and covariates, respectively. Engineering knowledge and historical data are the crucial factors for determining the interactions and associations between independent variables throughout the different types of model from linear to nonlinear.

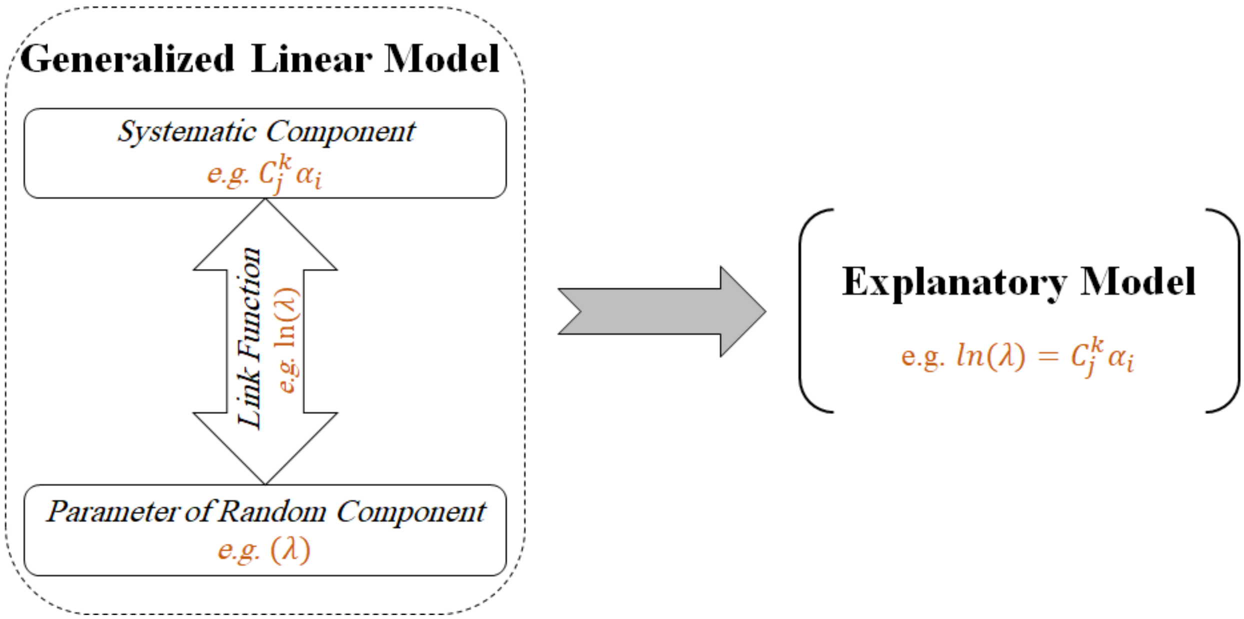


Figure 4.1 Different stages of GLM to define the explanatory model.

Following, the random components, refers to the probability distribution followed by response variable, required to be defined. According to the GLM, the probability distribution of dependent variable is given to be of exponential family. For instance the response variable, R, can follow Poisson distribution, $f(x;\lambda)$, with the parameter of interest λ.

Upon specifying the random and systematic components, the canonical link function can describe the link between the independent variables and the parameter of interest of the response distribution. Consequently, different potential explanatory models such as linear, quadratic (in specific PV), etc., can be established in order to describe the associations between these categorical variables.

Based on the type of random component, systematic component and link function, specific GLM such as multinomial response, loglinear, etc., can be applied to describe the quality of associations between covariates and response variable. As an illustration, given that the response variable follows a Poisson distribution, $f(x;\lambda)$, the explanatory model is :

$$f(R) = \alpha_0 + \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} \alpha_i C_j^k \tag{4.2}$$

where the response variable, R, can be defined as hazard rate. Furthermore, the canonical link function and λ would be achieved by Poisson regression as:

$$\ln(\lambda) = \alpha_0 + \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} \alpha_i C_j^k \tag{4.3}$$

and

$$\lambda = \exp(\alpha_0 + \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} \alpha_i C_j^k) \tag{4.4}$$

where α_0 is the intercept, α_i are the coefficients and C_j are the covariates like process variables.

4.4 Regression tool: hierarchical Bayesian approach

According to the characterized GLM, a great deal of regression tools has been applied, yet associated uncertainty has not been properly investigated. The aim of regression tools is to interpret the intercept and coefficient factors. In this study, for the purpose of incorporating the fluctuation associated with calculation, Bayesian inference with hierarchy levels is employed. Figure 4.2. illustrates the sequence of establishing the hierarchical Bayesian regression using MCMC simulation with sampling hyper-parameters.

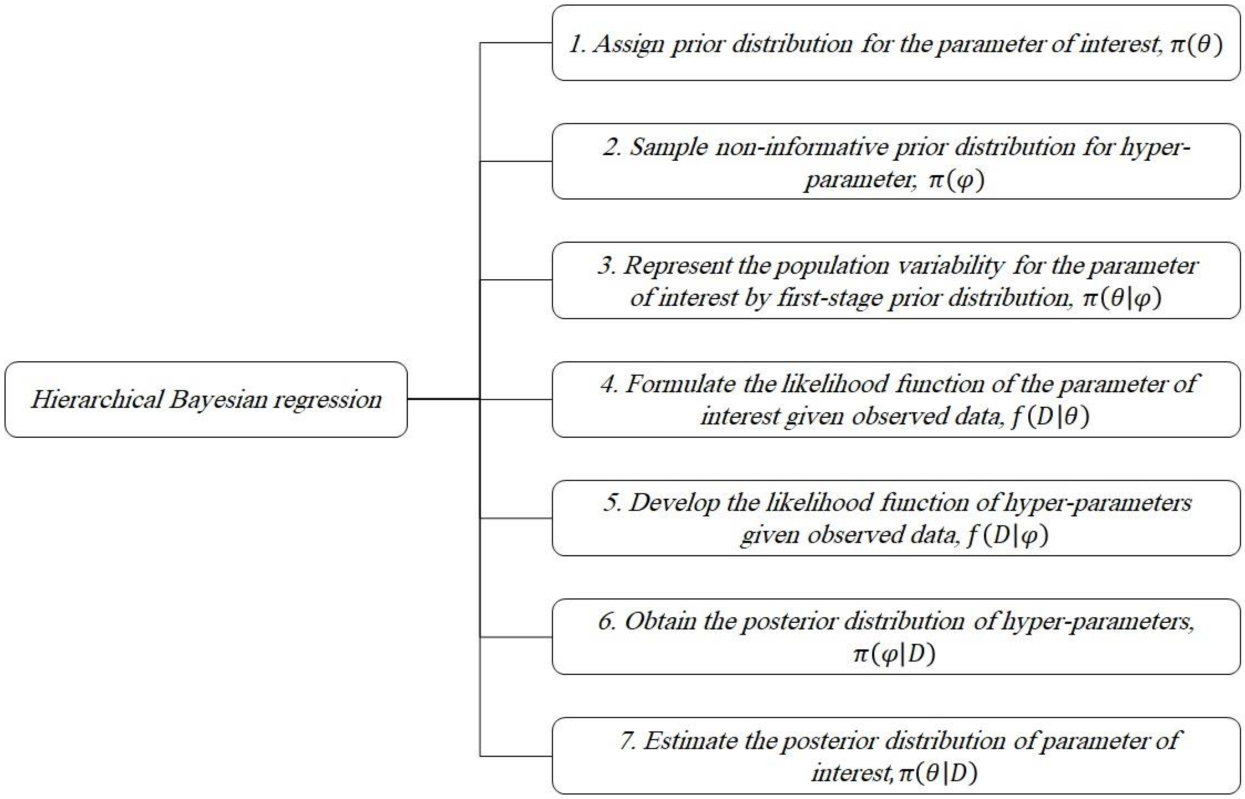


Figure 4.2 The sequence of the hierarchical Bayesian regression procedure.

Adopting the HBA as a regression tool leads to assume that the parameter of interest of dependent variable, θ, is not a certain value, rather follow a distribution with a prior probability of:

$$\pi_0(\theta) = \int_{\emptyset} \pi_1(\theta|\varphi)\, \pi_2(\varphi) d\varphi \tag{4.5}$$

where, $\pi_1(\theta|\varphi)$ is the first-stage prior representing the population variability in θ; φ denotes a vector of hyper-parameters and its uncertainty is represented by $\pi_2(\varphi)$ as the hyper-prior distribution. Through the Bayesian regression approach, the hyper-parameters are represented by the intercept and coefficient factors, α_i. Therefore, the prior distribution of the parameter of interest, $\pi_0(\theta)$, is:

$$\pi_0(\theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \pi_1(\theta|\alpha_0, \alpha_1, \dots, \alpha_i)\, \pi_2(\alpha_0, \alpha_1, \dots, \alpha_i) \partial\alpha_0 \partial\alpha_1 \, \dots \partial\alpha_i \tag{4.6}$$

In HBA, the hyper-parameter is sampled from a distribution given for modelling the uncertainty associated with population variability existing in data source of prior distribution of θ. The hyper-parameter distribution is preferably considered to be of non-informative one because of two main reasons; first, through the MCMC simulation, this distribution is able to generate data within any particular range since there is no preference for data generation. Second and more important; using Bayesian updating, the prior probability distribution does not affect the posterior probability distribution strongly if it is properly sampled by non-informative distribution. As a result, the nature of data including the uncertainties will be reflected by the posterior distribution (Yu et al. 2017). The uniform distribution, Jeffrey's prior, diffuse gamma and diffuse normal distribution are the typical choice of non-informative distribution for hyper-parameter. A diffuse normal prior suggested by (Kelly and Smith 2011) is employed to allow the results to be compared with other parameter estimator approaches such as maximum likelihood estimation. Accordingly, in the present study, an independent diffusive normal prior is applied as the prior distribution of hyper-parameters (described by Equation 6) for developing the HBA script in OpenBUGs software.

$$\varphi \sim N(\mu, \sigma^2) \tag{4.7}$$

In order to utilize the prior distribution of hyper-parameters the observation data of dependent and independent variables must be captured. The first-stage prior distribution, $\pi_1(\theta|\varphi)$, and the likelihood function of the parameter of interest given observed data, $l(D|\theta)$, are accordingly addressed through the GLM by the resulted explanatory model determined as Equation 4.2. Considering the estimation of likelihood function of the parameter of interest given observed data, $f(D|\theta)$, the next step is to obtain the likelihood function of the hyper-parameter given observed data, $f(D|\varphi)$, with respect to the parameter of interest, θ. Again, if the intercept and coefficient factors represent the hyper-parameters, the likelihood function can be expressed as:

$$f(D|\alpha_0, \alpha_1, \dots, \alpha_i) = \int_\theta f(D|\theta)\, \pi(\theta|\alpha_0, \alpha_1, \dots, \alpha_i) \partial\theta \tag{4.8}$$

Based on Bayes's theorem, the posterior probability distribution of hyper-parameters can be obtained as:

$$\pi(\alpha_0, \alpha_1, \dots, \alpha_i | D) = \frac{f(D|\alpha_0, \alpha_1, \dots, \alpha_i)\pi(\alpha_0, \alpha_1, \dots, \alpha_i)}{\iint \dots \int f(D|\alpha_0, \alpha_1, \dots, \alpha_i)\pi(\alpha_0, \alpha_1, \dots, \alpha_i)\, \partial\alpha_0 \partial\alpha_1 \; \dots \partial\alpha_i} \tag{4.9}$$

Since the hyper-parameters have conjugate priors, simulating MCMC to the joint probability of hyper-parameters, $(\alpha_0, \alpha_1, \dots, \alpha_i)$, is an approach to solve the analytically intractable integrals in the dominator. Finally, the posterior predictive distribution specifying the population variability in the parameter of interest, θ, would be achieved by marginalization over hyper-parameters as:

$$\pi(\theta|D) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \pi(\theta|\alpha_0, \alpha_1, \dots, \alpha_i)\, \pi\big((\alpha_0, \alpha_1, \dots, \alpha_i)\big|D\big) \partial\alpha_0 \partial\alpha_1 \; \dots \partial\alpha_i \tag{4.10}$$

Equally, the typical intractable integral in Equation 4.10 can be computed via MCMC simulation based on sampling hyper-parameters through their joint probability distribution.

The derived directed acyclic graphic (DAG) model for the developed hierarchical Bayesian regression is depicted in Figure 4.3, where C_j and R, represented as data, D, are covariates and regressand, respectively. α_i is coefficient and intercept factors, given as hyper-parameter, and finally θ denotes the parameter of interest of dependent variable.

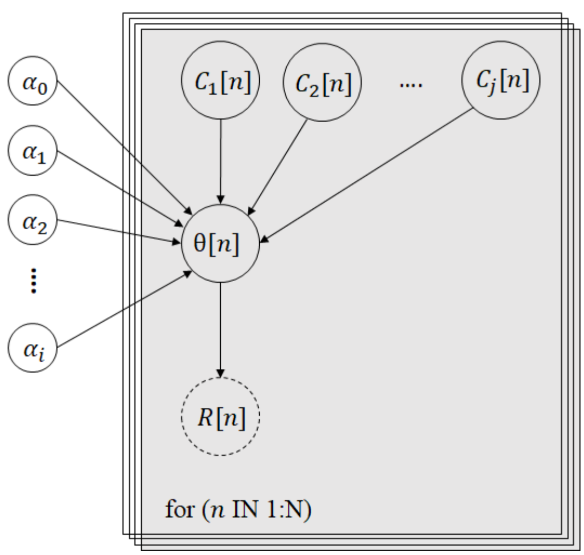


Figure 4.3 Directed acyclic graphic (DAG) model for the developed hierarchical Bayesian regression.

4.5 Application of methodology; Case study

The health condition of an operation can be dynamically demonstrated basing on different risk factors like the stochastic behavior of process variables. In this regard, the presented method is applicable in a wide range of engineering processes, e.g., chemical process, energy supply, etc. In this section, an online probabilistic reliability assessment has been executed on a NGRMS operating in Italy according to the different steps of the developed methodology. This assessment, as a verification of the proposed framework, focuses on adopting the pressure as a process variable and capturing how the *pressure behavior* can affect the *trend of failure function*.

4.5.1 Scenario development

As plotted through schematic plan in Figure 4.4, a standard NGRMS is made up of five main sections: heater, filter, metering, regulator and odorization. It is established with two aims; first, to regulate the Natural Gas outlet pressure to a setting value and, second, to measure the regulated gas flow. In order to reduce the downtime of the operation, one or two redundancy lines are set up. More details i.e., a good introduction on NGRMS operations, is provided in (BahooToroody et al. 2019a, BahooToroody et al. 2019b).

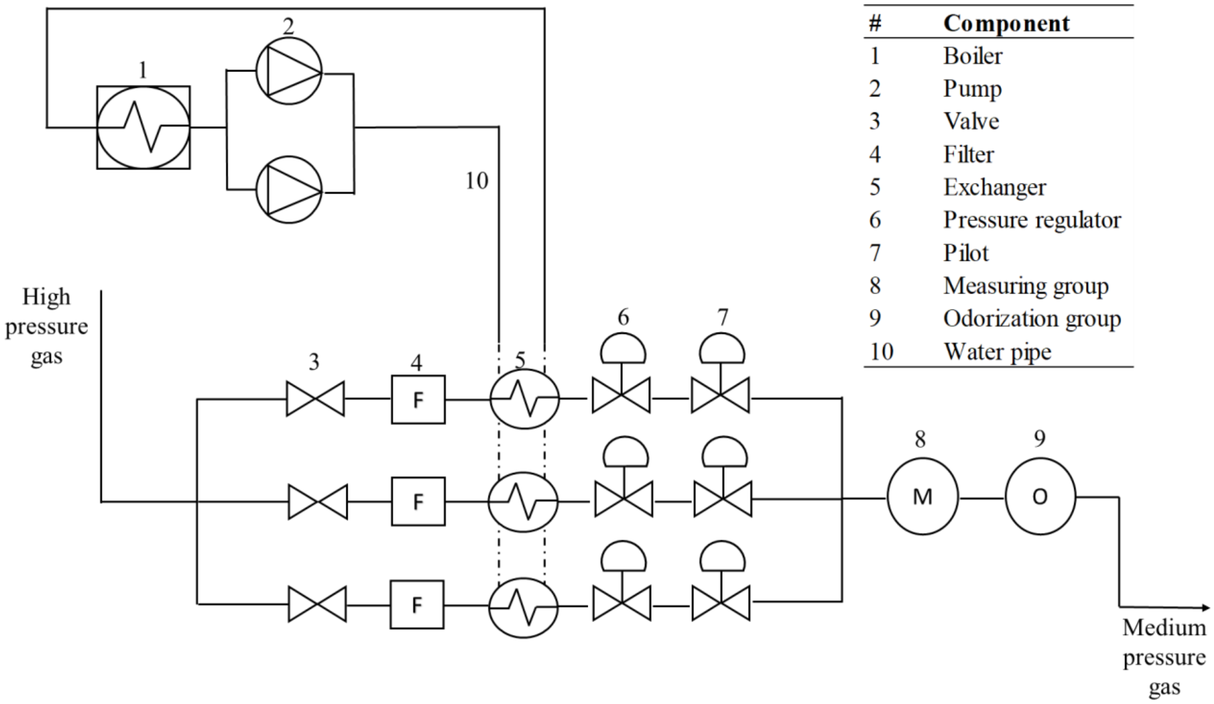


Figure 4.4 Simple system architecture of NGRMS.

Herein, the pressure data collected from the NGRMS operation is placed under scrutiny to predict any potential degradation in the given process. The failure is accounted as any exceedance from the operational safety limit by considered process variable. Figure 4.5 illustrates the time series pressure data of 631 days as well as the safe operational threshold which is set as the pressure of 5 bar. This operational safety limit was assigned based on both the Italian National Gas distribution regulations (UNI 2009, UNI 2009, UNI 2012) and the characteristic of selected NGRMS. It is worthwhile to mention that the random noise in operational pressure data is filtered from nonstationary and nonlinear raw data by adopting Empirical Mode Decomposition (EMD) method to be sure that the reported over-pressure data is true. The explanation of EMD is beyond the scope of this paper and readers are referred to the following researches for detailed discussions on this topic (Huang et al. 1998, Wu and Huang 2004, Wu et al. 2007, Li and Pandey 2017, BahooToroody et al. 2019b).

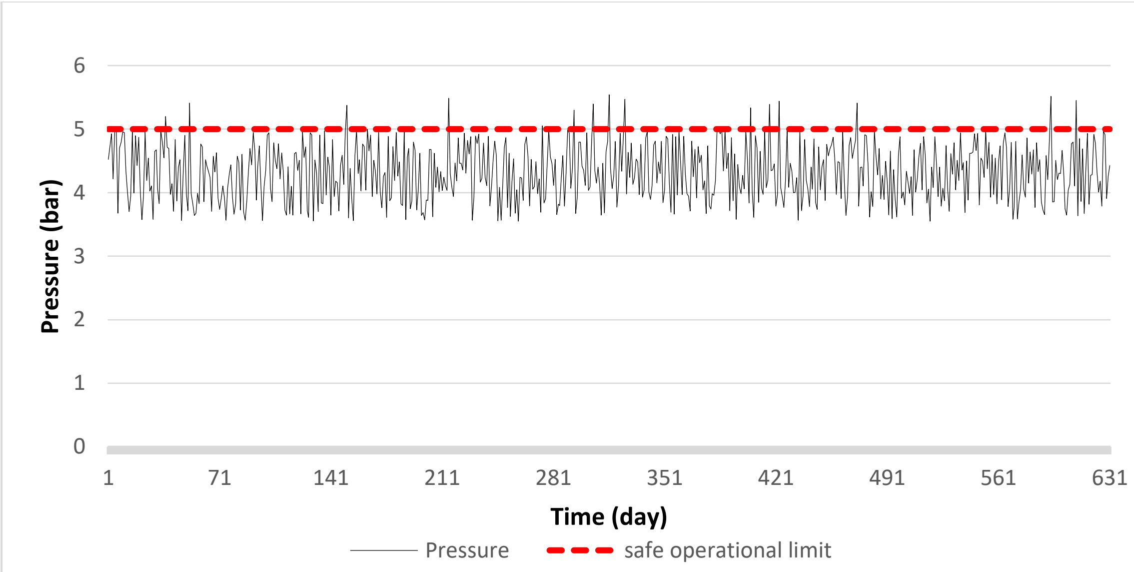


Figure 4.5 Pressure observation data collected from NGRMS.

4.6 Standard explanatory model; normal-binomial model

Upon characterizing the components and random system, an explanatory model can be generated. As this study focused on investigating the impact of pressure trend on the failure function, there is one predictor variable. The resulting component system is given by: $a + bx$, where a and b are intercept and slope coefficient, respectively, and also x denotes the observed pressure data.

The regressand, defined as the exceedance rate, is considered to be the binomial random variable whose the probability of exceedance is π. Subsequently, the exceedances rate variable, E, follows a binomial distribution given by:

$$E \sim B(n, \pi); P(E = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k} \tag{4.11}$$

where n is the number of trials corresponding to the collected pressure records (631 data points). Each of these trials can end in a pressure lower or higher than the assigned safety threshold. Also, k represents the number of exceedances from the safety limit. Accordingly, the canonical link function is the logit function, $logit(\pi)$, expressed as:

$$logit(\pi) = \ln\left(\frac{\pi}{1 - \pi}\right) \tag{4.12}$$

where π can be obtained as:

$$\pi = \frac{e^{(a+bx)}}{1 + e^{(a+bx)}} \tag{4.13}$$

Consequently, the standard explanatory model would be achieved by:

$$logit(\pi) = a + bx \tag{4.14}$$

4.7 Hierarchical Bayesian regression; sampling the coefficients

The resulted explanatory model is applied in the hierarchical Bayesian inference paradigm to predict the posterior mean of a, b and π as intercept, coefficient and parameter of interest of the binomial aleatory model, respectively.

Based on aforementioned steps in section 4.4, an independent diffusive normal prior is employed as the prior distribution of intercept and coefficient. Then, the number of times that the pressure exceeded safety limit were recorded and entered into HBA in order to work out the likelihood function and predict the posterior distribution of the hyper-parameter, a and b. Three chains with over-dispersed initial value of a and b were used in MCMC sampling to check the convergence and to estimate posterior distribution of hyper-parameters by simulating 1000 burn-in iterations followed by 10^5 iterations through each chain. Figure 6 shows the predicted posterior probability of intercept and coefficient. The confirmation of simulation process is also depicted by the trace plot of hyper-parameters in the same figure. Furthermore, the statistical summary of hyper-parameters, a and b, is listed in Table 4.1.

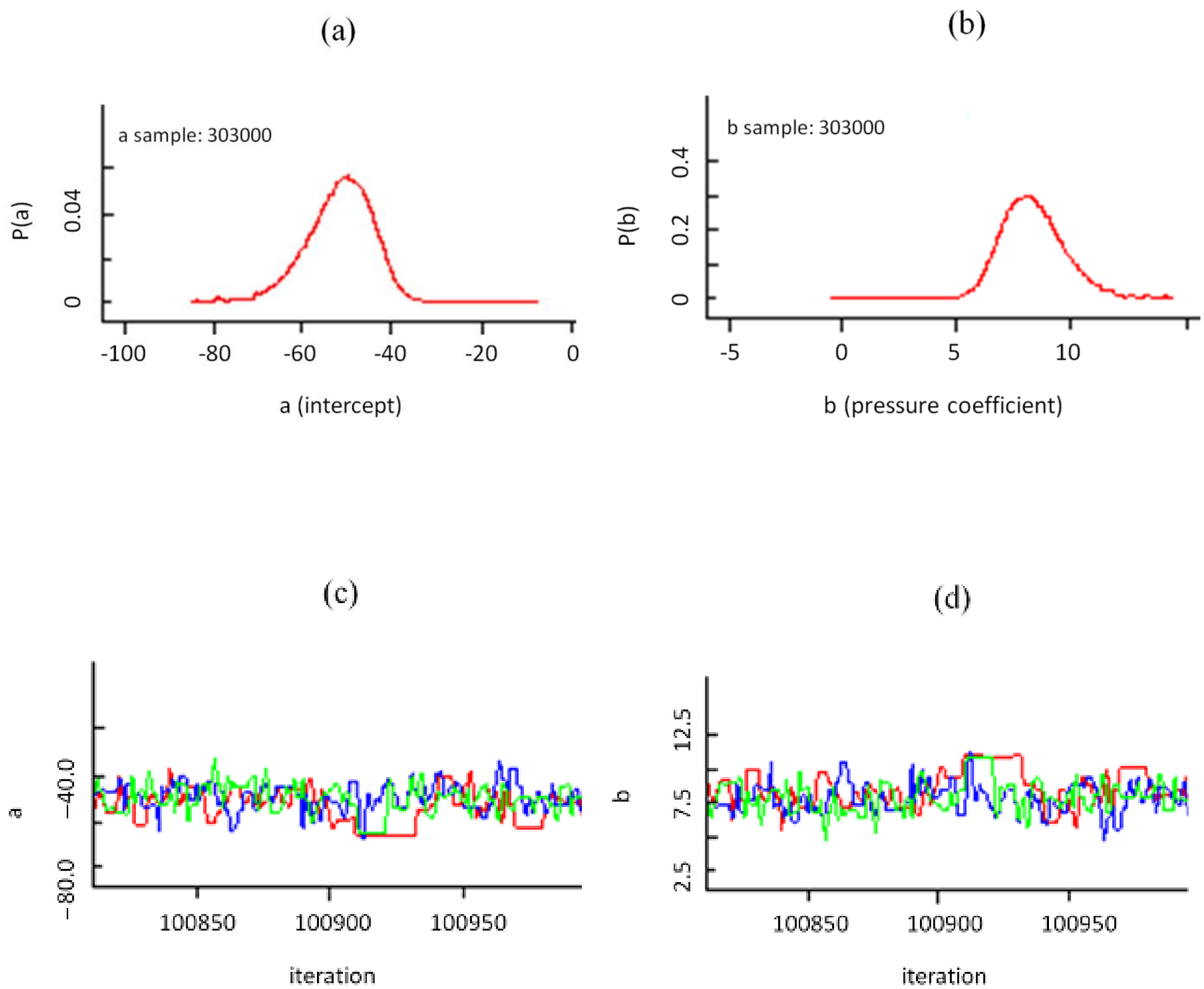


Figure 4.6 Predicted posterior distribution for (a) intercept, (b) pressure coefficient, as well as trace plot for (c) intercept and (d) pressure coefficient.

The validity of predicted model is examined by estimation of Bayesian p-value through the Bayesian chi-square statistic. While as suggested by Kelly and Smith (2009, 2011), the mean of p-value should be near 0.5, the subsequent p-value for the developed HBA script in present study has a mean of 0.7433 proving the validity of the developed Bayesian regression model.

Table 4.1 Statistical summary of predicted hyper-parameters.

	Mean	Standard deviation	2.5 percentile	median	97.5 percentile
a **(intercept)**	-51.9	7.676	-68.95	-51.2	-39.1
b **(pressure coefficient)**	8.343	1.423	5.956	8.22	11.48

The intercept and coefficient factors reported in Table 4.1 and also recorded pressure data are incorporated into the explanatory model to obtain the corresponding logit function, $logit\ (\pi)$, according to Equation 4.14. The resulting standard link function is depicted in Figure 4.7. As it can be viewed in this figure, characterizing the intercept and slop coefficient through a confidence interval with a normal distribution resulted in a normal logit function with a mean of -15.895 and standard deviation of 3.759. (see Figure 4.7).

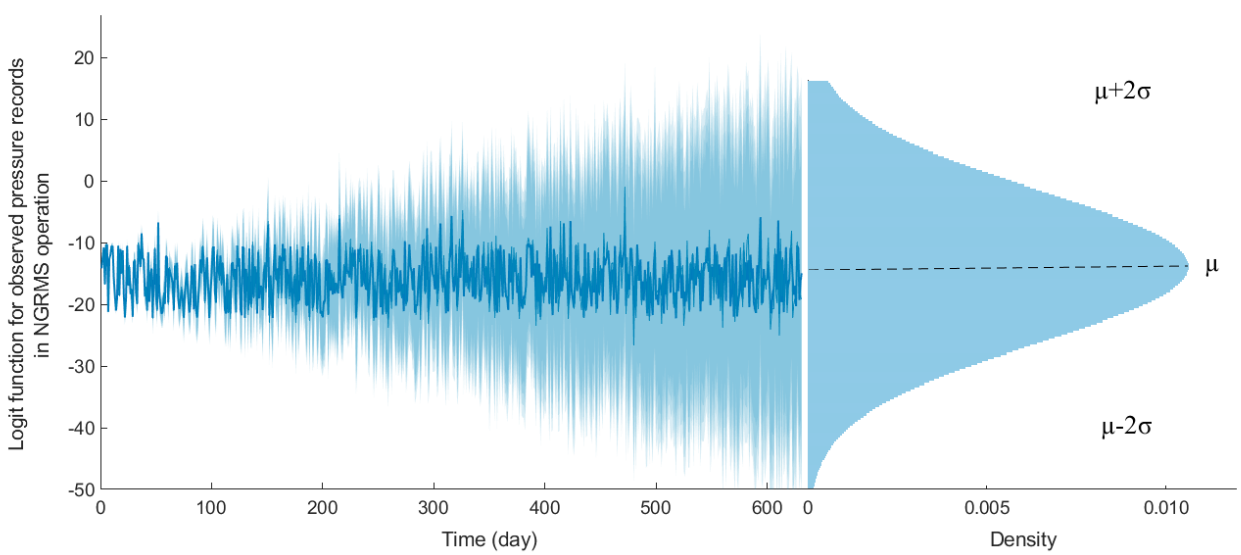


Figure 4.7 Illustration of canonical link function according to the posterior hyper-parameter distribution and pressure condition monitoring data as well as its assigned probability density function.

The obtained probability of exceedance, π, describes the binomial probability density function (PDF) of pressure exceedance from safe operational limit. The estimated PDF is shown in Figure 4.8. The binomial Probability of different number of failures picked up in two exceedances from acceptable risk level, with the probability of slightly higher than 25% (it then declines steadily). One exceedance and three exceedances ranked second and third, respectively, in the binomial probability of failure. Based on the provided data from the considered case study, it is inferred that experiencing higher than 5 failures is almost unlikely referred as rare event. The presented PDF can be correspondingly applied for establishing the risk mitigation strategies.

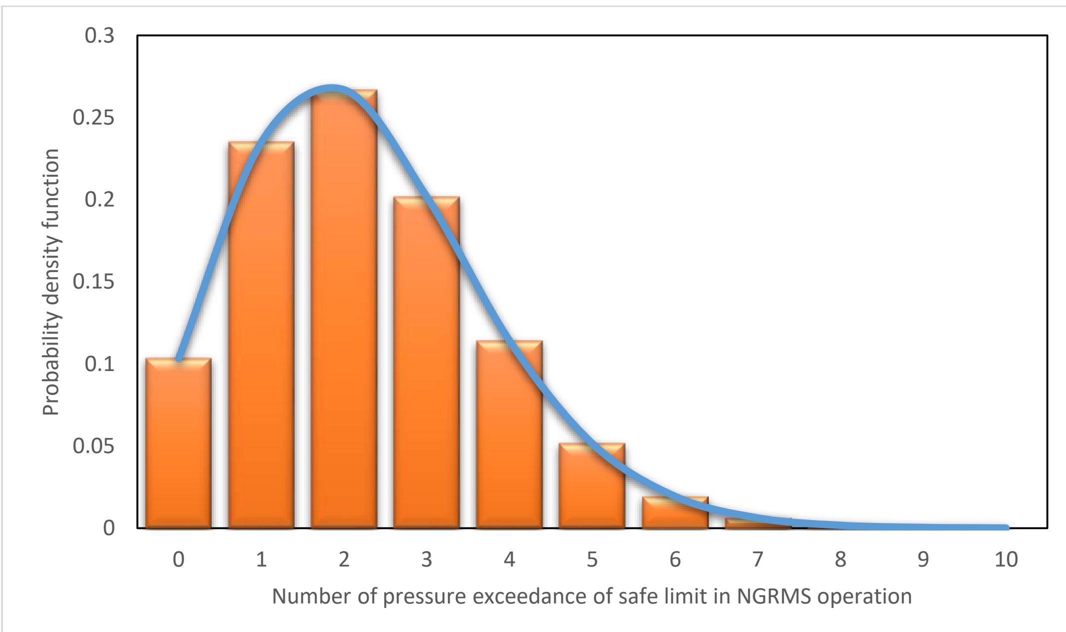


Figure 4.8 Probability density function of pressure exceedances from safe operational limit.

In order to quantify the on-line probability of reliability based on the observation made during the operation (depicted in Figure 4.5), the exceedance probability, π, for each day was obtained (as discussed earlier by Equation 4.13) and shown in Figure 4.9. The results suggest that approximately in a slightly higher percentile than the 95th one of the collected pressure data (617 days), the probability of reliability is almost zero. Furthermore, the probability of having a risk of exceedance for the remaining days (14 days) is in an interval of $[10^{-4}, 3.5\times10^{-3}]$ per day. For instance, the maximum probability of failure appraised at almost 3.5×10^{-3} is expected to be experienced by the 315th working day. The likelihood of considered failure, given by the proposed model may lead to develop decision making approaches aiming at reducing the risk associated with the process.

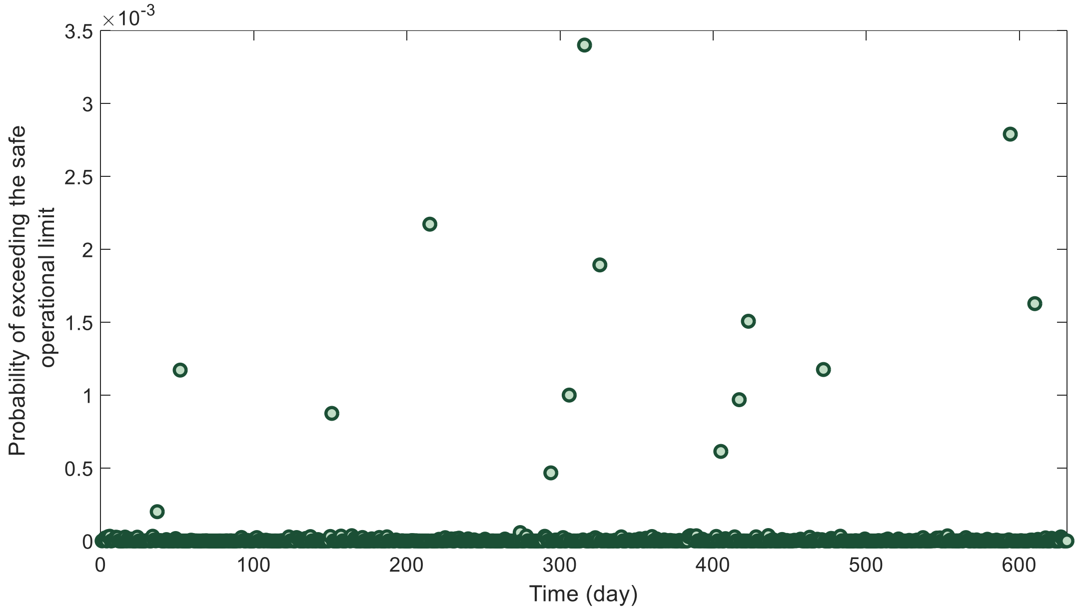


Figure 4.9 Illustration of on-line reliability prediction for the application of NGRMS operation.

Other crucial results, applicable for a risk remediation program, were figured out with inverse-binomial function as stated by Equation 4.15 corresponding with the provided pressure data. Accordingly, the number of exceedances from acceptable risk level given 5th, 50th and 95th percentile probability of failure was predicted to evolve additional data for risk-based decision making.

$$k \sim binoinv(P, n, \pi) \tag{4.15}$$

In this equation, k is the number of exceedances, P represents the probability function of exceedances with parameter of interest π and n is the number of collected pressure data points.

The maximum possible number of failures in the presented case study is estimated at five, as shown in Figure 4.10. It means, if the probability of exceedance from safety threshold is given as 0.95, then the pressure will exceed the threshold 5 times throughout the process, on average once in approximately four months. Conversely, considering a probability of failure of 0.05, no exceedance will be experienced. To describe the failure trend and make figure 4.10 more clear, the time series prediction of the number of exceedances given the probability of failure is 0.5 is illustrated. In this case, two exceedances were predicted; the first exceedance was estimated to be occurred after 193 days while the second and last failure was predicted to be on the 468th day.

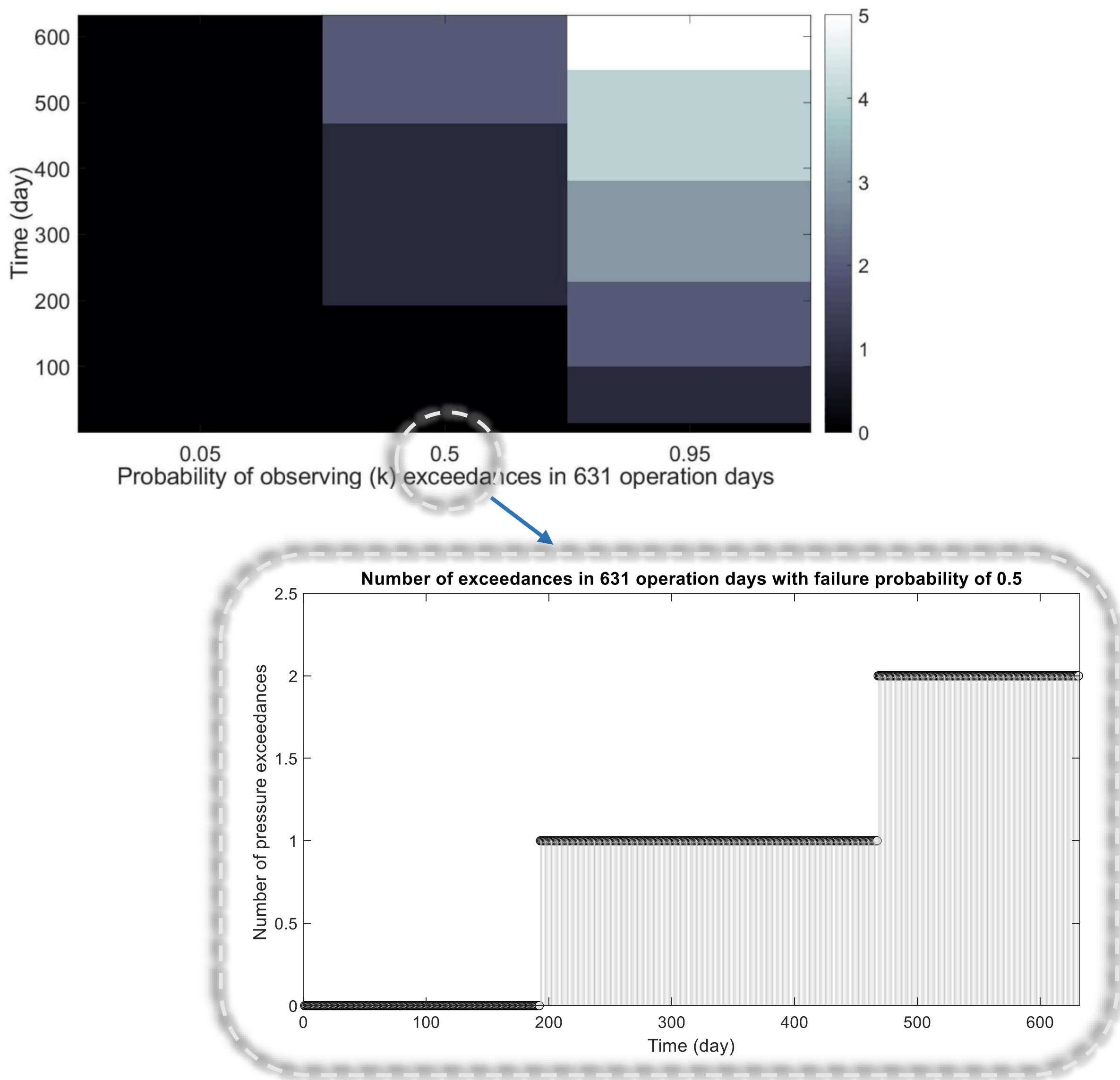


Figure 4.10 Number of exceedances from acceptable risk level based on different probabilities of failure.

4.8 Conclusions

A comprehensive methodology for predicting the trends of safety indicators affected by stochastic risk factors through the supply energy services sector was developed in this study. Given the uncertainty and complication associated with captured operational data, in the developed model, Bayesian inference with hierarchical structure and GLM was integrated to define a regression function. Accordingly, the random component, systematic component and link function were specified to characterize the explanatory model. Herein, the response variable followed a binomial distribution and the resulting link function and explanatory model were logit and logistic regression. Through the Bayesian regression approach, the hyper-parameters were represented by the intercept and coefficient factors, α_i. These factors were then sampled on the basis of operational observations by MCMC simulation to predict their posterior distribution. The obtained regression function is capable for forecasting on-line reliability assessment and consequently developing more efficient remediation plans. As a case study, a stochastic pressure trend of a NGRMS was investigated. Any exceedances from the operational safety threshold have been accounted as a failure. The predictions suggest that in 2% of the operational period, the probability of failure appraised at an interval of [10^{-4}, 3.5×10^{-3}] per day. In order to boost the uncertainty quantification and as a potential future research direction, it is suggested to consider a continuous random component through the GLM.

5. Multi-Level Optimization of Maintenance Plan for Natural Gas System Exposed to Deterioration Process

Abstract

In this paper, a risk-based optimization methodology for a maintenance schedule considering Process Variables (PVs), is developed within the framework of asset integrity assessment. To this end, an integration of Dynamic Bayesian Network, Damage Modelling and sensitivity analysis are implemented to clarify the behaviour of failure probability, considering the exogenous undisciplinable perturbations. Discrete time case is considered through measuring or observing the PVs. Decision configurations and utility nodes are defined inside the network to represent maintenance activities and their associated costs. The regression analysis is considered to model the impact of perturbations on PVs for future applications. The developed methodology is applied to a case study of Chemical Plant (Natural Gas Regulating and Metering Stations). To demonstrate the applicability of the methodology, three cases of seasonal observations of specific PV (pressure) are considered. The proposed methodology could either analyse the failure based on precursor data of PVs or obtain the optimum maintenance schedule based on actual condition of the systems.

Keywords: Risk-based maintenance, Regression tools, Dynamic Bayesian network, Influence diagram, Asset integrity assessment

5.1 Introduction

Since the operational fields of natural gas distribution networks extend far beyond the border of the above ground plant, the safety target community is not limited to the firm’s assets but also includes human life and the environment. Over the past few years, significant attention has been paid by researchers to the inclusion of these aspects in the safety and risk assessment of gas distribution pipelines (Dawotola et al. 2013; De Rademaeker et al. 2014; Mannan 2012; Pasman 2015, Han et al., 2011, Jo et al., 2005). A more recent conducted study by Zarei et al. (2017) developed a comprehensive quantitative dynamic risk assessment framework to alleviate the associated failure with natural gas distribution network. Up to now, many methodologies have been developed to undertake comprehensive risk analysis of an industrial plant. Tixier et al. (2002) identified 62 methodologies divided into three different phases (identification, evaluation and hierarchy). In order to understand their key features and to categorize them into different classes, the paper examines input data, utilized methods and obtained output data.

There is also a great deal of research on asset integrity management and optimization of maintenance plans (Adriaan et al. 2010; Ahmed et al. 2015; Arunraj and Maiti 2007; Azadeh et al. 2015; Khan et al. 2006). This has resulted in many innovative methodologies being developed for asset maintenance in the process industry, where the most common classification of the policies based on the time of application and the geographic location of an asset for single or multi-units, are corrective maintenance (CM), preventive maintenance (PM), predictive maintenance, and proactive maintenance (Barnard 2006; Iqbal et al. 2016; Khan et al. 2004; Moubray 1991).

The last two categories have attracted significant attention from researchers for increasing both effectiveness and efficiency of integrity management (Khan and Haddara 2004). Abbassi et al. (2016) developed a risk-based model to integrate predictive and preventive maintenance strategies in an optimal way. It was concluded that the risk-based methodology developed using Bayesian Network (BN) maintains the desired availability and safety level while minimizing the maintenance cost. Bhandari et al. (2016) proposes a methodology for the design of an optimum maintenance program integrating a dynamic risk-based approach in BN. Their method is based on failure prediction and utilizes precursor information in order to revise the risk profile of the system.

BN as a parametric and non-parametric probabilistic method, has been widely used for risk and reliability assessment of complex engineering systems Barua et al. 2016; Kabir et al. 2015; Yu et al. 2017). Khakzad et al. (2013) demonstrated and compared the application of bow-tie and BN models in conducting quantitative risk analysis of offshore drilling operations. The results of their study show that BN provides more efficient potential than bow-tie models for probabilistic analysis, since it can consider common cause failures and conditional dependencies along with the ability to perform probability updating and sequential learning based on accident precursors data or new available evidence.

Dynamic Bayesian Network (DBN) is a practical extension of static BN whenever an evolving phenomenon must be modelled. In many cases, such as deterioration processes, capturing the dynamic (temporal) behaviour is an important aspect of a modelling process. Daniel Straub (2009) developed a methodology for stochastic modelling of degradation processes. The proposed framework facilitates a robust reliability analysis and Bayesian updating of the model with measurements, monitoring and inspection results. This makes the method highly applicable to near-real time condition monitoring and integrity management.

Another extension to BNs are Influence Diagrams (IDs) which are utilized for developing decision support tools. Conventional graphic-based approaches to decision issues, like Decision Trees, suffer from a number of weaknesses including poor efficiency in representing decision issues with large numbers of parameters and the need for reliable prior information. However, IDs are an alternative which are widely established in engineering applications (Abaei et al. 2017; Arzaghi et al. 2017; Friis-Hansen 2000; Luque and Straub 2013).

Although, a number of researches are conducted for improving the performance of industrial operations using advance probability models, however, little attention has been paid to considering perturbation effects that may be involved in the long-term trend of the industrial process due to its uncertain nature. Therefore, it is essential to have a more in-depth study on estimating the long term trend of the observation data as a unique operational function that can show the performance of the process over the time. This will help to improve the previous studies more accurately capturing fluctuations and the uncertainty of operational parameters. This is a crucial step for a reliable failure modelling of the processes and a better solution platform for decision making problems. Also, BN is applied less for considering the impact of exogenous undisciplined perturbations as one of the important concepts of dynamic reliability. Other tools such as diffusion equations and Monte Carlo simulations etc. are widely used to solve these issues (Gao et al. 2011; Rief 1984; Roos et al. 2008). It should be noted that the present study does not aim at developing a fault detection method. Therefore, there is no specific failure event such as a leakage or crack to be detected by the proposed methodology. However, a mathematical perturbation model is developed based on Fourier series and observation data to predict the lifetime of the process that can assist in monitoring the process using a DBN as an inspection tool.

The present paper focuses on adopting the Process Variables (PVs) and assessing how their variations can be used for determining the optimum maintenance schedule. That is, what temperature or pressure, for instance, can change the failure rate of a component in the system for which a maintenance task may be essential. Among all contributors, the perturbation plays a pivotal role. It is the amount of deviation from expected steady state condition of normal operation. A DBN is established to model the damage and the estimation of failure probability distribution, considering the observed trends in PVs. The DBN is then extended to an ID for decision making regarding the optimum maintenance interval as well as the maintenance type. A risk-based approach is selected for proposing the methodology and to demonstrate its application. Developing a risk-based maintenance policy for a Natural Gas Reduction Station in Florence, Italy is considered.

5.2 Methodology

In this study, a framework for stochastic modelling of dynamic processes using a DBN is developed. To this end, the fundamentals of BNs are discussed first, then the steps of the developed methodology (see Figure 5.1) are discussed in detail in the following sections. The model can be used in different applications for estimating the failure rates based on precursor data and for optimising the maintenance schedules using a risk-based approach.

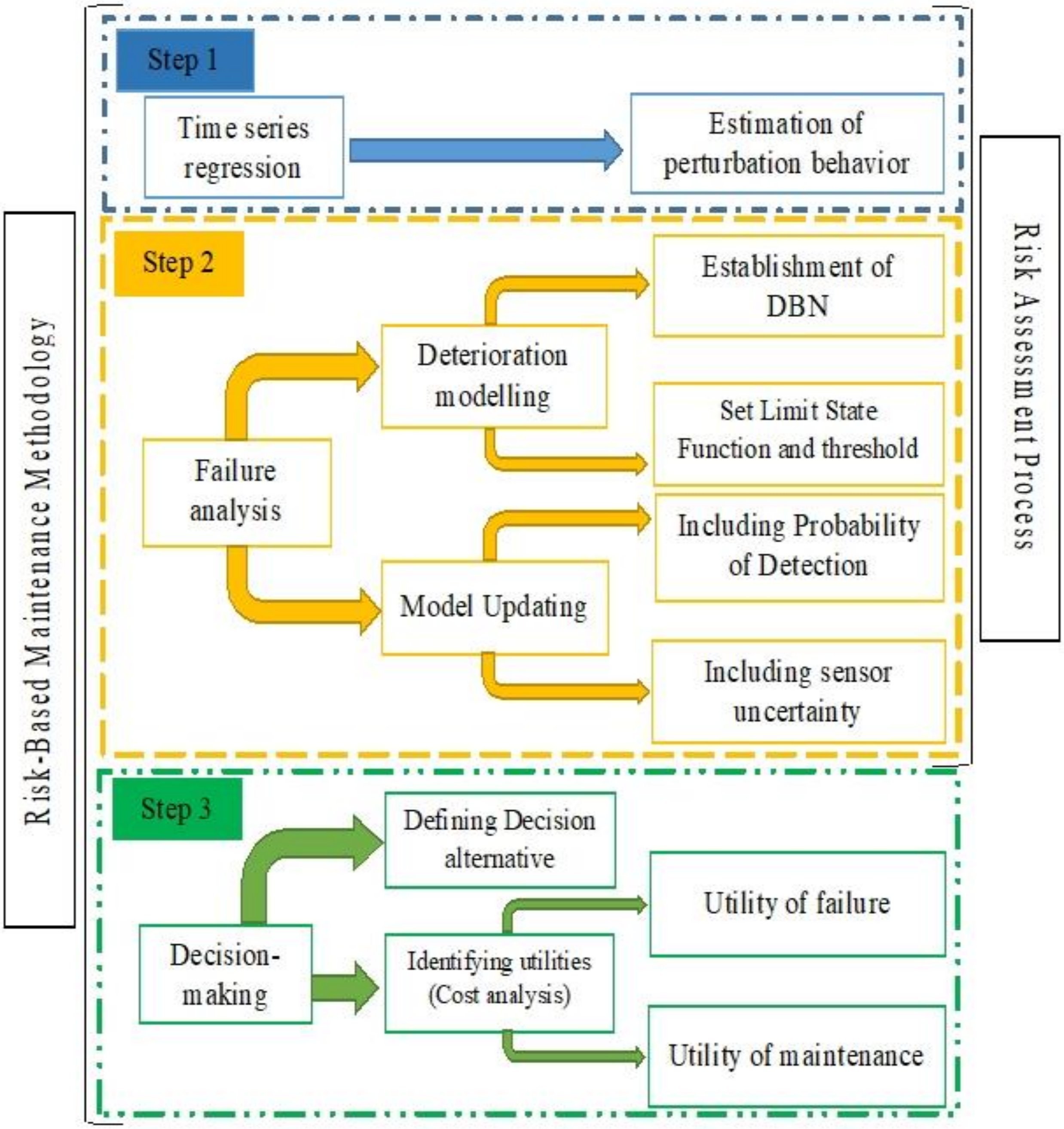


Figure 5.1 Developed methodology for maintenance planning based on PV behavior and impacts of exogenous perturbations.

5.3 Tools

5.3.1 Bayesian Network

A detailed discussion on probabilistic knowledge elicitation using BN with a wide range of engineering applications is presented by (Barber 2012, Neapolitan 2004, Nielsen 2009) BN is a strong tool to incorporate the deterministic data into the probabilistic model with robust connections to graph theory. Based on the capability of including different types of uncertainty (aleatory and epistemic), BN is recognised as a promising method for risk analysis of complex systems. BN is also able to incorporate both causes and consequences of the failure event in a single network.

BN is a Directed Acyclic Graph (DAG) in which the nodes (random variables) are inter-connected with arcs that represent probabilistic dependencies among variables. For instance, Figure 5.2 presents a schematic BN where node X_3 is a child of X_1 and X_2; nodes X_1 and X_2 are considered as parent nodes of X_3. Each node consists of a conditional probability table (CPT). Based on the conditional independencies and the chain rule, BN estimates the joint probability distribution of a set of random variables (Barber 2012) given in Equation 5.1.

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^{n} P\left(X_1 \middle| Pa(X_1)\right) \tag{5.1}$$

As an example, the joint probability distribution of the random variables X_1, X_2 and X_3 shown in Figure 5.2 is estimated by $P(X_1, X_2, X_3) = P(X_1)P(X_2)P(X_3|X_1, X_2)$

In case new information becomes available for one or more chance nodes, BN is able to update the joint probability distribution based on the Bayes' theorem given in Equation 5.2 (Nielsen 2009):

$$P(X|E) = \frac{P(X, E)}{\sum_X P(X, E)} \tag{5.2}$$

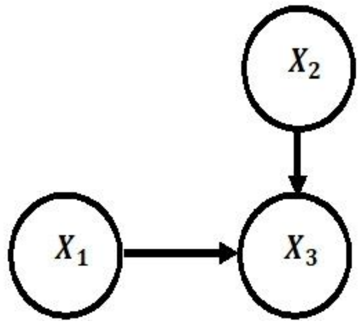


Figure 5.2 A schematic Bayesian Network.

5.3.2 Dynamic Bayesian Network

A DBN represents a stochastic process as a sequence of several time slices, each consisting of inter-dependent nodes. As an illustration, if the BN in Figure 5.2 is expanded into multiple time slices $t = \{1, \dots, T\}$, a DBN will be constructed, as shown in Figure 5.3.

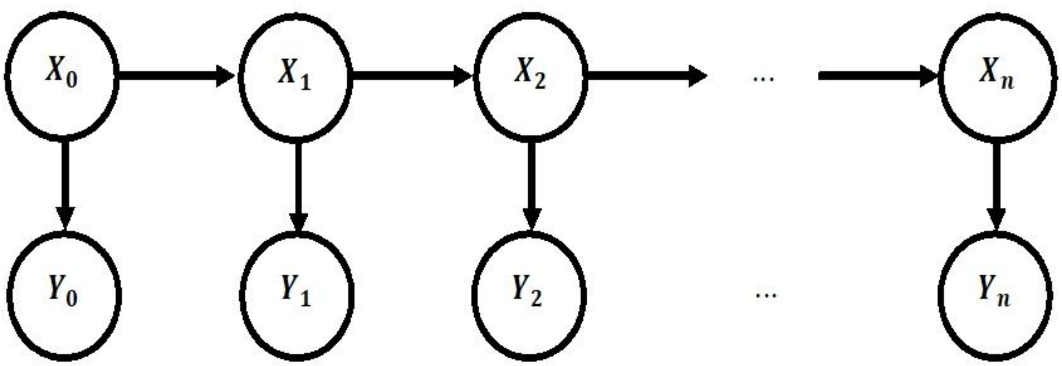


Figure 5.3 Example of Dynamic Bayesian Network.

5.3.3 Influence Diagram

An ID can be established by including utility nodes (diamonds) and decision nodes (rectangles) into a BN (see Figure 5.4). A decision node consists of several decision alternatives available to the user. Since the parents of a decision node incorporate the required information at the time of decision making, the arc pointing to a decision node is an information arc, not an expression of probabilistic dependency. Utility nodes are the descendants of either chance nodes and/ or decision nodes and have no successors. The utility values (including benefits or losses) of a utility node are determined as the preference of the user/operator over each configuration of the decision alternatives and those chance nodes that are the parents of the utility node. Once the ID is completely formed for a decision issue, the expected utility of each decision alternative can be estimated. The optimal decision is the one that maximizes the total expected utility, in agreement with classical decision analyses.

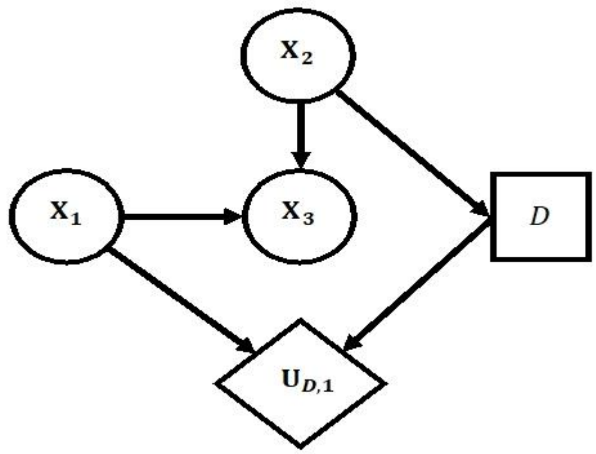


Figure 5.4 An Influence Diagram including utility and decision nodes ($\boldsymbol{X}$: chance nodes, $\boldsymbol{D}$: decision node, $\boldsymbol{U}$: utility node).

5.4 Function time series prediction

The proposed methodology aims at developing a dynamic model that represents the changes in PVs over time. Both mathematical and statistical modelling are applied to predict the behaviour of PVs (more explanation is referred to in (Roffel and Betlem 2006; Tinsley and Brown 2000)). However, implementing an appropriate model for monitoring the nonlinear behaviour of the industrial process is highly important that can assist for minimizing the associated uncertainties. It is then necessary to consider a more reliable model that can represent this stochastic behaviour and provide a tool for better understating the complexity of the process. This will assist the operators and risk and safety managers for a more robust risk analysis and improving the lifetime reliability and availability of the industrial operations. For this purpose, a new perturbation model is developed in this paper to incorporate the fluctuations and uncertainty of the observation data as a basis for conducting future failure analysis and decision-making. As a result, the framework will be able to monitor the nonlinearity of the observation data by considering perturbation model over the life-time of the project for a better understanding of the operational performance.

5.5 Failure analysis

Consider a DBN model that describes the condition of PVs before and after applying a set of perturbations. Failure analysis has been developed to assess the related failures. For the purpose of this class, two subsections will be presented in detail.

5.5.1 PVs Monitoring Mechanism modelling

The generic DBN model for stochastic modelling of PVs is represented in Figure 5.5. The proposed DBN is applied as a generalization of Markov process models. In a Markov process, the future is independent on the past, given the present, as given in Equation 5.3. Here the Markov process is modelled as a chain of nodes that represent the PV.

$$P(X_{t+1}|X_t, \dots, X_0) = P(X_{t+1}|X_t) \tag{5.3}$$

In order to ensure that the DBN is homogenous with identical time slices, the arcs connecting nodes $[\mathcal{E}_1, \dots, \mathcal{E}_T]$ are considered. The transition between these nodes are modelled with diagonal matrices resulting in $\mathcal{E}_t = \mathcal{E}_{t-1}, t = \{2, \dots, T\}$ (similar assumption is implemented for Ω). This is performed to facilitate the model building process and for a better graphical presentation of the model. As suggested by Daniel Straub (2009), these arcs have no impact on the computational efficiency of the model.

Although the model here is proposed in general, the numbers of PVs can vary based on the demand of application with times as:

$$K^j(t) \in K^j = \{K_1^j, K_2^j, \dots, K_t^j\}; j = 1, \dots, n \tag{5.4}$$

For instance, temperature or vibration can be the PV modelled with this method. Owing to the fact that updating in the light of new evidence is counted as a feature of the present model, observations can be adopted from each $K^j(t)$, as given in Equation 5.5:

$$O^j(t) \in O^j = \{O_1^j, O_2^j, \dots, O_t^j\}; j = 1, \dots, n; \tag{5.5}$$

where j is the number of observation and i indicates the PV being monitored. The same condition is assumed for the extent and type of perturbation variables such as system perturbation and exogenous perturbation, see Equation 5.6.

$$\mathcal{E}_q; q = 1,2, \dots, x \tag{5.6}$$

The model has n time slices representing the entire process time divided into discrete number of time steps. All the distributions of variables with continuous analytical expression are discretized into a number of mutually exclusive states. The univariate discretization is proposed so that the continuity in the probability distributions is achieved precisely. More detail of the discretization process of the variables is explained in the following sections.

5.5.2 Model specification

Each perturbation parameter has a stationary process and consequently its probability distribution does not change when shifted in time ($\mathcal{E}_t = \mathcal{E}_{t-1} = \mathcal{E}$, $t = 2, \dots, T$). Therefore, the parameters of the suitable probability distribution must be estimated only once and the CPT of perturbation can be filled after discretization of the final distribution. It is suggested that for the sake of simplicity and without loss of generality, the perturbation data be fitted to a Normal distribution.

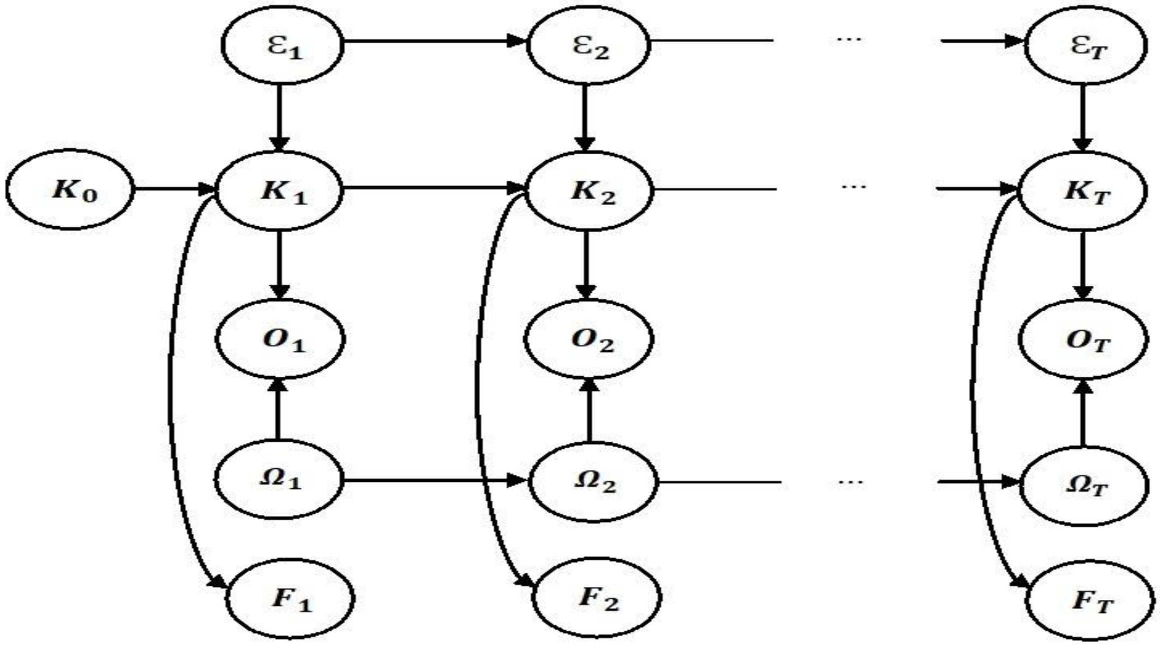


Figure 5.5 Developed DBN for stochastic modeling of process of PVs under impact of exogenous perturbations. Network nodes are, Ɛ: perturbations, K: state of PV, O: observations, Ω: sensor uncertainty, F: Failure.

To obtain the Probability Density Function (PDF) of PV elements in the first time slice, the historical data should be analysed. The available database contains lower and higher bounds and fault threshold rates. Based on the trend and the extreme values, the most suitable distribution for the data can be figured out by several methods such as Maximum Likelihood Estimation (MLE), or Least-Squares Estimation (LSE). MLE has been recommended in previous research (Myung 2003), since it has many features such as efficiency in the calculations, consistency and parameterization invariance. As a result, the MLE is adopted in the present study and the PDF of PVs is accordingly discretized.

The CPT of PVs ($P(K_i \mid K_i - 1,\ \mathcal{E}_i)$) is defined with binary values based on the limit state concept. These binary values are presented in $N \times N \times M$ transition probabilities, where N and M are the state numbers of K_i and $\mathcal{E}_i$, respectively. Limit state function is discussed further, later in this section. In order to fill the transitional CPTs, it is necessary to define a safe operational interval for the considered PV. For instance, the interval $[a, b]$ can be chosen to determine whether the PV is within this interval. It is through this comparison that CPTs can be filled, as given in Equation 5.7:

$$CPT = \begin{cases} if\ a \leq PV \leq b \quad then & 0 \\ else & 1 \end{cases} \tag{5.7}$$

The DBN model used in the proposed methodology provides the user with an opportunity to consider new evidence to update the probability distributions. Observations can be made from many strategies such as real time monitoring and failure monitoring. In the present study, inspection results are incorporated into the network and the uncertainty associated with the results is characterized by Probability of Detection (PoD), Daniel Straub (2004) provides a number of PoD functions based on empirical methods. A common approach to define the PoD function is the one-dimensional exponential threshold model, previously used by several researchers (Ambühl 2017; J. S. S. Nielsen, J. D. 2011; J. S. Nielsen and Sørensen 2017; D. Straub 2004) ,and given by:

$$PoD(D) = P_0 \left[1 - \exp\left(-\frac{D}{\lambda}\right)\right] \tag{5.8}$$

where D is as the actual perturbation, $P_0$the maximum probability of detection and λ is the expected value of the smallest detectable perturbation.

In order to complete the PoD model, probability distributions are discretized into E states. It should be noted that the number of states for node O should be the same as the states of node K. The discretized probabilities are set in the first column of the $N \times E$ $(N = E)$ matrix. The perturbation in the former states of PV cannot be detected as damage in the latter states of inspection (for example the perturbation value in $K_1$is not detectable in O_2 or O_3). So, the final CPT of $P(O_i \mid K_i)$ is as follows:

$$\begin{bmatrix} PoD_1 & 0 & 0 & \ldots. & 0 & 0 \\ PoD_2 & PoD_2 & 0 & \ldots. & 0 & 0 \\ . & . & . & \ldots. & . & . \\ . & . & . & \ldots. & . & . \\ . & . & . & \ldots. & . & . \\ . & . & . & \ldots. & . & . \\ PoD_{N-1} & PoD_{N-1} & PoD_{N-1} & \ldots. & PoD_{N-1} & 0 \\ PoD_N & PoD_N & PoD_N & \ldots. & PoD_N & 1 \end{bmatrix} \tag{5.9}$$

The method for estimating PoD in other time slices (from the second time slice onwards) is different from the first, since these nodes have an extra parent node which is the node incorporating the uncertainty of sensors. Although PoD function is applied to model the reliability of inspection, the uncertainty of sensor values can be represented in three forms, from no attention to uncertainty at all, to the highest resolution of uncertainty information: point uncertainty, interval uncertainty and probabilistic uncertainty (Cheng 2003). In the present paper, probabilistic uncertainty approaches are adopted.

Considering Ω_i, the model reflects the reliability of sensors as well. As a general concept of this work (as done for perturbation), normal distribution is proposed as the suitable probability distribution being fitted to uncertainty of sensors, however, other distributions can be adopted based on the available data and characteristics of sensors. This parameter is time-invariant, so, the calculation of PDF and discretized value of probability must be done only once for the whole process.

Assuming that N and E are the state numbers of O_i and K_i subsequently, and L is the number of states (S) in node Ω_i, the final CPT of $P(O_i| \, K_i \, , \, \Omega_i)$ is shown in Equation (10) in the form of $[N \times E] \times [L]$:

$$\begin{bmatrix} 1-\frac{1}{N}\Omega & 0 & 0 & \dots. & 0 & 0 \\ 0 & 1-\frac{2}{N}\Omega & 0 & \dots. & 0 & 0 \\ 0 & 0 & 1-\frac{3}{N}\Omega & \dots. & 0 & 0 \\ . & . & . & \dots. & . & . \\ . & . & . & \dots. & . & . \\ . & . & . & \dots. & . & . \\ 0 & 0 & 0 & \dots. & 1-\frac{N-1}{N}\Omega & 0 \\ \frac{1}{N}\Omega & \frac{2}{N}\Omega & \frac{3}{N}\Omega & \dots. & \frac{N-1}{N}\Omega & 1 \end{bmatrix} \times [S_1, S_2, \dots, S_L] \tag{5.10}$$

Failure probability is assessed using limit state function (Kamphuis 2010). This approach is adopted as follows here:

$$G = C - PR \tag{5.11}$$

where G is failure function, C is critical PV interval and PR is the actual PV interval. Consequently, the conditional probability of failure $P(F_i \mid K_i)$ in the DBN is expressed with two states of Safe and Fail, 1 (fail) if $G \leq 0$ and 0 (safe) when $G > 0$.

5.6 Decision making support tool

The next stage of the methodology is to develop an ID for optimising the maintenance. The ID developed in this paper (see Figure 5.6) incorporates the socio-economic aspects including operation and maintenance costs into the decision-making process. As discussed in section 5.5.3, two additional node types, utility and decision nodes, are added to the DBN for constructing the ID.

The decision node (M_i) characterizes different decision alternatives (repair, replace, continue without any maintenance actions). It is made based on the results of inspections and subsequently it affects its descendent including the chance node ($K'(t)$). The nodes $K'(t)$ are introduced into the network for discriminating between the status of the PV before and after a decision regarding maintenance. In case the decision is made to continue without any maintenance actions, the CPTs of $K'(t)$ nodes are identical to that of $K(t)$ node from the same time slice. This is while the CPTs vary if maintenance (repair or replace) is carried out in the previous time slice.

It should be noted that if a decision is made for repairing or replacing the system, the state of PV will be recovered to its initial time steps. The value of recovery is directly depending on the norms and practices in the industry of interest.

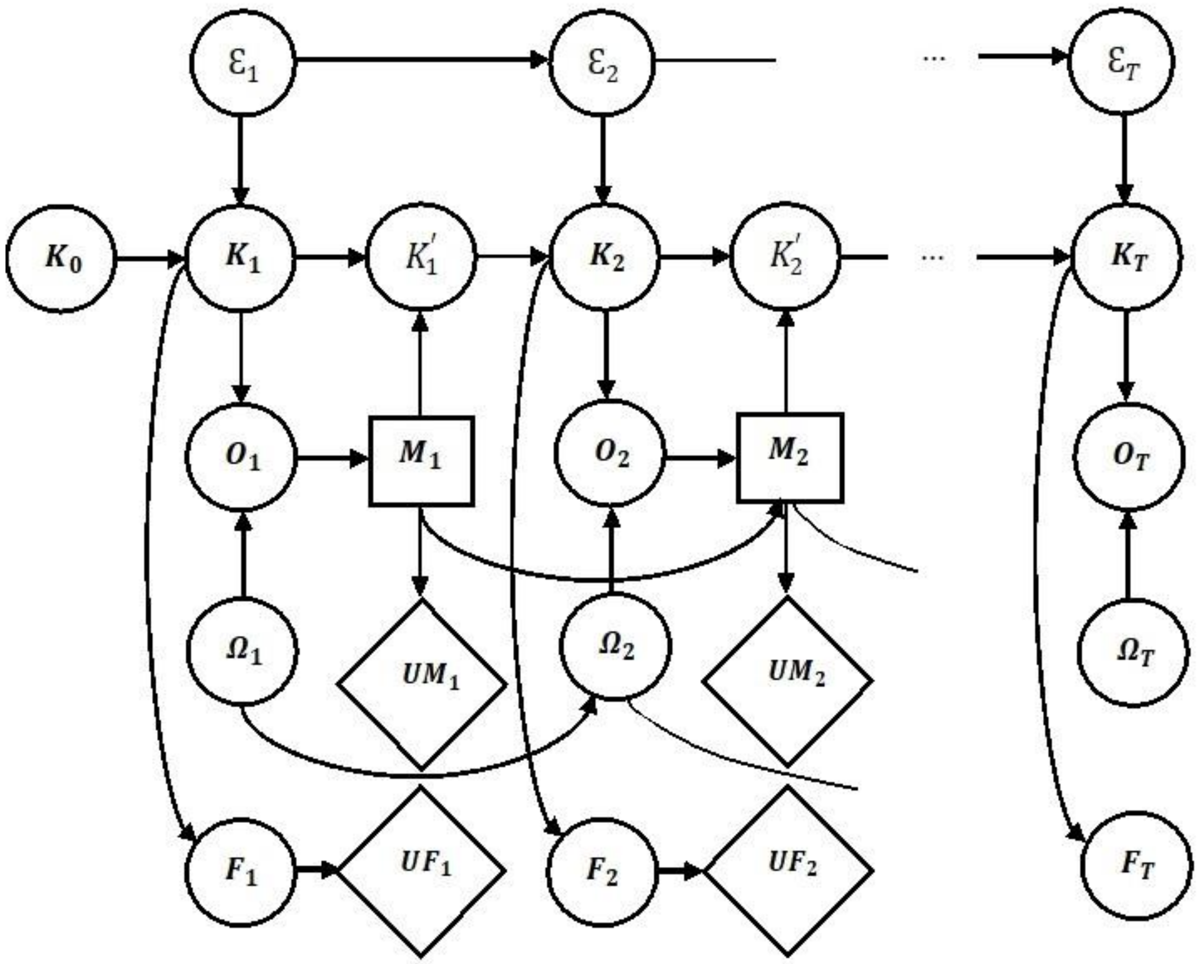


Figure 5.6. The ID of the multi-criteria decision-making model developed for maintenance planning of a stochastic process. Network nodes are, $\mathcal{E}$: perturbations, K: PVs condition, K': PVs condition after maintenance, O: observations, Ω: device uncertainty, F: Failure, M: Maintenance, UM: utility of maintenance, UF: Utility of Failure.

The costs associated with system failure are accounted for using the utility of failure (UF_i node) in the network. The initial time slices can be filled based on the amount of collected data from the structure. The inevitable cost of failure in the future periods of operation is given by (Usher et al. 1998):

$$CF_j = \overline{CF}.\overline{h}_j(1+f)^j \tag{5.12}$$

Where cost of failure (CF) in period (j) is estimated by a simple Rate of OCcurrence Of Failures (ROCOF) constant, $\overline{CF}$ (in units of \$/unit-failure-rate) multiplied by the average ROCOF, ($\bar{h}$). It is also assumed that the cost of a failure taking a place in future will be subjected to inflation at a rate of f percent in the considered period of j. For the sake of simplification, it is suggested a linear approximation be considered for the average of ROCOF (Referred to (Usher et al. 1998) for more explanations).

The utility values developed for maintenance alternatives are suggested to be evaluated in detail separately for each configuration. The cost of replacement of the equipment is estimated as:

$$CR_j = \overline{CR}(1+r)^j \tag{5.13}$$

where CR is a constant cost for replacing the equipment. In the present paper, the values for CR are adopted from historical data. Similar to the case of failure, a separate inflation rate (r) is considered for replacements over the period j.

Finally, if the system requires a repair, the regular cost for this activity (CR') will be affected by an inflation rate of r' percent per period, therefore the cost of repair is given by Equation 5.14.

$$CR'_j = \overline{CR'}(1+r')^j \tag{5.14}$$

5.7 Application of developed methodology: Case study

An example of Natural Gas Regulating and Metering Stations (NGRMS) is given to show how the application of developed ID in risk-based maintenance can be applied. GeNie software is used as a tool for modelling the ID. A detailed discussion on application of each step of the developed methodology in the case study is discussed in the following sections.

5.7.1 Scenario development

NGRMS are established in the gas distribution networks to reduce the Natural Gas outlet pressure to a setting value. To handle the process, there are two regulating streams with two regulators arranged in a series for each line. One is the main regulator and the other is used as a control/slam shut valve. Through the normal operation, one line is working while the other line is on stand-by. If main and slam regulator (both) fail, the other standby line starts to work. As illustrated in Figure 5.7, the standard configurations of lines in NGRMS are made up of control valve, pressure regulator-passive controller, main pressure regulator with a built-in slam-shut valve and filters.

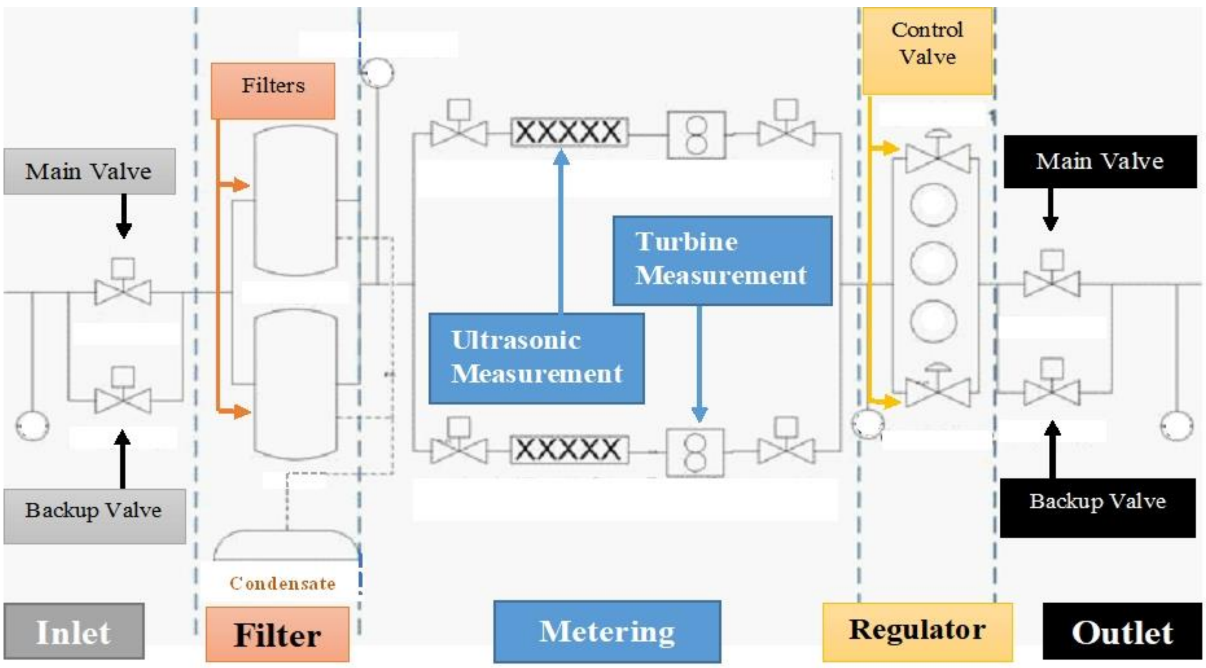


Figure 5.7 Simple System Architecture of NGRM stations.

The requirements for particular characteristics of delivered gas, including gas specification, odorisation and pressure are specified in both local and international regulations. (see (UNI 2009a, 2009b, 2012)). Such regulations are issued to obtain smooth operations with the lowest possible number of maintenance interruptions, failure losses and accidental damages. In the present study, the pressure is applied as PV to analyse the deterioration process and to finally achieve the optimal time schedule of maintenance.

5.8 Function prediction

To set up the decision making process, the pressure values during 36 months of process operation were taken into account. The time series predictions are depicted in Figure 5.8 along with historical and validation data. The historical data is predicted by different regression tools to find the most suitable one based on their predictive performance. In this study, the competitive evaluation of models summed up and selected Fourier as the suitable one due to stable results across samples with below representation (see Equation 5.15).

$$P(t) = \alpha_0 + \sum_{i=1}^{\infty}\left(\alpha_i \cos\frac{i\pi t}{L} + \beta_i \sin\frac{i\pi t}{L}\right) \tag{5.15}$$

Where first term of Fourier equation (α_0) is actually the expected amount of observed pressure since it is defined by Equation 5.16.

$$\alpha_0 = \frac{1}{T}\int_0^T P(t) = \bar{P} \tag{5.16}$$

Based on historical data, it can be reckoned that although the process engineering gives protection for pressure behaviours against perturbations, there are nevertheless a wide range of perturbations in the pressure. These perturbations can be modelled by regression tools of pressure through time. As a result, α_i and β_i are considered as independent perturbation parameters and follow from:

$$\mathcal{E} = \{\alpha_i, \beta_i;\ i = 0, 1 \ldots, 6\} \tag{5.17}$$

$$\alpha_i = \frac{1}{T}\int_0^T P(t) \cos\frac{i\pi t}{L} \tag{5.18}$$

$$\beta_i = \frac{1}{T}\int_0^T P(t)\sin\frac{i\pi t}{L} \tag{5.19}$$

By taking pressure (including perturbation) into account, the condition of pressure is predicted in time based on its initial treatments. Each α and β in the model experienced a normal distribution with specific mean (μ) and standard deviation (σ^2). So Equation 5.15 can be represented as:

$$P(t) = \bar{P} + \sum_{i=1}^{\infty}\left(\frac{1}{\sqrt{2\pi\sigma_i{}^2}}e^{-\frac{(\alpha_i-\mu_i)}{2\sigma_i{}^2}}\cos\frac{i\pi t}{L} + \frac{1}{\sqrt{2\pi\sigma_i{}^2}}e^{-\frac{(\beta_i-\mu_i)}{2\sigma_i{}^2}}\sin\frac{i\pi t}{L}\right) \tag{5.20}$$

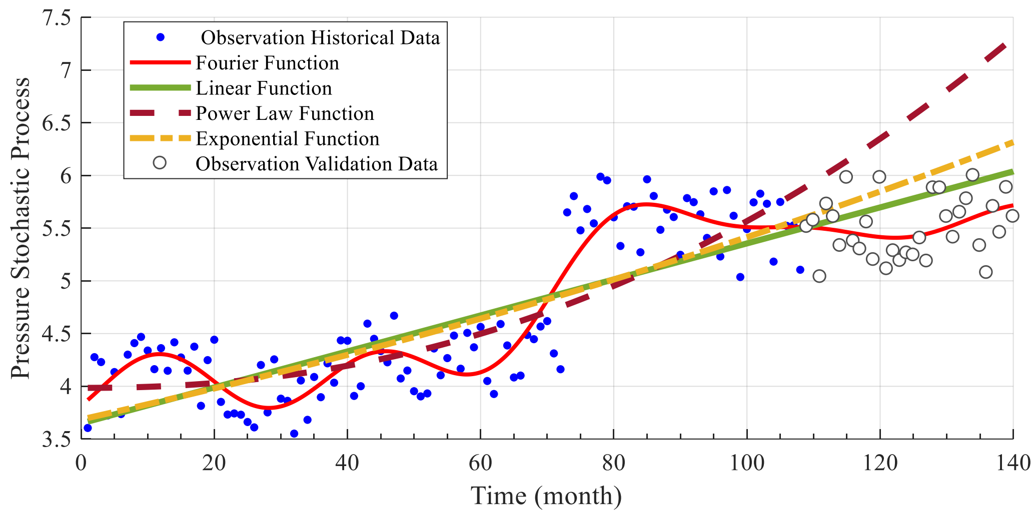


Figure 5.8 Time series ahead prediction of Pressure treatment.

5.9 Pressure monitoring model

To demonstrate the time dependent stochastic modelling of a PV (pressure), a DBN model is developed (see Figure 5.9). For the purpose of this study, the model is simplified by analysing the pressure behaviour for a period of four seasons (each season representing a time slice; $P_0, P_1, \ldots, P_4$) with influence of exogenous perturbations on it. Although in reality, the system will often be maintained (repaired or replaced) at a fixed time, especially after detection of failure, in the method presented in section 5.4, it is assumed that the system has not been maintained for three years.

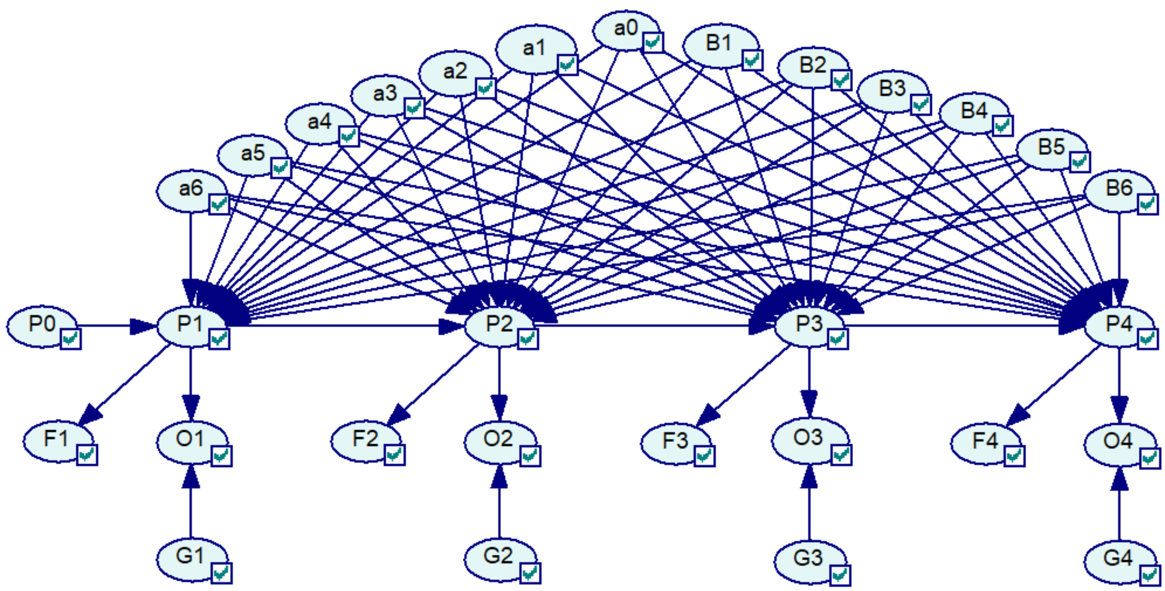


Figure 5.9 Developed DBN model including exogenous perturbations for four seasons. Network nodes are, $\{\boldsymbol{\alpha_0}, \boldsymbol{\alpha_1}, \dots, \boldsymbol{\alpha_6}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \dots, \boldsymbol{\beta_6}\}$:perturbations, $\boldsymbol{P_0}$: Initial Pressure condition, $\boldsymbol{P}$: Pressure condition, $\boldsymbol{O}$: observations, $\boldsymbol{G}$: device uncertainty, $\boldsymbol{F}$: Failure.

The parameter P_0 accounts for the initial pressure (historical data of pressure) and has a Weibull distribution with scale and shape parameters of A and B respectively. It is assumed that both parameters have negligible deviation and a constant rate of $A = 4.455$ and $B = 10.244$. On the contrary, other parameters applied in the model are normally distributed. The distribution of each parameter mentioned in Table 5.1 is split into a specific number of intervals. In addition to these parameters, the observation node is similarly discretised using 10 intervals while avoiding round-off errors by using MATLAB software (see Figure 5.10).

It should be mentioned that pressure size in the following time slices are discretized using the same uniform interval lengths as P_0. A detailed discussion of the sequence of filling the CPTs for all parameters (pressure size $P\ (P_i \mid P_{i-1}, \mathcal{E}_i\ )$, observation $P\ (O_i \mid P_i, \Omega_i\ )$, and failure $P\ (F_i \mid P_i)$) is stated in section 5.5.2.

Table 5.1 Parameters of stochastic modelling of pressure with univariate perturbations variables.

Variable	Description	Distribution (discretized interval)	Mean	Standard deviation
$\boldsymbol{\alpha_0}$	1st Perturbation parameter	Normal (5)	4.885	3.118
$\boldsymbol{\beta_1}$	2nd Perturbation parameter	Normal (5)	-0.638	3.031
$\boldsymbol{\alpha_1}$	3rd Perturbation parameter	Normal (5)	-0.6945	2.8985
$\boldsymbol{\beta_2}$	4th Perturbation parameter	Normal (5)	-0.11	0.3829
$\boldsymbol{\alpha_2}$	5th Perturbation parameter	Normal (5)	-0.7374	5.6264
$\boldsymbol{\beta_3}$	6th Perturbation parameter	Normal (5)	0.2462	1.7332
$\boldsymbol{\alpha_3}$	7th Perturbation parameter	Normal (5)	-0.2168	2.5258
$\boldsymbol{\beta_4}$	8th Perturbation parameter	Normal (5)	0.0346	3.5336
$\boldsymbol{\alpha_4}$	9th Perturbation parameter	Normal (5)	-0.0044	0.9305
$\boldsymbol{\beta_5}$	10th Perturbation parameter	Normal (5)	0.1719	1.1268
$\boldsymbol{\alpha_5}$	11th Perturbation parameter	Normal (5)	0.0515	0.6279
$\boldsymbol{\beta_6}$	12th Perturbation parameter	Normal (5)	0.1096	0.561
$\boldsymbol{\alpha_6}$	13th Perturbation parameter	Normal (5)	-0.0989	2.4919
Ω	Devices uncertainty	Normal (3)	0.0002	0.05

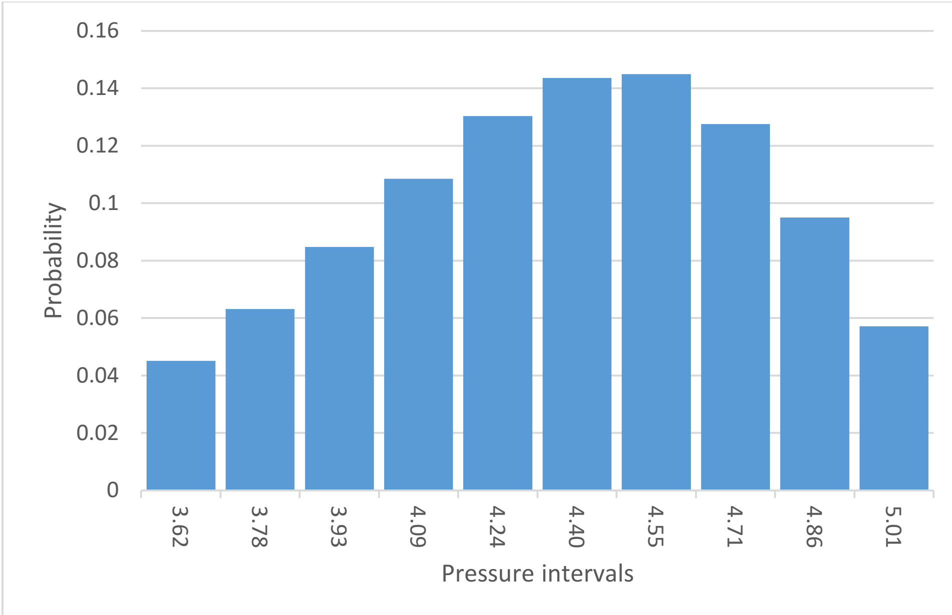


Figure 5.10 Discretized Weibull Distribution of initial pressure size.

5.10 Utility efficiency

Recognizing the optimal maintenance strategies and times are conducted by an extension of DBN into ID (see Figure 5.11; due to space limitation nodes P_0 to P_2 of the decision model are only depicted). To evaluate the effect of maintenance on process, the maintenance alternatives are defined consequently in three actions including continue, repair and replace. The elements of drawn ID were previously introduced in section 5.6.

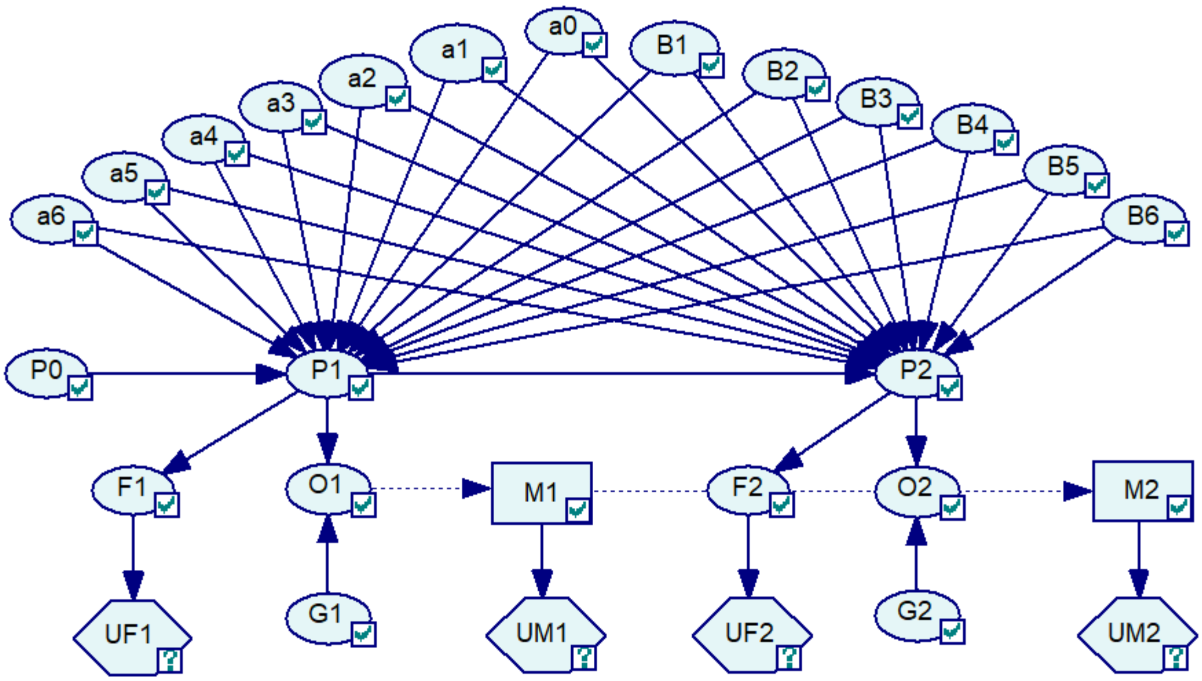


Figure 5.11 Developed Influence Diagram for maintenance planning considering exogenous perturbations. Network nodes are $\{\alpha_0, \alpha_1, \ldots, \alpha_6, \beta_1, \beta_2, \ldots, \beta_6\}$:perturbations, P_0: Initial Pressure condition, P: Pressure condition, O: observations, G: device uncertainty, F: Failure M: Maintenance, UF: Utility of Failure UM: Utility of Maintenance.

Iqbal et al. (2016) presented a comprehensive review on inspection and maintenance policies for oil and gas pipelines. They defined the repair of a unit as Imperfect maintenance after which, although the unit is not taken into account as new, it is supposed to be younger than before. The replacement is also assumed to be established either at complete failure or after fixed number of failures. To improve the effectiveness of the decision making process, the hybrid policies have been examined with mentioned decision alternatives later.

Based on aforementioned assumptions and Equation 5.13 and Equation 5.14, the costs associated with repair and replacement are compared and depicted in Figure 5.12. The line graph illustrates the repair value and bar chart represents the replacement expenses for the entire domain of pressure (as illustrated in Figure 5.10. the pressure variable is discretised in 10 intervals). Units are measured in Euros.

Overall, the expected cost is changing through different intervals for repair, while experiencing constant rate for replacement. It is proved again that the most desired pressure value is starting from the third and finishing at the fifth interval as the repair costs decline and rise significantly before and after these intervals. Additionally, it is necessary to note that steady rate of replacement cost does not mean that for any conditions of pressure in any time of replacement, the expected cost would be the same. This will be explained in detail in the following section by considering different occasions.

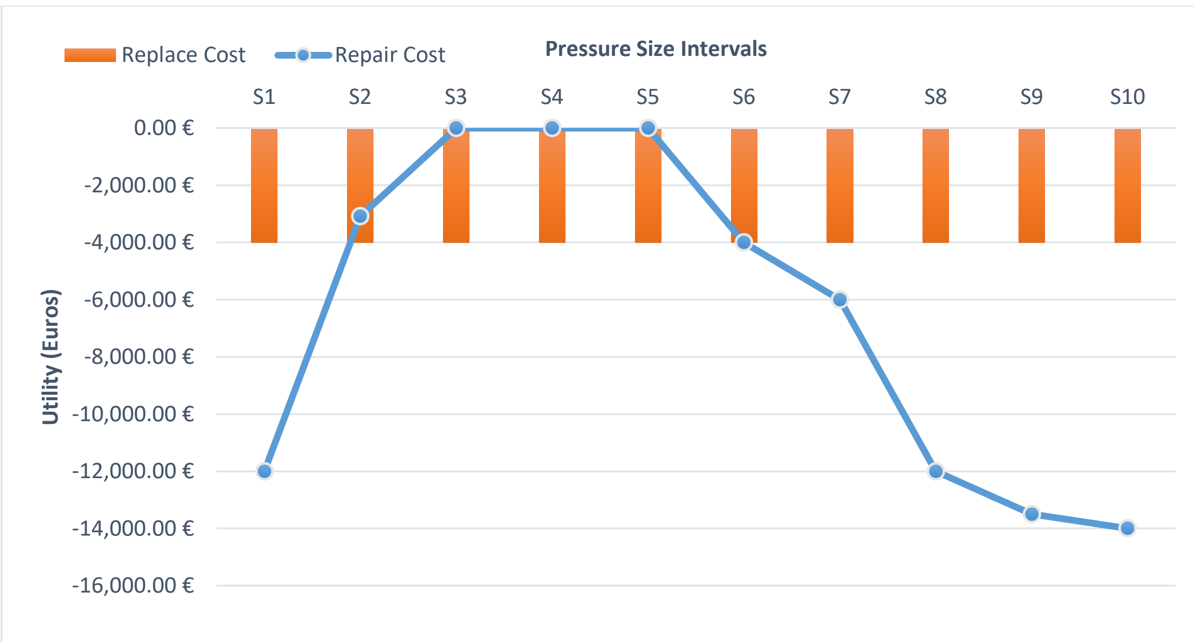


Figure 5.12 Utility value of maintenance alternatives, repair and replace, for each interval of pressure.

5.11 Influence Diagram application: results

To assess the advantage of the developed methodology, three different seasonal inspection cases were considered. To make clear reported data in Table 5.2, in case B, the observations are made with a pressure in state 2 followed by state 7 of pressure intervals in the third season. The health of the system is not monitored for the second and last seasons.

Table 5.2 Observations of pressure size in the NGRM station. Three cases of different monitoring results were considered. Note: the cells with dashes illustrate times where monitoring is not performed.

Month	3	6	9	12
Case A	State 1	State 6	State 8	-
Case B	State 2	-	State 7	-
Case C	State 8	-	-	-

The line graphs in Figure 5.13, Figure 5.14 and Figure 5.15 depict the Expected Utility (EU) for three maintenance alternatives (repair, replace and continue) based on inspection results reported in Table 5.2 over a period of one year.

Starting with case A, the deterioration is mapped through the gradual increase of pressure from its 1st state to 6th and lastly 8th state. Although at the beginning of the considered period, continuing the operation is the most beneficial option, the subsequent drop of EU in this line at the second season implies that this is not an appropriate alternative after six months. Based on results depicted in Figure 5.13**,** it is deduced that the optimal strategy is continuing at the first season, followed by repairing at the second stage. The utility of all three options for the final season is predicted to be approximately equal. Ultimately, according to the model, the maximum benefits are achieved if in the 3rd and 4th seasons replace and continue alternatives are applied respectively, where the EU reaches a peak of 12000 and 1000 Euros.

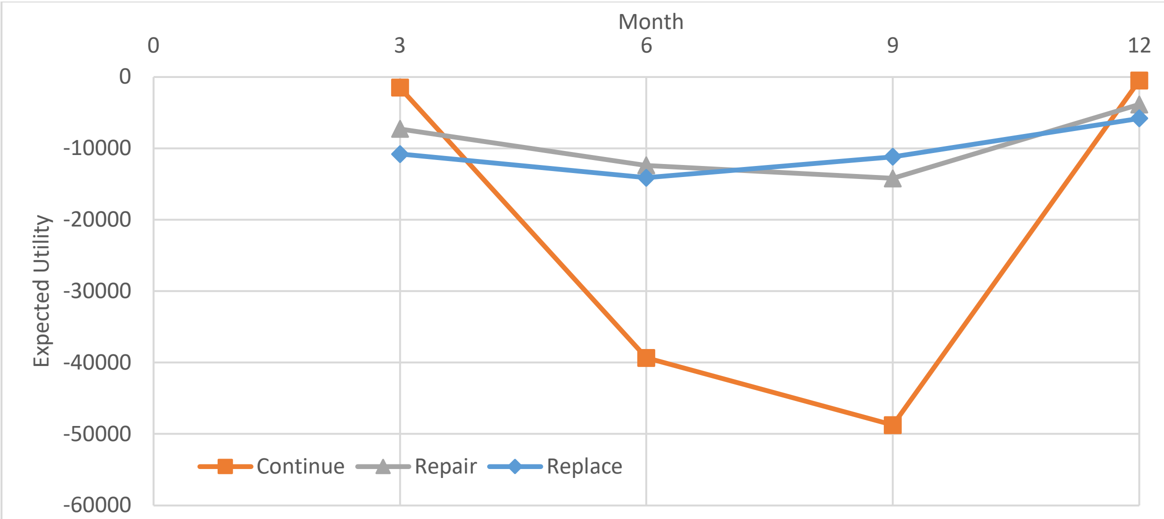


Figure 5.13 Expected utilities of three decision alternatives: Replace, Repair and Continue operation for case (A) with different pressure size incidents as detailed in table 5.2.

In case B, continuing the operation is considered as the best configuration of maintenance decision in the first 2 seasons. As can be seen in Figure 5.14, since the pressure experienced its 7^{th} interval at the end of the 9^{th} month, it is proposed that the system must undergo a repair process at the third season to recover its healthy state. This action has the maximum EU of approximately 15000 € at 3^{rd} season NGRMS. Similar to case A, it is predicted that conducting the suitable maintenance policy optimizes the EU in the upcoming season.

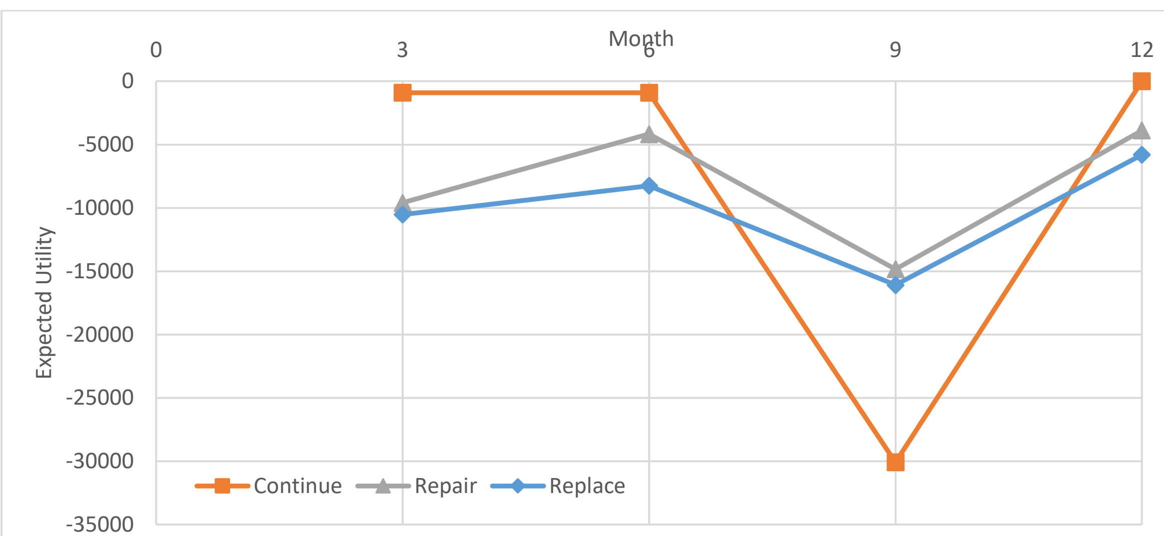


Figure 5.14 Expected utilities of three decision alternatives: Replace, Repair and Continue operation for case (B) with different pressure size incidents as detailed in table 5.2.

Considering Case C (illustrated in Figure 5.15), the EU of continue option, deviates noticeably over the period given, while for the other two alternatives it changes minimally. Considering the status of pressure in the first season, the model assesses the EU of replace as the optimal alternative where it reaches a peak of about 4000 Euros. Executing this decision configuration will result in a surge in EU of other options in the future, this trend can be seen chiefly through continue to the end of studied time.

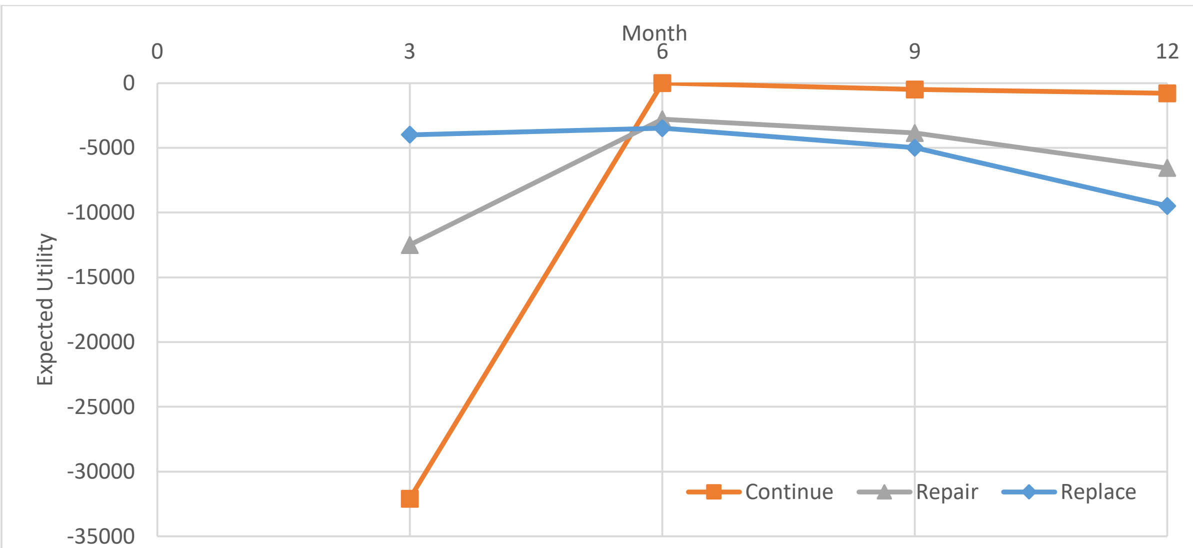


Figure 5.15 Expected utilities of three decision alternatives: Replace, Repair and Continue operation for case (C) with different pressure size incidents as detailed in table 5.2.

5.12 Conclusion

A novel methodology using Markov degradation model as an underlying principle of decision making is developed to estimate the optimal maintenance time schedule. The treatment of PVs under the influence of perturbations in time series has been analysed applying DBN and ID. Furthermore, the proposed approach enables investigating uncertainty related to parameters, models and historical data through limit state function. The failure mode has also been explained in a limit state equation. The model has been enabled to update the probability based on new observation of system. The reliability of inspection has been characterized by PoD through one-dimensional exponential threshold model. In addition to model the reliability of inspection, the uncertainty of sensor values is also represented. The expected cost associated with failures and maintenances is estimated considering inflation. The study has been implemented on actual examples of stochastic deterioration process of Natural Gas Regulating and Metering Stations (NGRMS) in order to validate the proposed method using real field data. The pressure has been taken into account as PV. The Fourier series is used as the regression tool to predict the trend of pressure considering perturbation parameters in time. To examine the method, three different seasonal inspection cases have been introduced into the network to determine the optimum maintenance times and strategies. The proposed framework highlights that repairing the components in second season for the first time is the most economic decision for case A. this is while the optimum approach for case B is that the system undergoes a repair process at the third season. In case C also the model assesses the EU of replace as the most cost-effective strategy on the 3rd month of the operation. The present framework is able to perform a risk-based maintenance planning for the industrial processes aiming at minimizing the operational and maintenance costs. A non-stationary model can be developed in order to predict the trend of PVs in the process, since condition monitoring data tend to have a more non-linear behaviour. Therefore, it is recommended to apply a noise reduction approach for filtering the observations data and smoothing the unexpected fluctuations.

6. Conclusions and Recommendations

The proposed research attempts to undertake risk-based asset integrity modelling of engineering process. With four main objectives, different frameworks developed to evaluate the reliability engineering operation subjected to random stochastic behavior. To deal with the first objective a differences between application of two mostly utilized assumptions in failure modelling, MR and PR, have been addressed in this paper. This was carried out in a case study of natural gas regulation and measurement plant by MLE and Bayesian inference method. The final results highlighted that relaxing the renewal process assumption (constant failure rate) and taking the time dependency between the observed failure times into account, results in a more precision of failure modelling where the shape parameter value of Weibull distribution in both parameter estimation approaches (HBM, MLE) are higher than 1, confirming that the number of failure events are dependent upon time. This study can help asset managers to optimize the reliability assessment of repairable systems based on available data. To address the second objective, a methodology for time dependent reliability assessment of engineering operations were presented by considering a strategy for noise reduction in monitoring demanding parameters. The considered data had nonlinear and non-stationary nature, so it could not be analysed by a standard method, e.g., SPC or LSR. Subsequently, in order to remove the noise from the raw data in the observation process, EMD was selected as the statistical tool to filter out the data. The results show that the expected time for exceeding the safe limit is 50 weeks with a credible interval of (38, 64) weeks for the 2.5 and 97.5 percentile of estimated distribution, respectively. The predicted exceedance distributions facilitate the exploration of the onset of deterioration. The developed methodology is capable of being considered as a predictive tool for estimating lifetime condition of an engineering process,

The third objective is achieved by developing a comprehensive methodology for predicting the trends of safety indicators affected by stochastic risk factors through the supply energy services sector. Given the uncertainty and complication associated with captured operational data, in the developed model, Bayesian inference with hierarchical structure and GLM was integrated to define a regression function. The obtained regression function is capable for forecasting on-line reliability assessment and consequently developing more efficient remediation plans. The predictions suggest that in 2% of the operational period, the probability of failure appraised at an interval of [10-4, 3.5×10-3] per day. The final objective is achieved by assessing a novel methodology using Markov degradation model as an underlying principle of decision making is developed to estimate the optimal maintenance time schedule. The treatment of PVs under the influence of perturbations in time series has been analysed applying DBN and ID. Furthermore, the proposed approach enables investigating uncertainty related to parameters, models and historical data through limit state function. The failure mode has also been explained in a limit state equation. The model has been enabled to update the probability based on new observation of system. The reliability of inspection has been characterized by PoD through one-dimensional exponential threshold model. In addition to model the reliability of inspection, the uncertainty of sensor values is also represented. To examine the method, three different seasonal inspection cases have been introduced into the network to determine the optimum maintenance times and strategies. The proposed framework highlights that repairing the components in second season for the first time is the most economic decision for case A. this is while the optimum approach for case B is that the system undergoes a repair process at the third season. In case C also the model assesses the EU of replace as the most cost-effective strategy on the 3rd month of the operation. The present framework is able to perform a risk-based maintenance planning for the industrial processes aiming at minimizing the operational and maintenance costs.

6.1 Recommendations

In order to boost the risk based asset integrity modelling and as a potential future research direction, it is suggested to:

- effectively using the diverse, high-dimensional, high-velocity condition monitoring data of industrial assets to improve their availability and resilience.
- design of robotic systems that can self-certify and guarantee their safe operation; Self-diagnosis of faults and self-healing
- extending the concept of IoT to develop the ‘social network of things’. If machines can report their ‘status’ into a common data-sharing platform – or social network – it becomes possible to create a single view of how the whole factory is running.

7. Bibliography

1. Abaei, M. M., E. Arzaghi, R. Abbassi, V. Garaniya and S. Chai (2018a). A novel approach to safety analysis of floating structures experiencing storm. Ocean Engineering 150: 397-403.
2. Abaei, M. M., E. Arzaghi, R. Abbassi, V. Garaniya, M. Javanmardi and S. Chai (2018b). Dynamic reliability assessment of ship grounding using Bayesian Inference. Ocean Engineering 159: 47-55.
3. Abaei, M. M., E. Arzaghi, R. Abbassi, V. Garaniya, S. Chai and F. Khan (2018c). A robust risk assessment methodology for safety analysis of marine structures under storm conditions. Ocean Engineering 156: 167-178.
4. Abaei, M. M., R. Abbassi, V. Garaniya, S. Chai and F. Khan (2018). Reliability assessment of marine floating structures using Bayesian network. Applied Ocean Research 76: 51-60.
5. Abaei, M. M., E. Arzaghi, R. Abbassi, V. Garaniya and I. Penesis (2017). Developing a novel risk-based methodology for multi-criteria decision making in marine renewable energy applications. Renewable energy 102: 341-348.
6. Abbassi, R., J. Bhandari, F. Khan, V. Garaniya and S. Chai (2016). Developing a quantitative risk-based methodology for maintenance scheduling using Bayesian network. Chemical Engineering Transactions 48: 235-240.
7. Ahmed, Q., Moghaddam, K. S., Raza, S. A., & Khan, F. I. (2015). A multi-constrained maintenance scheduling optimization model for a hydrocarbon processing facility. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 229(2), 151-168.
8. Allgood, G. O. and B. R. Upadhyaya (2000). Model-based high-frequency matched filter arcing diagnostic system based on principal component analysis (PCA) clustering. Applications and Science of Computational Intelligence III, International Society for Optics and Photonics.

9. Ambühl, S., & Sørensen, J. D. (2017). On Different Maintenance Strategies for Casted Components of Offshore Wind Turbines.
10. Antonino-Daviu, J., S. B. Lee and E. Strangas (2017). Guest editorial special section on advanced signal and image processing techniques for electric machines and drives fault diagnosis and prognosis. IEEE Transactions on Industrial Informatics, Institute of Electrical and Electronics Engineers.
11. Antonov, A. and V. Chepurko (2017). Statistical analysis of data on failures of the nuclear plant equipment in conditions of a non-homogeneous flow of events. Part 2. Nuclear Energy and Technology 3(3): 206-210.
12. Aremu, O. O., Palau, A. S., Parlikad, A. K., Hyland-Wood, D., & McAree, P. R. (2018). Structuring Data for Intelligent Predictive Maintenance in Asset Management. IFAC-PapersOnLine, 51(11), 514-519.
13. Artigao, E., S. Martín-Martínez, A. Honrubia-Escribano and E. Gómez-Lázaro (2018). Wind turbine reliability: A comprehensive review towards effective condition monitoring development. Applied energy 228: 1569-1583.
14. Arunraj, N. S., & Maiti, J. (2007). Risk-based maintenance—Techniques and applications. Journal of hazardous materials, 142(3), 653-661.
15. Arzaghi, E., Abaei, M. M., Abbassi, R., Garaniya, V., Chin, C., & Khan, F. (2017). Risk-based maintenance planning of subsea pipelines through fatigue crack growth monitoring. Engineering Failure Analysis, 79, 928-939.
16. Arzaghi, E., M. M. Abaei, R. Abbassi, V. Garaniya, J. Binns, C. Chin and F. Khan (2018a). A hierarchical Bayesian approach to modelling fate and transport of oil released from subsea pipelines. Process Safety and Environmental Protection 118: 307-315.
17. Arzaghi, E., R. Abbassi, V. Garaniya, J. Binns and F. Khan (2018b). An ecological risk assessment model for Arctic oil spills from a subsea pipeline. Marine pollution bulletin 135: 1117-1127.
18. Asher, H. and H. Feingold (1984). Repairable system reliability: Modeling, Inference, Misconceptions and their cause, Marcel Dekker, New York.
19. Averill, A., J. Ingram, P. Holborn, P. Battersby and C. Benson (2018). Application of Bayesian methods and networks to ignition hazard event prediction in nuclear waste decommissioning operations. Process Safety and Environmental Protection 116: 396-404.

20. Azadeh, A., Gharibdousti, M. S., Firoozi, M., Baseri, M., Alishahi, M., & Salehi, V. (2016). Selection of optimum maintenance policy using an integrated multi-criteria Taguchi modeling approach by considering resilience engineering. The International Journal of Advanced Manufacturing Technology, 84(5-8), 1067-1079.
21. Back, J. A., L. P. Tedesco, R. F. Molz and E. O. B. Nara (2016). An embedded system approach for energy monitoring and analysis in industrial processes. Energy 115: 811-819.
22. BahooToroody, A., M. M. Abaei, E. Arzaghi, F. BahooToroody, F. De Carlo and R. Abbassi (2019a). Multi-level optimization of maintenance plan for natural gas system exposed to deterioration process. Journal of hazardous materials 362: 412-423.
23. BahooToroody, A., M. M. Abaei, F. BahooToroody, F. De Carlo, R. Abbassi and S. Khalaj (2019b). A condition monitoring based signal filtering approach for dynamic time dependent safety assessment of natural gas distribution process. Process Safety and Environmental Protection.
24. Baptista, M., S. Sankararaman, I. P. de Medeiros, C. Nascimento Jr, H. Prendinger and E. M. Henriques (2018). Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. Computers & Industrial Engineering 115: 41-53.
25. Barabadi, A., J. Barabady and T. Markeset (2014). Application of reliability models with covariates in spare part prediction and optimization–a case study. Reliability Engineering & System Safety 123: 1-7.
26. Barber, David. Bayesian reasoning and machine learning. Cambridge University Press, 2012.
27. Barnard, I. (2006). Asset Management—An Insurance Perspective. In Engineering Asset Management (pp. 44-53). Springer, London.
28. Barua, S., Gao, X., Pasman, H., & Mannan, M. S. (2016). Bayesian network based dynamic operational risk assessment. Journal of Loss Prevention in the Process Industries, 41, 399-410.
29. Bechhoefer, E. R. and J. T. Taylor (2011). Data acquisition system for condition-based maintenance, Google Patents.
30. Behmanesh, I., B. Moaveni, G. Lombaert and C. Papadimitriou (2015). Hierarchical Bayesian model updating for structural identification. Mechanical Systems and Signal Processing 64: 360-376.

31. Benghanem, M. (2009). Measurement of meteorological data based on wireless data acquisition system monitoring. Applied energy 86(12): 2651-2660.
32. Bhandari, J., Arzaghi, E., Abbassi, R., Garaniya, V., & Khan, F. (2016). Dynamic risk ‐ based maintenance for offshore processing facility. Process Safety Progress, 35(4), 399-406.
33. Bin, G., J. Gao, X. Li and B. Dhillon (2012). Early fault diagnosis of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network. Mechanical Systems and Signal Processing 27: 696-711.
34. Box, G. E., G. M. Jenkins, G. C. Reinsel and G. M. Ljung (2015). Time series analysis: forecasting and control, John Wiley & Sons.
35. Breslow, N. E. and D. G. Clayton (1993). Approximate inference in generalized linear mixed models. Journal of the American statistical Association 88(421): 9-25.
36. Carden, E. P. and J. M. Brownjohn (2008). ARMA modelled time-series classification for structural health monitoring of civil infrastructure. Mechanical systems and signal processing 22(2): 295-314.
37. Chang, Y.-P. (2001). Estimation of parameters for nonhomogeneous Poisson process: Software reliability with change-point model. Communications in Statistics-Simulation and Computation 30(3): 623-635.
38. Cheng, R., & Prabhakar, S. (2003). Managing uncertainty in sensor database. ACM SIGMOD Record, 32(4), 41-46.
39. Chetouani, Y. (2014). A sequential probability ratio test (SPRT) to detect changes and process safety monitoring. Process Safety and Environmental Protection 92(3): 206-214.
40. Chitsazan, N., A. A. Nadiri and F. T.-C. Tsai (2015). Prediction and structural uncertainty analyses of artificial neural networks using hierarchical Bayesian model averaging. Journal of Hydrology 528: 52-62.
41. Choi, W., H. Kikumoto, R. Choudhary and R. Ooka (2018). Bayesian inference for thermal response test parameter estimation and uncertainty assessment. Applied Energy 209: 306-321.
42. Collacott, R. (2012). Mechanical fault diagnosis and condition monitoring, Springer Science & Business Media.

43. Crow, L. H. (1975). Reliability analysis for complex, repairable systems, ARMY MATERIEL SYSTEMS ANALYSIS ACTIVITYABERDEEN PROVING GROUND MD002E
44. Dar, M. A., D. Gowsami and A. Chaturvedi (2015). Testing Effort Dependent Software Reliability Growth Model with Dynamic Faults for Debugging Process. International Journal of Computer Applications 113(11).
45. Dawotola, A. W., Trafalis, T. B., Mustaffa, Z., Van Gelder, P. H. A. J. M., & Vrijling, J. K. (2012). Risk-based maintenance of a cross-country petroleum pipeline system. Journal of pipeline systems engineering and practice, 4(3), 141-148.
46. De Rademaeker, E., Suter, G., Pasman, H. J., & Fabiano, B. (2014). A review of the past, present and future of the European loss prevention and safety promotion in the process industries. Process Safety and Environmental Protection, 92(4), 280-291.
47. Deloux, E., B. Castanier and C. Bérenguer (2009). Predictive maintenance policy for a gradually deteriorating system subject to stress. Reliability Engineering & System Safety 94(2): 418-431.
48. Dewey, H. H., DeVries, D. R., & Hyde, S. R. (2019). Uncertainty Quantification in Prognostic Health Management Systems. Paper presented at the 2019 IEEE Aerospace Conference.
49. El-Gheriani, M., F. Khan, D. Chen and R. Abbassi (2017). Major accident modelling using spare data. Process Safety and Environmental Protection 106: 52-59.
50. Elkin, P. L., D. R. Schlegel, M. Anderson, J. Komm, G. Ficheur and L. Bisson (2018). Artificial Intelligence: Bayesian versus Heuristic Method for Diagnostic Decision Support. Applied clinical informatics 9(02): 432-439.
51. Fast, M. and T. Palme (2010). Application of artificial neural networks to the condition monitoring and diagnosis of a combined heat and power plant. Energy 35(2): 1114-1120.
52. Fouladirad, M., A. Grall and L. Dieulle (2008). On the use of on-line detection for maintenance of gradually deteriorating systems. Reliability Engineering & System Safety 93(12): 1814-1820.
53. Friis-Hansen, A., Friis-Hansen, P., & Pedersen, P. T. (2001). Bayesian networks as a decision support tool in marine applications. Technical University of Denmark (DTU).

54. Gao, W., Wu, D., Song, C., Tin-Loi, F., & Li, X. (2011). Hybrid probabilistic interval analysis of bar structures with uncertainty using a mixed perturbation Monte-Carlo method. Finite Elements in Analysis and Design, 47(7), 643-652.
55. Giurgiutiu, V., A. Zagrai and J. Jing Bao (2002). Piezoelectric wafer embedded active sensors for aging aircraft structural health monitoring. Structural Health Monitoring 1(1): 41-61.
56. Gonzalez-Bustamante, J., J. Sala, L. López-González, J. Miguez and I. Flores (2007). Modelling and dynamic simulation of processes with 'MATLAB'. An application of a natural gas installation in a power plant. Energy 32(7): 1271-1282.
57. Gowid, S., R. Dixon and S. Ghani (2015). A novel robust automated FFT-based segmentation and features selection algorithm for acoustic emission condition based monitoring systems. Applied Acoustics 88: 66-74.
58. Guisan, A., T. C. Edwards Jr and T. Hastie (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological modelling 157(2-3): 89-100.
59. Guk, E., J.-S. Kim, M. Ranaweera, V. Venkatesan and L. Jackson (2018). In-situ monitoring of temperature distribution in operating solid oxide fuel cell cathode using proprietary sensory techniques versus commercial thermocouples. Applied Energy 230: 551-562.
60. Hajati, T., C. Langenbruch and S. Shapiro (2015). A statistical model for seismic hazard assessment of hydraulic‐fracturing‐induced seismicity. Geophysical Research Letters 42(24): 10,601-610,606.
61. Han, D., S. An and P. Shi (2016). Multi-frequency weak signal detection based on wavelet transform and parameter compensation band-pass multi-stable stochastic resonance. Mechanical Systems and Signal Processing 70: 995-1010.
62. Han, Z. Y., & Weng, W. G. (2011). Comparison study on qualitative and quantitative risk assessment methods for urban natural gas pipeline network. Journal of hazardous materials, 189(1-2), 509-518.
63. Hanski, J., & Ojanen, V. (2019). Sustainability in strategic asset management frameworks: A systematic literature review. International Journal of Strategic Engineering Asset Management.

64. Hameed, Z., Y. Hong, Y. Cho, S. Ahn and C. Song (2009). Condition monitoring and fault detection of wind turbines and related algorithms: A review. Renewable and Sustainable energy reviews 13(1): 1-39.
65. Heaton, J., Parlikad, A. K., & Schooling, J. (2019). Design and development of BIM models to support operations and maintenance. Computers in Industry, 111, 172-186.
66. Heng, A., A. C. Tan, J. Mathew, N. Montgomery, D. Banjevic and A. K. Jardine (2009). Intelligent condition-based prediction of machinery reliability. Mechanical Systems and Signal Processing 23(5): 1600-1614.
67. Hermans, T., F. Nguyen, M. Klepikova, A. Dassargues and J. Caers (2018). Uncertainty Quantification of Medium‐Term Heat Storage From Short‐Term Geophysical Experiments Using Bayesian Evidential Learning. Water Resources Research 54(4): 2931-2948.
68. Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung and H. H. Liu (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, The Royal Society.
69. Iqbal, H., Tesfamariam, S., Haider, H., & Sadiq, R. (2017). Inspection and maintenance of oil & gas pipelines: a review of policies. Structure and Infrastructure Engineering, 13(6), 794-815.
70. Jardine, A. K., D. Lin and D. Banjevic (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical systems and signal processing 20(7): 1483-1510.
71. Jo, Y. D., & Ahn, B. J. (2005). A method of quantitative risk assessment for transmission pipeline carrying natural gas. Journal of hazardous materials, 123(1-3), 1-12.
72. Kabir, G., Sadiq, R., & Tesfamariam, S. (2016). A fuzzy Bayesian belief network for safety assessment of oil and gas pipelines. Structure and Infrastructure Engineering, 12(8), 874-889.
73. Khakzad, N., Khan, F., & Amyotte, P. (2013). Quantitative risk analysis of offshore drilling operations: A Bayesian approach. Safety science, 57, 108-117.

74. Khan, F. I., Haddara, M. M., & Bhattacharya, S. K. (2006). Risk‐based integrity and inspection modeling (RBIIM) of process components/system. Risk Analysis: An International Journal, 26(1), 203-221.
75. Khan, F. I., Sadiq, R., & Haddara, M. M. (2004). Risk-based inspection and maintenance (RBIM): multi-attribute decision-making with aggregative risk analysis. Process safety and environmental protection, 82(6), 398-411.
76. Khan, F. I., & Haddara, M. (2004). Risk‐based maintenance (RBM): A new approach for process plant inspection and maintenance. Process safety progress, 23(4), 252-265.
77. Kamphuis, J. W. (2010). Introduction to coastal engineering and management (Vol. 30). World Scientific.
78. Kans, M. (2008). An approach for determining the requirements of computerised maintenance management systems. Computers in Industry 59(1): 32-40.
79. Kastner, G. (2018). Sparse Bayesian time-varying covariance estimation in many dimensions. Journal of Econometrics.
80. Kelly, D. L. and C. L. Smith (2009). Bayesian inference in probabilistic risk assessment—the current state of the art. Reliability Engineering & System Safety 94(2): 628-643.
81. Kelly, D. and C. Smith (2011). Bayesian inference for probabilistic risk assessment: A practitioner's guidebook, Springer Science & Business Media.
82. Kruger, U. and G. Dimitriadis (2008). Diagnosis of process faults in chemical systems using a local partial least squares approach. AIChE Journal 54(10): 2581-2596.
83. Kumar, N. and S. Chakraborti (2015). Improved phase I control charts for monitoring times between events. Quality and Reliability Engineering International 31(4): 659-668.
84. Lei, Y., J. Lin, Z. He and M. J. Zuo (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mechanical systems and signal processing 35(1-2): 108-126.
85. Leoni, L., A. BahooToroody, F. De Carlo and N. Paltrinieri (2019). Developing a risk-based maintenance model for a Natural Gas Regulating and Metering Station using Bayesian Network. Journal of Loss Prevention in the Process Industries 57: 17-24.
86. Li, B. and M. D. Pandey (2017). An advanced statistical method to analyze condition monitoring data collected from nuclear plant systems. Nuclear Engineering and Design 323: 133-141.

87. Li, Q., F. Guo, S. G. Klauer and B. G. Simons-Morton (2017). Evaluation of risk change-point for novice teenage drivers. Accident Analysis & Prevention 108: 139-146.
88. Liang, Z., Liu, B., Xie, M., & Parlikad, A. K. (2019). Condition-based maintenance for long-life assets with exposure to operational and environmental risks. International Journal of Production Economics.
89. Liu, B., Liang, Z., Parlikad, A. K., Xie, M., & Kuo, W. (2020). Condition-based maintenance for systems with aging and cumulative damage based on proportional hazards model Value Based and Intelligent Asset Management (pp. 211-231): Springer.
90. Liu, B., S. Riemenschneider and Y. Xu (2006). Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum. Mechanical Systems and Signal Processing 20(3): 718-734.
91. Liu, J., R. Seraoui, V. Vitelli and E. Zio (2013). Nuclear power plant components condition monitoring by probabilistic support vector machine. Annals of Nuclear Energy 56: 23-33.
92. Louit, D. M., R. Pascual and A. K. Jardine (2009). A practical procedure for the selection of time-to-failure models based on the assessment of trends in maintenance data. Reliability Engineering & System Safety 94(10): 1618-1628.
93. Loutas, T., G. Sotiriades, I. Kalaitzoglou and V. Kostopoulos (2009). Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements. Applied Acoustics 70(9): 1148-1159.
94. Lunn, D., D. Spiegelhalter, A. Thomas and N. Best (2009). The BUGS project: evolution, critique and future directions. Statistics in medicine 28(25): 3049-3067.
95. Luque, L. and D. Straub (2013). Algorithms for optimal risk-based planning of inspections using influence diagrams. 11th IWP.
96. Mahadevan, S. and R. Rebba (2005). Validation of reliability computational models using Bayes networks. Reliability Engineering & System Safety 87(2): 223-232.
97. Majeske, K. D. (2007). A non-homogeneous Poisson process predictive model for automobile warranty claims. Reliability Engineering & System Safety 92(2): 243-251.
98. Mannan, Sam, ed. Lees' (2012). Loss prevention in the process industries: Hazard identification, assessment and control. Butterworth-Heinemann.
99. Márquez, A. C., Macchi, M., & Parlikad, A. K. Value Based and Intelligent Asset Management.

100. Miao, Q., X. Zhang, Z. Liu and H. Zhang (2017). Condition multi-classification and evaluation of system degradation process using an improved support vector machine. Microelectronics Reliability 75: 223-232.

101. Mishra, M., J. Martinsson, M. Rantatalo and K. Goebel (2018). Bayesian hierarchical model-based prognostics for lithium-ion batteries. Reliability Engineering & System Safety 172: 25-35.

102. Mitchell, K., S. Sana, V. S. Balakrishnan, V. S. Rao and H. J. Pottinger (1999). Microsensors for health monitoring of smart structures. Smart Structures and Materials 1999: Smart Electronics and MEMS, International Society for Optics and Photonics.

103. Modarres, M., M. P. Kaminskiy and V. Krivtsov (2016). Reliability engineering and risk analysis: a practical guide, CRC press.

104. Montiel, H., J. A. Vilchez, J. Arnaldos and J. Casal (1996). Historical analysis of accidents in the transportation of natural gas. Journal of Hazardous Materials 51(1-3): 77-92.

105. Moore, W. and A. Starr (2006). An intelligent maintenance system for continuous cost-based prioritisation of maintenance activities. Computers in industry 57(6): 595-606.

106. Moubray, J., & Lanthier, J. (1991). reliability centered maintenance ii. Butterworth-Heinemann.

107. Myung, I. J. (2003). Tutorial on maximum likelihood estimation. Journal of mathematical Psychology, 47(1), 90-100.

108. Nair, K. K., A. S. Kiremidjian and K. H. Law (2006). Time series-based damage detection and localization algorithm with application to the ASCE benchmark structure. Journal of Sound and Vibration 291(1-2): 349-368.

109. Nandi, S., H. A. Toliyat and X. Li (2005). Condition monitoring and fault diagnosis of electrical motors—A review. IEEE transactions on energy conversion 20(4): 719-729.

110. Neapolitan, R. E. (2004). Learning Bayesian Networks, vol. 38 Pearson Prentice Hall. Upper Saddle River, NJ.

111. Neil, M., M. Tailor, D. Marquez, N. Fenton and P. Hearty (2008). Modelling dependable systems using hybrid Bayesian networks. Reliability Engineering & System Safety 93(7): 933-939.

112. Nielsen, J. S. and J. D. Sørensen (2017). Bayesian estimation of remaining useful life for wind turbine blades. Energies 10(5): 664.

113. Nielsen, J. J., & Sørensen, J. D. (2011). On risk-based operation and maintenance of offshore wind turbine components. Reliability Engineering & System Safety, 96(1), 218-229.

114. Nielsen, T. D., & Jensen, F. V. (2009). Bayesian networks and decision graphs. Springer Science & Business Media.

115. Nikula, R.-P., M. Ruusunen and K. Leiviskä (2016). Data-driven framework for boiler performance monitoring. Applied energy 183: 1374-1388.

116. Paltrinieri, N. and F. Khan (2016). Dynamic Risk Analysis in the Chemical and Petroleum Industry: Evolution and Interaction with Parallel Disciplines in the Perspective of Industrial Application, Butterworth-Heinemann.

117. Pasman, H. J. (2015). Risk Analysis and Control for Industrial Processes-Gas, Oil and Chemicals: A System Perspective for Assessing and Avoiding Low-Probability, High-Consequence Events. Butterworth-Heinemann.

118. Peng, W., H.-Z. Huang, Y. Li, M. J. Zuo and M. Xie (2013). Life cycle reliability assessment of new products—A Bayesian model updating approach. Reliability Engineering & System Safety 112: 109-119.

119. Peng, W., L. Shen, Y. Shen and Q. Sun (2018). Reliability analysis of repairable systems with recurrent misuse-induced failures and normal-operation failures. Reliability Engineering & System Safety 171: 87-98.

120. Pesinis, K. and K. F. Tee (2017). Statistical model and structural reliability analysis for onshore gas transmission pipelines. Engineering Failure Analysis 82: 1-15.

121. Pham, H. T. and B.-S. Yang (2010). Estimation and forecasting of machine health condition using ARMA/GARCH model. Mechanical Systems and Signal Processing 24(2): 546-558.

122. Quy, N., J. Vrijling, P. v. Gelder and R. Groenveld (2006). On the assessment of ship grounding risk in restricted channels. 8th international conference on marine sciences and technologies.

123. Quy, N., J. Vrijling and P. Van Gelder (2008). Risk-and simulation-based optimization of channel depths: Entrance channel of Cam Pha Coal Port. Simulation 84(1): 41-55.

124. Raouf, A., Duffuaa, S., Ben‐Daya, M., Tsang, A. H., Yeung, W., Jardine, A. K., & Leung, B. P. (2006). Data management for CBM optimization. Journal of quality in maintenance engineering.

125. Rasmekomen, N., & Parlikad, A. K. (2016). Condition-based maintenance of multi-component systems with degradation state-rate interactions. Reliability Engineering & System Safety, 148, 1-10.

126. Rebba, R. and S. Mahadevan (2008). Computational methods for model reliability assessment. Reliability Engineering & System Safety 93(8): 1197-1207.

127. Rief, H. (1984). Generalized Monte Carlo perturbation algorithms for correlated sampling and a second-order Taylor series approach. Annals of nuclear energy, 11(9), 455-476.

128. Roffel, B. and B. Betlem (2007). Process dynamics and control: modeling for control and prediction, John Wiley & Sons.

129. Roos, H. G., Stynes, M., & Tobiska, L. (2008). Robust numerical methods for singularly perturbed differential equations: convection-diffusion-reaction and flow problems (Vol. 24). Springer Science & Business Media.

130. Ross, S. M. (2014). Introduction to probability models, Academic press.

131. Sheu, S.-H., T.-H. Liu, Z.-G. Zhang and H.-N. Tsai (2018). The generalized age maintenance policies with random working times. Reliability Engineering & System Safety 169: 503-514.

132. Sinha, B. (2002). Trend prediction from steam turbine responses of vibration and eccentricity. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy 216(1): 97-103.

133. Siu, N. O. and D. L. Kelly (1998). Bayesian parameter estimation in probabilistic risk assessment1. Reliability Engineering & System Safety 62(1-2): 89-116.

134. Slimacek, V. and B. H. Lindqvist (2016). Nonhomogeneous Poisson process with nonparametric frailty. Reliability Engineering & System Safety 149: 14-23.

135. Spiegelhalter, D., A. Thomas, N. Best and D. Lunn (2007). OpenBUGS user manual, version 3.0. 2. MRC Biostatistics Unit, Cambridge.

136. Srinivasan, R., & Parlikad, A. K. (2013). Value of condition monitoring in infrastructure maintenance. Computers & Industrial Engineering, 66(2), 233-241.

137. Srinivasan, R., & Parlikad, A. K. (2020). An approach to value-based infrastructure asset management Value Based and Intelligent Asset Management (pp. 123-138): Springer.

138. Straub, D. (2009). Stochastic modeling of deterioration processes through dynamic Bayesian networks. Journal of Engineering Mechanics 135(10): 1089-1099.

139. Straub, D. (2004). Generic approaches to risk based inspection planning for steel structures (Vol. 284). vdf Hochschulverlag AG.

140. Straub, D. and A. Der Kiureghian (2010). Bayesian network enhanced with structural reliability methods: methodology. Journal of engineering mechanics 136(10): 1248-1258.

141. Tahan, M., E. Tsoutsanis, M. Muhammad and Z. A. Karim (2017). Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review. Applied energy 198: 122-144.

142. Tinsley, H. E. and S. D. Brown (2000). Multivariate statistics and mathematical modeling. Handbook of applied multivariate statistics and mathematical modeling, Elsevier: 3-36.

143. Tixier, J., Dusserre, G., Salvi, O., & Gaston, D. (2002). Review of 62 risk analysis methodologies of industrial plants. Journal of Loss Prevention in the process industries, 15(4), 291-303.

144. Too, E. G., Betts, M., & Arun, K. (2006). A strategic approach to Infrastructure Asset Management.

145. Toroody, A. B., M. M. Abaiee, R. Gholamnia and M. J. Ketabdari (2016a). Epistemic-based investigation of the probability of hazard scenarios using Bayesian network for the lifting operation of floating objects. Journal of Marine Science and Application 15(3): 250-259.

146. Toroody, A. B., M. M. Abaiee, R. Gholamnia, M. B. Torody and N. H. Nejad (2016b). Developing a risk-based approach for optimizing human reliability assessment in an offshore operation. Open Journal of Safety Science and Technology 6(01): 25-32.

147. UNI (2009a), 'Impianti di ricezione, prima riduzione e misura del gas naturale. progettazione,costruzione,collaudo', UNI 9167 ; 2009.

148. UNI (2009b), 'UNI EN 334 2009_Regolatori di pressione del gas per pressioni di entrata fino a 100 bar ', UNI EN 334 ; 2009.

149. UNI (2012), 'infrustrutture del gas Condotte con pressione massima di esercizio non maggiore di 16 bar, part 1 ', UNI EN 12007-1.

150. Usher, J. S., Kamal, A. H., & Syed, W. H. (1998). Cost optimal preventive maintenance and replacement scheduling. IIE transactions, 30(12), 1121-1128.

151. Valdés, R. M. A., V. F. G. Comendador, L. P. Sanz and A. R. Sanz (2018). Prediction of aircraft safety incidents using Bayesian inference and hierarchical structures. Safety science 104: 216-230.

152. Van Horenbeek, A., Pintelon, L., & Muchiri, P. (2010). Maintenance optimization models and criteria. International Journal of System Assurance Engineering and Management, 1(3), 189-200.

153. Wagenmakers, E.-J., M. Marsman, T. Jamil, A. Ly, J. Verhagen, J. Love, R. Selker, Q. F. Gronau, M. Šmíra and S. Epskamp (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. Psychonomic bulletin & review 25(1): 35-57.

154. Wang, X., U. Kruger and B. Lennox (2003). Recursive partial least squares algorithms for monitoring complex industrial processes. Control Engineering Practice 11(6): 613-632.

155. Weidl, G., A. L. Madsen and S. Israelson (2005). Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. Computers & chemical engineering 29(9): 1996-2009.

156. Wu, Z. and N. E. Huang (2004). A study of the characteristics of white noise using the empirical mode decomposition method. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, The Royal Society.

157. Wu, Z., N. E. Huang, S. R. Long and C.-K. Peng (2007). On the trend, detrending, and variability of nonlinear and nonstationary time series. Proceedings of the National Academy of Sciences 104(38): 14889-14894.

158. Xin, P., F. Khan and S. Ahmed (2017). Dynamic hazard identification and scenario mapping using Bayesian network. Process Safety and Environmental Protection 105: 143-155.

159. Yam, R., P. Tse, L. Li and P. Tu (2001). Intelligent predictive decision support system for condition-based maintenance. The International Journal of Advanced Manufacturing Technology 17(5): 383-391.

160. Yang, M., F. I. Khan and L. Lye (2013). Precursor-based hierarchical Bayesian approach for rare event frequency estimation: a case of oil spill accidents. Process safety and environmental protection 91(5): 333-342.

161. Yu, H., F. Khan and B. Veitch (2017). A flexible hierarchical Bayesian modeling technique for risk analysis of major accidents. Risk analysis 37(9): 1668-1682.

162. Zarei, E., M. Yazdi, R. Abbassi and F. Khan (2018). A hybrid model for human factor analysis in process accidents: FBN-HFACS. Journal of Loss Prevention in the Process Industries.

163. Zarei, E., Azadeh, A., Khakzad, N., Aliabadi, M. M., & Mohammadfam, I. (2017). Dynamic safety assessment of natural gas stations using Bayesian network. Journal of hazardous materials, 321, 830-840.

164. Zeger, S. L. and M. R. Karim (1991). Generalized linear models with random effects; a Gibbs sampling approach. Journal of the American statistical association 86(413): 79-86.

165. Zhang, J., P. Wang, R. X. Gao and R. Yan (2018). An Image Processing Approach to Machine Fault Diagnosis Based on Visual Words Representation. Procedia Manufacturing 19: 42-49.

166. Zhang, S., Ma, L., Sun, Y., & Mathew, J. (2007). Asset health reliability estimation based on condition data.

167. Zhou, S., B. Sun and J. Shi (2006). An SPC monitoring system for cycle-based waveform signals using Haar transform. IEEE Transactions on Automation Science and Engineering 3(1): 60-72.

168. Žvokelj, M., S. Zupan and I. Prebil (2010). Multivariate and multiscale monitoring of large-size low-speed bearings using ensemble empirical mode decomposition method

combined with principal component analysis. Mechanical Systems and Signal Processing 24(4): 1049-1067.