Prognostics and Health Management for Multi-Component Systems

Roy Assaf
University of Salford

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

School of Computing, Science & Engineering
Autonomous Systems and Robotics

February 2018

Contents

List of Publications

The content of this thesis builds on and extends the work presented in the following publications by the author:

Journal Papers

- Roy Assaf, Phuc Do Van, Samia Nefti-Meziani, and Philip Scarf. ”Wear rate-state interactions within a multi-component system: a study of a gearbox accelerated life testing platform.” Sage Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability. **In press**

- Phuc Do, Roy Assaf, Phil Scarf and Benoit Iung. ”Condition-based maintenance for a two-component system with dependencies with application to a Gearbox platform.” Submitted to Elsevier Journal of Reliability Engineering and System Safety. **In revision**

- Roy Assaf, Phuc Do Van, Philip Scarf, and Samia Nefti-Meziani. ”Wear rate-state interaction modelling for a multi-component system: Models and an experimental platform.” IFAC-PapersOnLine, 49(28) 232-237. (11)

Papers in Peer-Reviewed Conferences

- Roy Assaf, Phuc Do Van, Philip Scarf, and Samia Nefti-Meziani. ”Diagnosis for systems with multi-component wear interactions.” In IEEE International Conference on Prognostics and Health Management (PHM), 2017. (12) **Received best paper award**

- Roy Assaf, Samia Nefti-Meziani, and Philip Scarf. ”Unsupervised learning for improving fault detection in complex systems.” In IEEE International Conference on Advanced Intelligent Mechatronics (AIM), 2017. (10)

Other publications by the author outside the scope of this thesis

- Stefania Russo, Roy Assaf, and Samia Nefti-Meziani. ”Towards a practical Implementation of EIT-based Sensors using Artificial Neural Networks.” In IEEE SENSORS, 2017. (210)

List of Figures

List of Tables

List of Algorithms

Acknowledgements

I want to express my gratitude to my supervisors, Prof. Phil Scarf and Prof. Samia-Nefti Meziani for giving me the chance to be part of a great European project over the last three years. I want to especially thank Phil for his friendship, trust and support especially during the final stages of my PhD. Also, I very much appreciate the help and support of Dr. Phuc Do who has contributed greatly to the guidance of this thesis.

I appreciate the continued support of my close friends and colleagues Dr. Theo Theodoridis, Saber Mahbouboubi, Alexandre Bousaid and Stefania Russo. Theo has been a great mentor and has helped on many occasions. And Stefania has been constantly loving, kind and supportive. This has played an important role motivating me and will continue to do so in the future.

I would also like to thank my family for always being therefor me, and for their love: Sami, Rita, Cynthia and Cindy. I am also grateful to the “Society of Friends of the White Knowledge“ for showing me the path to esoteric science and self knowledge.

Finally I would like to mention that the research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/ under REA grant agreement number 608022.

Abbreviations and Notations

	Abbreviations
AI	Artificial Intelligence
ALT	Accelerated Life Testing Platform
AM	Amplitude Modulation
ANN	Artificial Neural Networks
AR	Auto Regressive Model
ARMA	Auto-Regressive Moving Average
CAD	Computer Aided Design
CBM	Condition Based Maintenance
CNN	Convolutional Neural Network
DAQ	Data Acquisition System
DFT	discrete Fourier transform
DL	Deep Learning
DNN	Deep Neural Networks
DWT	discrete wavelet transform
EM	Expectation Maximisation
EMD	Empirical Mode Decomposition
EM	expectation maximisation
FCM	Fuzzy C-means
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FRMS	Frequency Root Mean Square
FT	Fourier transform

G	Acceleration of Gravity
GMM	Gaussian Mixture Model
HHT	Hilbert-Huang Transform
HMM	Hidden Markov Models
HSHL	Low Speed Low Load cycle
HSMM	Hidden Semi-Markov Model
HT	Hilbert transform
ICA	Independent Component Analysis
ICPHM	International Conference on Prognostics and Health Management
IFFT	Inverse Fast Fourier Transform
IIR	Infinite Impulse Response
IJPHM	international Journal of Prognostics and Health Management
IOT	Internet Of Things
KF	Kalman Filter
LPP	Locality Preserving Projections
LSLL	High Speed High Load cycle
MA	Moving Average
MAD	Median Absolute Deviation
MDF	Medium Density Foam
MED	Median Value
ML	Machine learning
MLE	maximum likelihood estimation
OEM	Original Equipment Manufacturers
PDF	Probability Density Function
PF	Particle Filter
PF	Particle Filter
PHM	Prognostics Health Management
PCA	Principal Component Analysis
RBF	Radial Basis Function

RMS	root mean square
RPM	Revolutions Per Minute
RUL	Remaining Useful Lifetime
SDK	Software Development Kit
SMC	Sequential Monte Carlo
SNR	Signal-to-Noise Ratio
SOM	Self Organising Maps
STFT	short time Fourier transform
SVM	Support Vector Machine
TBM	Time Based Maintenance
TSA	Time Synchronous Average
UKF	Unscented Kalman Filter
UPF	Unscented Particle Filter
WPD	Wavelet Packet Decomposition
WT	Wavelet Transform

	Notations
a	cost-saving factor for joint replacement (when components 1 and 2 are replaced together)
b	duration-saving factor for joint replacement
C	class or cluster
C^{∞}	long-run expected maintenance cost per time unit (cost-rate)
$CS_{-,-}$	cost-saving of joint replacement
C_I	cost of an inspection
C_p^i, C_c^i	cost of preventive and corrective replacement of component i respectively
c_d	downtime cost-rate of the system
d_i	duration of a replacement for component i
n_c	Number of components of a multi-component system

n_p	Number of particles of a particle filter
T_i	the time of the ith inspection of the system
ΔT	inter-inspection interval
$x^i_{T_k}$	state (degradation level) of component i at time T_k
L^i	failure threshold of component i
m^i_p, m^i_o	preventive and opportunistic maintenance thresholds for component i
w	A frequency range
w_c	filter cut-off frequency
X^i_t	Degradation state of a component i at time t
X^i_t	Degradation influence of component j on a component i
α^i, β^i	shape and rate parameter of gamma distribution associated with component i

Abstract

The ever increasing number of manufacturing requirements is pushing original equipment manufacturers (OEM) to design more complex systems to meet industrial needs. These systems are being fitted with more components which bear stochastic and economic dependencies. Therefore maintaining such systems is becoming more and more of a challenge, especially due to their degradation processes becoming highly stochastic in nature.

This thesis is concerned with the prognostics and health management (PHM) of such complex multi-component systems, whereby signal processing and health indicator extraction, diagnostics, prognostics and maintenance decision making in light of present stochastic and economic dependencies are considered.

We introduce several novel approaches for dealing with systems that have multiple components. We first introduce a gearbox accelerated life testing platform that was designed with the objective of gathering experimental data for multi-component degradation models, for the reason that multi-component systems with inter-dependencies follow a highly stochastic degradation process which depends to an extent on their complex mechanical design. We then present our methodology for extracting accurate health indicators from multi-component systems by means of a time-frequency domain analysis. This sets the stage for degradation modelling, and so we show the development of a generic degradation model in which the degradation process of a component may be dependent on the operating conditions, the component's own state, and the state of the other components. We then show how to fit the models to data using particle filter. This method is then used for the data generated by the gearbox. Afterwards a diagnostic procedure is presented and uses Gaussian mixture models. This is used to uncover accelerated wear processes that take place when old worn out components are coupled with new healthy components. Finally economic dependency is considered where combining multiple maintenance activities has lower cost than performing maintenance on components separately. To select a component or components to be preventively maintained, adaptive preventive maintenance and opportunistic maintenance rules are proposed. A cost model is developed to find the optimal values of decision variables.

In our work, we find that stochastic dependencies between components lead to accelerated degradation which causes unexpected faults and failures, and consequent economic losses.

Although this work deals with stochastic dependence between components, it involves some engineering knowledge of the systems under study, and this makes application of the models on a large scale challenging to automate. Therefore, we make recommendation for future research that includes the development of end-to-end learning techniques such as deep learning. In doing so we can potentially use the time wave data and automatically extract the most relevant features for doing accurate prognostics, and therefore health management, of such systems.

The research work in this thesis was motivated by the problems faced by industrial partners such as the world leading food system manufacturing company Marel in the Netherlands, which were part of the sustainable manufacturing and advanced robotics training network in Europe (SMART-e).

Chapter 1

Introduction

Contents

1.1 Motivation

From a physics point of view, and especially considering the second law of thermodynamics, we know that the sum of the entropy of interacting thermodynamic systems has to increase with time. Therefore in a general sense, systems that have interactions in the likes of transfer of energy, heat and work, will eventually degrade. Unfortunately all our machinery, be it mechanical or electronic, fall under the category of such systems, thus their degradation processes are inevitable. However these degradation processes can be slowed or even in some cases stalled through the act of maintenance.

And so naturally one would come to think about automating such interventions. This becomes especially attractive when one considers some self-healing properties that occur in nature, taking for example the self-healing capabilities of biological systems.

In (74), a self-repairing algorithm for self-reconfigurable modular robots is presented. The robot is able to determine which of its modules is faulty, eject it and replace it with a spare module. This gives the robot the capability to detect and recover from failures. A similar approach is presented in (169). First, the work reviews several types of self-repairing systems. Then the authors propose their approach for a self-repairing machine which presents component redundancy, namely, it is made of only one type of unit. The system can change functionality by modifying the local connection of its units and therefore its global shape. When the system detects failure, the faulty units are automatically disconnected and redundant ones are used and moved to restore its original shape and functionality. In (6), the concept of a self-healing vehicle is presented as a vehicle capable of predict and detect faults and perform corrective interventions. The authors propose a preliminary view of how the dynamic system architecture should look like for such a vehicle. In particular, some key capabilities should be: on-board diagnostics and prognostics of vehicle's electronics; using components and architectures that facilitate self healing; and perform diagnostics of mechanical components for monitoring their wearing out processes.

Some self-healing polymers and composites are reviewed in (31), this work discusses the self-healing capabilities of such materials in response to impact, puncture, corrosion and during fatigue. However, although research is being carried on such materials, they are yet to be incorporated as integral parts of industrial machinery components. Also, in a recent study (78), the authors review the currently existing self-healing technologies in software engineering, materials, mechanics and electronics. They state that research on self-healing mechanisms is still considered to be in their infancy when it comes to engineering problems.

Consequently when we refer to the maintenance literature, we see that the recent major advancements in the field fall under the following two categories: condition based maintenance (CBM) (92, 115, 193), which in contrast to older maintenance strategies is proactive in nature, and aims to carry out maintenance interventions only when needed, see Section 2.2 for more detail; and, more recently, prognostics and health management (PHM) (255, 140, 128, 190, 139), which has many similarities with CBM, and which is principally seen as a key enabler for CBM (245).

PHM is an engineering approach which allows system health state assessment in real-time, as well as predicting its future health states. Thus in contrast to CBM, PHM is more concerned with the actual health indicator extraction from the acquired signals, and puts a lot of emphasis on the prognostics step which is essential to performing optimal maintenance decision making. This will be topic considered in this thesis, and will be described in more detailed in Section 2.3.

The key idea here is prognostics, whereby the end of life (EOL) of components is predicted, and consequently the remaining useful lifetime (RUL) extracted, see Eq 1.1

$$RUL_k = t_{eol} - t_k \tag{1.1}$$

where RUL_k represents the remaining useful life at a time t_k, and t_{eol} denotes the predicted end of life.

The three principal elements of PHM can be seen in Figure 1.1. When considering a system to be maintained, PHM starts by performing health indicators extraction for the system at hand. These are then used to perform diagnostics and prognostics. These, in turn, are used to take the best maintenance decision. This is typically done with the objective of reducing maintenance costs or increasing the system reliability.

Often the degradation processes of components in a system are assumed to be independent, see (35, 176, 248). But since real world systems are usually complex and include multiple interacting components, such interactions can potentially affect overall system availability, and jeopardise the effectiveness of PHM and CBM.

In (78), the authors express their interest in investigating the claim that fail-

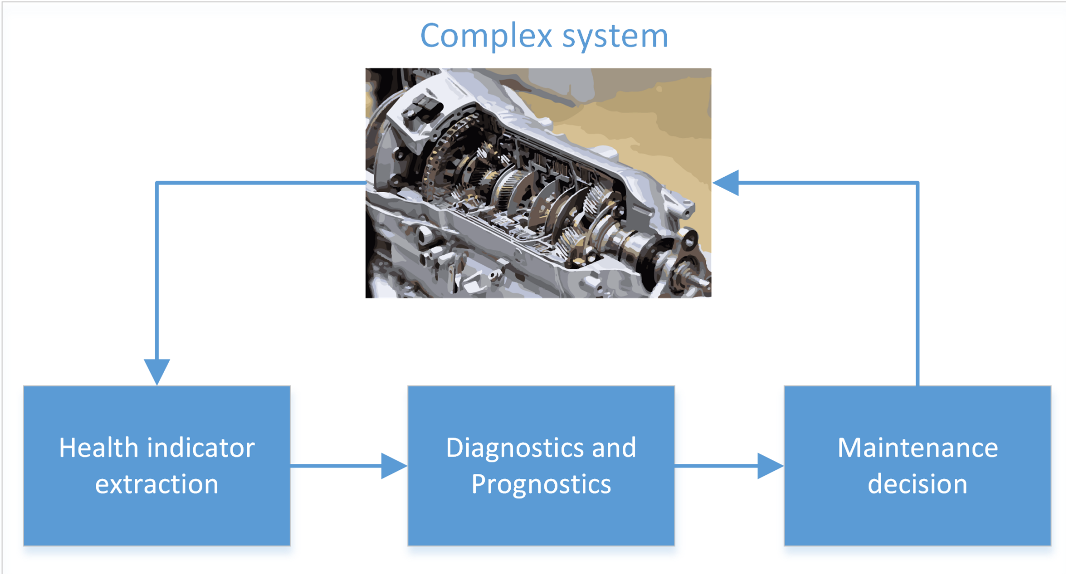

Figure 1.1: Principal elements of prognostics and health management

ures in a system are mostly mutually independent. They explain that it seems more likely that failures are correlated and that failures in some components might lead to failures in others. Also, recent CBM literature has been showing a growing interest in multi-component systems and their dependencies (125). The modelling of stochastic dependence, whereby the health state of some components can be affected by the health states of other components remains the least explored (178). This falls more under the PHM aspect, and literature on the topic are hardly found.

Therefore in this thesis we are concerned with the PHM of complex multi-component systems, and the health indicator extraction, diagnostics, prognostics, maintenance decision making and their sub-elements therein, in the light of the stochastic and economic dependencies.

1.2 Main Contribution

We make a number of contributions to prognostics and health management that deal with multi-component system degradation. These are the development of: an experimental platform; a methodology for health indicator extraction; multi-component degradation models and prognostics; clustering of degradation phases; and maintenance optimisation.

Due to the nature of the case study and data that are considered throughout this thesis, the scope of these contributions mainly covers multi-component rotating machinery which play an essential role in industrial applications (140, 102). The key contributions are presented in further detail next.

1.2.1 Multi-Component Experimental Platform

Although there exists a large number of experimental data-sets on degradation (113, 173), few represent multi-component degradation, whereby data are present for more than one component that are simultaneously degrading. And to our knowledge no experimental platforms have been specifically constructed or data-sets generated in order to study and represent the stochastic dependence between components within a system. Therefore we have designed and developed an experimental platform with the aim of providing such data that can allow us to study the degradation processes of interacting components. This has resulted in more insight about the true nature of such degradation. The design and development of the experimental platform along with its results are presented in Chapter 3. Another configuration of the same platform is presented in Chapter 4 and provides data that can be further analysed in the frequency domain.

1.2.2 Health Indicator Extraction

PHM is dependant on data acquisition and its processing (251). Key information regarding a machine's condition has to be acquired from sensors and then processed. The main approaches for dealing with sensor signals include time waveform analysis, frequency analysis and time-frequency analysis. However in the literature we find that these techniques are mostly applied for single components. Using the case study provided by the experimental platform presented in Section 4.3.2, we show how time-frequency analysis techniques can be adapted to analyse multi-component system signals, and therefore contribute to the PHM body of literature by providing a methodology to deal with multi-component signals in the presence of degradation. This is mainly described in Chapter 4.

1.2.3 Degradation Modelling

Multiple dependencies can exist between the multiple components of a complex system. In the literature we find that these are mainly grouped into three different types: structural dependencies; economic dependencies; and stochastic dependencies (237, 57, 177). Stochastic dependencies are concerned with the degradation or failure interactions between components. In Chapter 5, we present our generic degradation model where the degradation process of a component may be dependent on the operating conditions, the component's own state, and the state of the other components. This will be an integral part of the PHM procedure since it will allow us to model the degradation dependencies between components, and is later used for performing prognostics. This could also be used for simulating system degradation processes that would lead to choosing optimal maintenance policies.

1.2.4 Prognostics

PHM puts a lot of emphasis on the prognostic step, and rightfully so, since all decision regarding maintenance will be based on these predictions. Many approaches can be taken regarding prognostics, and are mainly separated into three categories, physics based, data-driven and hybrid models (128, 116). A particularly successful one is the stochastic filtering method known as particle filter (PF) which has lately gained a lot of attention in the field of prognostics (119). We use this state-of-the-art approach and apply it in a multi-component context using as a state model the degradation model developed in Chapter 5. This allows us to efficiently estimate the parameters of the model and in turn conduct accurate prognostics and EOL predictions.

1.2.5 Pattern Recognition

Although humans are capable of recognising and distinguishing a multitude of patterns that occur naturally, we still face many difficulties when trying to understand and discriminate patterns that occur artificially such as the ones present in signals and data generated by sensors. This becomes more of an overwhelming challenge when we deal with large amounts of data. Recently, there has been a surge in the number of applications that employ machine learning techniques for finding patterns in signals and data (51, 114). These patterns might describe specific situations, phases or events that can then be used to aid humans or machines in taking decisions. In Chapter 6, we demonstrate our diagnosis method. This uses unsupervised learning, particularly Gaussian mixture model (GMM) to distinguish different degradation phases using the case study of the gearbox experimental platform introduced in Chapter 4. Here we note that the proposed approach can be applied to any system consisting of multiple components. Therefore we contribute to the PHM literature by presenting a novel approach for detecting degradation dependency phases within a multi-component system.

1.2.6 Maintenance Optimisation

Economic dependencies arise when it is cost-effective to combine certain maintenance interventions (125). In Chapter 7 we consider this dependence and we describe our proposed maintenance policy and the optimisation process. Then the utility of this proposed maintenance policy is demonstrated on the data resulting from the gearbox experimental platform.

1.3 Thesis Outline

Having highlighted the purpose and scope of the thesis, the next Chapter 2 provides the introductory concepts and a literature review on CBM and related maintenance strategies; PHM; and multi-component system dependencies. The content in Chapters 3,4,5,6 and 7 contains the main body of the thesis, the details of the previously mentioned contributions are herein presented. These chapters begin with a chapter summary and end with a discussion around the results obtained.

Chapter 3 presents a gearbox accelerated life testing platform. This chapter includes the design and development of the platform, and primary results which show stochastic dependence between components. In Chapter 4, we present our methodology for extracting health indicators for multiple components within a system. The different aspects of achieving this are then presented. This is applied to vibration data resulting from the gearbox platform. The results are then shown, and stochastic dependence between components is clearly highlighted. In this chapter we obtain a time series representing the degradation trajectories of the system components which is used in the following chapters. In Chapter 5 we introduce our generic degradation model. This can be used to model stochastic dependence along with other factors that affect system degradation. The model is then fitted to the data generated by the gearbox using PF. This is used for predicting the EOL of components. In Chapter 6 we demonstrate our multi-component diagnostics approach. This is done using GMM and is applied on the data generated by the gearbox platform to show different degradation phases.

In Chapter 7 economic dependence is considered and a maintenance policy is proposed. This policy is studied using simulation obtained from the degradation model. This is followed by an analyses of the results. Finally, Chapter 8 concludes the thesis and discusses the findings and limitations of this work, and proposes future directions for research on multi-component system PHM.

Chapter 2

Background

Contents

2.1 Chapter Summary

This chapter reviews some of the literature, and presents some introductory concepts that are relevant to the work presented in this thesis. It covers an overview on different maintenance strategies; outlines the different aspects of prognostics and health management (PHM); and covers dependencies in systems with multiple components, discussing in more depth the topic of stochastic dependency which is fundamental to this thesis.

Furthermore, this chapter includes an extensive literature review on different prognostics approaches, since this is an integral aspect of PHM.

2.2 Maintenance Strategies

Failures are undesirable and have implications for safety and the environment. Therefore achieving high reliability and availability is important for industry (115). Nonetheless, all machinery will eventually degrade. Furthermore, it is hard to anticipate all the environment variables for these systems once put into operation, therefore a lot of randomness is introduced into their nominal degradation behaviour.

Maintenance is thus used to underpin system reliability, and to prevent unexpected faults and failures that might lead to downtime and economic losses, or even hazardous situations where the safety of the operators or the public can be jeopardised.

Maintenance planning has evolved trough three different stages (164, 115, 202): unplanned breakdown maintenance; planned scheduled maintenance; and condition based maintenance (CBM), this is where PHM is considered. A further illustration of the evolution of maintenance planning can be seen in Figure 2.1.

2.2.1 Unplanned Breakdown Maintenance

Unplanned breakdown maintenance is also known as reactive or corrective maintenance. It is the earliest form of maintenance, and as the name suggests, the machinery is operated until there is breakdown and therefore maintenance is performed after a failure has already occurred. For this reason this strategy is considered passive. And although this approach can have the advantages of a low cost policy and that it requires minimal management (164); these benefits can

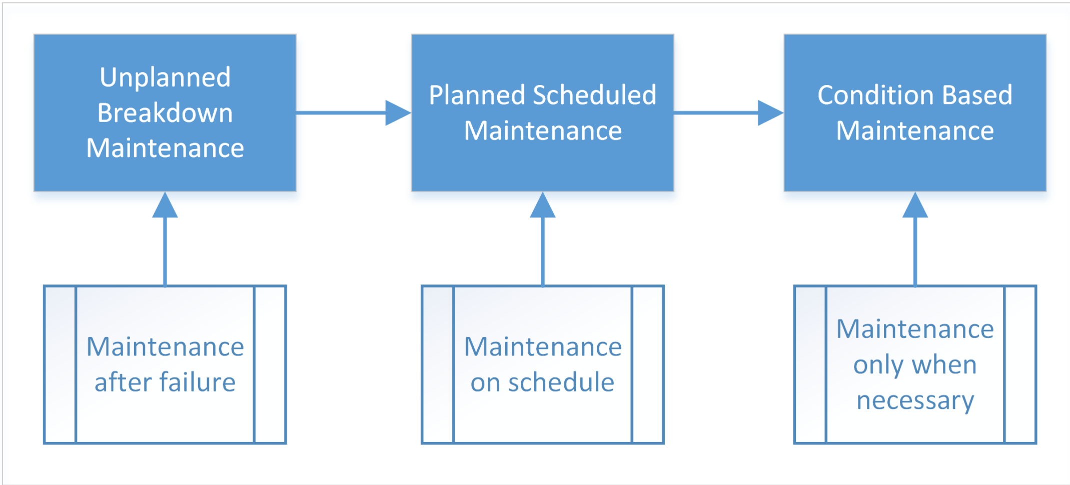

Figure 2.1: Evolution of maintenance in three different stages

be easily outweighed by the disadvantages. For instance it could result in high downtime and therefore great economic losses and cause serious damage to other healthy components of the system. In some cases failures can even cause severe damage to the surrounding environment jeopardising safety.

2.2.2 Planned Scheduled Maintenance

Planned scheduled maintenance is also known as preventive or time based maintenance (TBM). This strategy is an improvement over unplanned breakdown maintenance for degrading systems. The principle is that a machine is maintained periodically regardless of its condition or state. This allows for maintenance resources to be planned, such as spare part and maintenance crew availability. It can also decrease the number of unforeseen faults and failures. However, this strategy does not contribute to maximising the life of an asset, and in many cases leads to over-maintenance since maintenance is carried out regardless of the health state of the components of a system. This can lead to increased costs of maintenance.

2.2.3 Condition Based Maintenance

Condition based maintenance is also known as predictive maintenance. This maintenance strategy monitors the condition of a system and decides if maintenance needs to be done. It has been compared to the previously mentioned strategies in (130), and shown to offer many benefits to industry.

As an example, Figure 2.2 depicts how monitoring equipment using the internet of things (IOT) may impact on costs in industry (73).

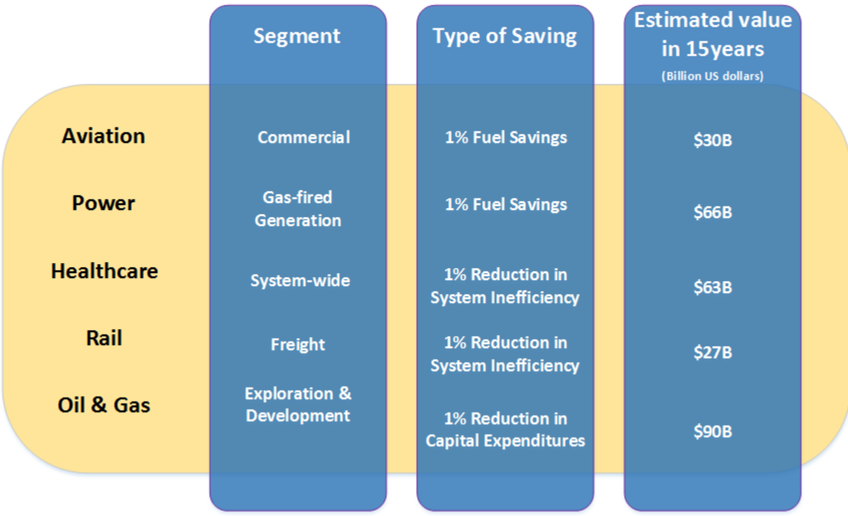

Figure 2.2: General Electric estimation (73) on how a 1% improvement from Industrial Internet across specific global industry sectors could save 276 billion USD

CBM is currently widely accepted as the best maintenance strategy in many cases (202). However, there are still significant challenges for its implementation (2).

Since CBM is based on the idea that maintenance is carried out only when needed, it is then firmly associated with condition monitoring. And so to render this strategy effective, reliable condition monitoring techniques are needed.

Furthermore, in order to optimally make decisions on maintenance and within a timely manner, predictions on the future health states of the system must be carried out. Moreover, in the context of maintenance it is specifically interesting to predict the end of life (EOL) of a component; this is known as prognostics, which is preceded by the step of health indicator extraction and which are the main concern of PHM, and therefore treated extensively in this thesis.

Predicting the end of life EOL allows for the extraction of the remaining useful lifetime (RUL) of the components of a system which then allows for effective decision making. This is further discussed in Chapter 7, and RUL will be discussed in more detail later in this chapter.

2.3 Prognostics and Health Management

PHM is getting substantial attention from the maintenance community recently, see (255, 140, 128, 190, 139). It is seen as a key enabler for CBM (245, 233). According to (244) it can be described as an emerging engineering discipline which studies and associates the degradation processes to system lifecycle management.

The disciplines of PHM and CBM share a lot of similarities. One attempt of aligning these disciplines is described in (240), where first the appropriate monitoring approach is adopted. This is then used to support optimal maintenance decision making through an asset's life cycle.

PHM is more concerned with the actual health indicator extraction from the acquired signals, and puts a lot of emphasis on the prognostics step which is essential to performing optimal maintenance decisions. Therefore, it allows for entirely benefiting from CBM.

The term prognostics and health management was first adopted in the joint strike fighter (JSF) F-35 program (36), whereby the proposed PHM concept, design, and architecture were described. Thus the development of PHM was initially driven by the aerospace and defence industries, as it is historically linked to safety and high maintenance cost issues (245). The U.S. Navy has been implementing the strategy of developing, integrating, and demonstrating diagnostics, prognostics, health monitoring, and estimating RUL for mechanical systems (98). Some weapon platforms such as in the JSF Program (103), are in fact already implemented with prognostics capabilities.

Many technical societies have lately focused on PHM. As mentioned in (128), the two most representative ones are the PHM society which holds yearly conferences on the topic since 2009 (227), and which publishes the international journal of prognostics and health management (IJPHM); And, the IEEE reliability society, which holds the yearly international conference on prognostics and heath management (ICPHM) since 2008 (226), and which is responsible for publishing the IEEE Transactions on reliability.

Recently, a great amount of research has been published on PHM. This shows that it is now being applied to various industries such as aerospace, nuclear, wind power, civil infrastructure, manufacturing and electronics (233, 276, 277).

PHM techniques have mostly advanced in the area of aerospace. Some examples are in detecting shaft unbalance, or gears and bearings degradation as in (284), where the authors work on estimating the RUL of bearings of a helicopter's oil cooler. Or in (245) for predicting cracks in the gear plate of helicopters. In the electronics industry, some failure prognosis of commercial notebook computers by monitoring life cycle temperature data are presented in (249). A review on the application of PHM in industrial electronics can be found in (276). Other applications include methodologies for engine health assessment and prediction as in (257). The manufacturing industry has also seen considerable amount of work in PHM with the goal of minimising downtime (115, 139).

PHM comprises of three main elements. These are: health indicator extraction; diagnostics and prognostics; and maintenance decision making. A flowchart of these three elements along with some of their sub-elements is presented in Figure 2.3. These will be detailed in the following sections.

2.3.1 Data Acquisition

The first required step for implementing PHM is data acquisition. This process can greatly affect the performance of prognostics and decision making. Usually data that are used in PHM come in two forms:

- Event data: information about events such as the installation of new components, faults, failures, etc.
- Condition monitoring data: information about general health condition of the component. These data could be vibration, temperature, pressure, electric currents. etc. usually acquired through sensors.

In the case of condition monitoring data, which are principally used for PHM, various sensors can be used for data collection, usually attached to the component to be monitored. These sensors include accelerometers, acoustic emission sensors, infrared sensors and ultrasonic sensors. An overview of several sensor systems and their performance needs for PHM applications can be found in (49). Also, the parameters to be measured and the electrical and physical attributes are presented. Once data are acquired from the system, they are then transmitted and stored into a workstation through a data acquisition (DAQ) system.

The advancement of new technologies for fast and low-noise DAQ systems, and the recent developments seen on the topic of IOT (15) are making PHM increasingly convenient to implement.

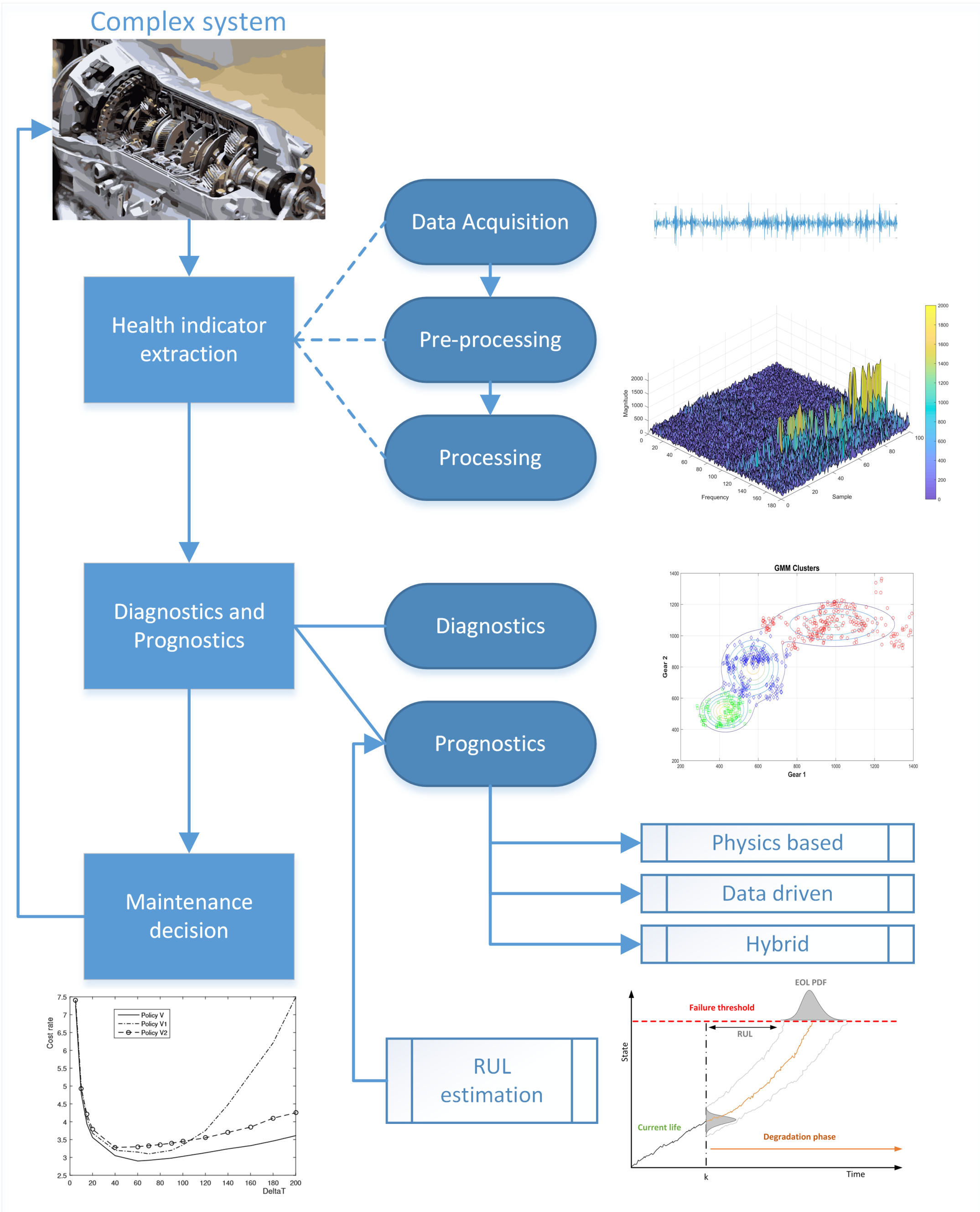

Figure 2.3: Flowchart of the different elements of PHM

2.3.2 Signal Processing

Data processing is the foundation for building reliable models. The data processing goal is to extract useful information from the raw signal data. These data are in fact usually complicated and might carry noise. The main purpose of signal processing is to reduce the noise; enhance the underlying information by understanding the processes that generated the data themselves; and reduce data size to enhance the degradation models' effectiveness. In general, the methods for analysing raw data can be further divided into two steps, they are signal pre-processing and signal processing.

Various signal processing techniques and algorithms have been proposed in the literature. A review on such techniques applied to fault diagnosis of rolling element bearings can be found in (199). These techniques fall mainly under three categories: time domain analysis, frequency–domain analysis, and time frequency domain analysis.

Time domain analysis deals directly on the original time sampled signals. In frequency domain analysis the signal is decomposed into its frequency components. This approach does not consider non-stationary measurement signals, which are usually generated by machinery components. For this reason, time-frequency domain analysis should be used when dealing with non stationary signals, as it is applied over both the time and frequency domains.

Nonetheless, all the above techniques have their own advantages and disadvantages, so experience and case-dependent knowledge are needed to choose the appropriate approach.

In Chapter 4, we include a more detailed discussion regarding signal preprocessing and processing. We then present our approach for performing signal processing and extracting health indicators in the case of multi-component systems.

2.3.3 Diagnostics

Diagnostics is the process of detecting several fault attributes given the processed signal; these attributes are then used and compared to previously designated fault types. Some sources such as (245) discriminate between three different stages of fault diagnosis: the detection of a fault, fault isolation, and fault identification.

The detection of a fault happens by comparing to baseline data. The baseline data is sensory data about the components under examination when in healthy condition. A fault would then appear as a result of a difference between the baseline data and the current signal. Fault isolation is about the fault size and its location. Finally, fault identification tries to establish the pattern and severity of the fault, whereby the pattern indicates any regularities that are exhibited by specific faults, and severity is used to classify the failure mode based on its consequence.

The topic of fault diagnosis is an old one and has seen its fair share of research, and with the emergence of new data mining algorithms and recent developments in PHM some surveys books can be found on the topic, such as (245, 80, 44).

If performed manually, fault diagnosticss requires expertise and case-knowledge of the area of study. And with the ever increasing complexity of machines, there comes a need for more condition monitoring data to be collected and effectively manipulated to be able to extract important features. This is promoting the use of big data analytics and machine learning techniques to deal with such an issue. In Chapter 6, we use such an approach by way of implementing unsupervised machine learning, namely Gaussian Mixture Models (GMM). This is used for performing system diagnosis and discriminating between different degradation phases in multi-component systems.

2.3.4 Prognostics

Similarly to forecasting, whereby past and present available data are analysed in order to predict future trends; prognostics follows the same process yet instead of just projecting trends into the future, it is more concerned with predicting the EOL time at which a specific failure threshold is reached, and consequently extracting RUL estimation as seen in Eq 1.1. This is depicted in Figure 2.4, where at time $t = k$ an attempt to predict the EOL is made. Since EOL is uncertain, it is usually represented by a probability density function (PDF), and consequently so is the RUL.

RUL estimation relates to a common question in industry, which is how long for can a component operate before reaching a certain failure threshold. Then, based on the RUL estimation, appropriate actions can be taken. Therefore it is the remaining time to maintenance from current time. Moreover, when consulting the literature RUL is usually more addressed than EOL, however as Eq 1.1 suggests, these terms are strongly related. Furthermore, the lower bound of a confidence interval of the RUL is usually considered for conservative purpose (128).

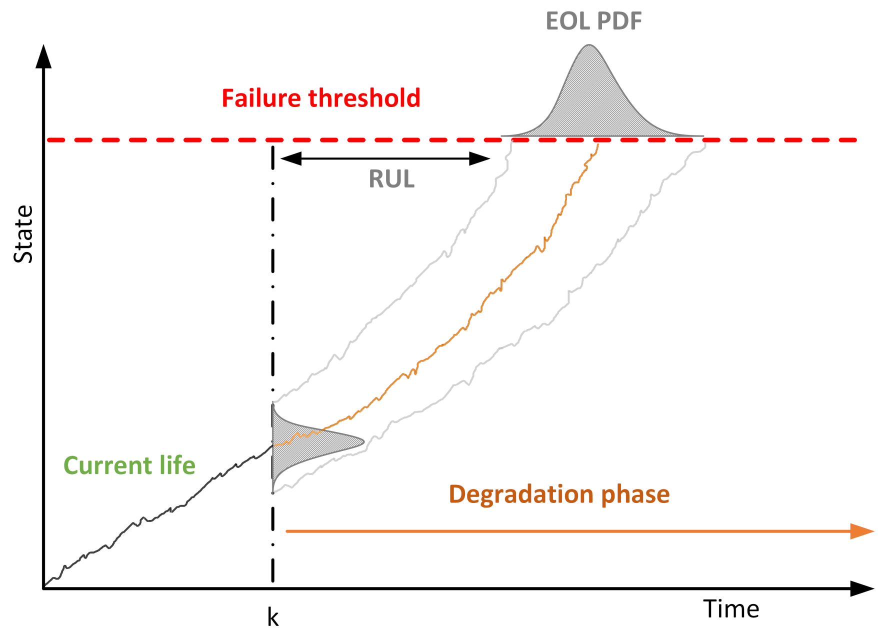

Figure 2.4: An illustration of prognostics and RUL estimation

This is crucial from both a cost-effective and a safety point of view, especially for critical equipment, such as aircraft engines, inertial navigation platforms used in aerospace and integral equipment on a production line.

Traditional methods for RUL estimation are heavily dependent on the time-to-failure data. However, these data are sometimes unavailable, as it is not always possible to have runs until system failure because of economic and safety issues. In such cases, data from degradation of components can be used as an alternative resource for RUL estimation. Several papers have reviewed and compared the main probabilistic prognostic methods for RUL estimation, see for example (223), and (135).

From the above mentioned, we conclude that prognostics is a vital step that helps industries manage their risks and prevents the occurrence of unforeseen components failures. And based on these predictions of future fault occurrences, maintenance and downtime costs can be minimised with CBM.

A holistic view of prognostics shows that it builds upon the following three aspects:

- State estimation: based on the collected data, this step is used to give an estimation of the degradation state of the component.

- State prediction: the task of state prediction is to predict the degradation tendency according to the information of the historical data.

- EOL and RUL prediction: serves for determining the time left under the degradation curve before final failure or before a predetermined failure threshold.

An extensive body of literature exists on prognostics approaches and applications, and therefore a considerable amount of review papers can be found where the classification of different prognosis approaches is presented (115, 133, 190, 193, 223, 225, 245, 248). Mainly, prognostics approaches can be categorised into three types: physics based prognosis, data-driven prognosis and hybrid prognostics. They will be further detailed and their literature reviewed in the following sections.

2.3.4.1 Physics Based Prognostics

Physics-based prognostics uses the physical model (or white-box model) of the degradation behaviour of the component under examination. This is then combined with the measured data to predict the degradation tendency and the RUL.

The approach takes into account the fundamental processes that create failures such as mechanical, electrical, chemical and thermal processes. Also, it requires the use of knowledge about geometry, and material properties of the monitored system (191, 192).

Works present in literature that use physics-based approaches are suitable in cases where the physics are known and the accuracy is of extreme importance as in the case of aerospace health management (207). These models are generally used for component level prognosis (260).

Although the physics-based prognostics method has its advantages, being very intuitive and a white box model, generalising these models can only occur for similar components. As one can imagine, the disadvantages of the approach are mainly caused by necessity of having a thorough knowledge of the component dynamics and physics. The model might also contain wrong assumptions and errors leading to inaccurate predictions.

If we take the topic of aircraft structural components for example, physics of failure models are developed by considering the effect of spectrum loading (34, 212). And most of these models are established on standardised test coupon geometry (37). Also these crack growth models of complex geometry are oversimplified, which might lead to a large inaccuracy when performing predictions on real systems. And although finite element models can be used to more accurately predict crack growth, this comes at a very high computational cost and is therefore impractical (230). Therefore these models are sometimes combined in a hybrid way with data-driven methods for updating the model parameters on-line.

2.3.4.2 Data-Driven Prognostics

These methods try to derive the future trend of the degradation process of the system from previous and current collected data. The assumption at the basis of the approach is that the statistical characteristics of the data from the components are consistent until a fault occurs. The EOL and RUL predictions are then formulated based on historical data (116).

Data-driven approaches are becoming more and more practical to implement as many libraries are now available for data mining and machine learning. Also the increasing computational power is making them more appealing to implement and are now considered as state-of-the-art in PHM. This is reflected by the extensive body of literature on the topic.

In this work, we use data-driven prognostics for our approach, this is specifically done using particle filters, and is presented in Chapter 5.

However, although data driven prognostics are showing considerable success recently, they still require knowledge about the system that is being monitored. This is for example needed for choosing the type of sensors used, and to estimate the amount of data needed for performing accurate prognostics.

Furthermore, since no physical knowledge is involved, this method could provide insights that are not considered in traditional modelling approaches, but might also lead to counter-intuitive results that do not point to the root cause of the failure.

We classify these data-driven approaches into two main classes, ML approaches and statistical approaches. However the difference can be blurred at times, and really depends on the referenced authors and what they consider their method to be. This matter can be highlighted when performing multivariate regression which is considered to be both a statistical approach and an ML approach.

Machine Learning Approaches ML approaches for prognosis are chosen when the complexity of the system is too difficult to capture with classical models. Also, as this artificial intelligence (AI) branch learns by example, usually a large amount of data is required for this approach to be effective.

Depending on the type of data, the learning process could be divided into supervised learning if the data are labelled, or unsupervised if the learning data are just made of inputs and with unknown output, therefore the task becomes mainly focused on clustering the data. Another less used variation is semi-supervised learning that uses few labelled data and mostly unlabelled data. It is important to note that even in beyond the field of PHM, supervised learning has attracted most of the attention since it had the most successful applications. When we talk about predicting future events, supervised learning is the ML approach commonly used, although some unsupervised and semi-supervised learning methods are sometimes combined with different approaches to achieve such a task (51, 170).

Some key insights for applying machine learning successfully are presented in (68). The main idea is that collecting large amounts of data is not enough, and that emphasis on feature selection and engineering is required in order to achieve good performance using ML. This however, to this day, remains the black art of this approach (68), and on its basis most of the PHM research using ML is published.

As the computational competencies of our systems advance and we are able to manage more and more data, which are also becoming widely available, ML approaches have been thriving in the field of PHM. Many reviews on these techniques have appeared in the last years (275, 186, 276, 278, 134, 225, 76). We are going to summarise the main approaches and some of their applications here. Note that the categories are classified without restriction on supervised, unsupervised, semi-supervised learning approaches.

In all of the ML approaches presented in the following, the accuracy of the results relies heavily on the features used and engineered to train the learning algorithm and to detect or estimate the fault propagation. It is crucial to use a scheme that is capable of selecting the most representative features. In (163) feature selection based on principal component analysis is presented and validated for both supervised and unsupervised ML approaches. Other works use wavelet packet transformations as in (146, 132), or Boltzmann machines (77).

We will not attempt to further divide ML methods into different classes, since in application the difference between the different ML approaches is very blurred and those classes suffer from a lack of consistency in the literature. We will however point out what occurs as the broader classes.

Some of the ML methods that are sometimes referred to as connectionist methods model the studied phenomena as the emergent processes of interconnected networks. The examples are used to learn the intricate relations in the data. One specific method stands out here. This is the universal function approximator, artificial neural networks (ANN) (285). ANNs have been recently used comprising of multiple hidden layers and referred to as deep neural networks (DNN), or deep Learning (DL) (138). Another important method that falls under this category is fuzzy logic (282).

In ML, Bayesian methods, usually refer to probabilistic graphical methods that are mainly used when uncertainty exists (131). Some approaches are: Markov Models as Hidden Markov Models (HMM) (86). Sometimes they encompass state estimation approaches or what is also referred to as stochastic filtering approaches such as the Kalman filter (KF), particle filter (PF) and their variants (225).

Another class of methods that can be found in the literature is instance based learning (IBL). These approaches, instead of performing explicit generalisation, compare new problem instances with obtained knowledge acquired during training that have been stored in memory. Some examples are the K-nearest neighbour algorithm (189) or case-based reasoning for advanced IBL(24).

Finally the combination of different approaches is referred to as ensemble methods (106, 19). These should not be confused with boosted methods, which are in some sense similar, and use the same method a multitude of times and in different feature combinations to reach more accurate results. A famous example of boosted methods is random forests, and in an oversimplified explanation it works by averaging the performance of many different decision trees to improve prediction accuracy (62).

In (84) the authors used feedforward backpropagation networks ANN to estimate the RUL of bearings using vibration information. The bearings are run until failure in their experimental setup and the data acquired are used to create a database of degradation signals. These are fitted with an exponential model in the form $\alpha e^{\beta t}$ where the parameters α and β are estimated. In (101) a feedforward neural network approach is validated on pump vibration data. The model is created using population characteristics and historical data. In (162) the authors use an ANN which has as inputs time and fitted measurements of Weibull hazard rates of RMS and kurtosis from its present and previous points, and uses normalised life percentage as output. The authors show that the accuracy of the prognosis is improved by using this method and these specific inputs. In (232) the authors use an analytical method using neural network modelling of vibration data in combination with short-time Fourier transform that is used to extract important features for the training. The model only uses data from a healthy motor component. Then the fault indicator is computed from fault condition used to generate analytical residuals. In (145), ANN is applied on vibration data for motor bearing fault diagnosis. The data have time-domain characteristics which are presented to the neural network. In (96) the authors present an optimised back propagation neural network method applied to fault detection in refrigerant flow air conditioning systems. The correlation analysis method is used to eliminate redundant variables. Then, the association rule mining method is used to optimise the features selection. In (165), unsupervised neural networks are used to generate an automatic algorithm and applied for an on-line fault detection of a three-phase induction motor. The approach uses stator currents as input variables. Then, the principal components of the stator current data are extracted using a Hebbian-based unsupervised neural network. These are used to verify the presence of the fault and its severity.

In (287), deep learning is investigated, specifically a deep convolutional neural network (CNN) is proposed for bearing fault diagnosis with proven ability to work under noisy environments and directly on raw signals. The importance of this approach is its adaptability. The approach however is based on the assumption that a big amount of data is available. In (146), vibratory and acoustic signals from a gearbox are used for fault diagnosis based on a deep random forest. A wavelet packet transform is used to extract information from the signals and two deep Boltzmann machines are used to develop a deep representations of these statistical parameters. The random forest is then applied. In this way, by fusing together the output of 2 deep representations, the authors are able to achieve 98 % accuracy for classification of faults in different conditions. (290) provides a survey of deep learning applications in PHM. Mainly, as in other fields such as computer vision, CNN seem to be the most successful in application. Furthermore, the authors also present the advantages and disadvantages of deep learning. They specifically discuss the difficulties that deep learning might face in PHM, most notably: the fact that DL is considered as a black box model, and so its validation for use on safety critical system might be jeopardised; also, that DL is by far the data driven approach that requires most data for performing efficiently, on top of that, data sets in PHM are usually unbalanced with very few fault and failure observation which might bias the models.

In (75) the authors use an unsupervised learning technique for identifying a fault. The feature vector characterises the status of the monitored equipment.

The feature space is decomposed into multiple fuzzy regions which represent the machine under different operating conditions. The model is essentially a set of evolving GMM fuzzy clusters, that are dynamically updated every time new features are available from data. In (201), the authors use a mathematical morphology technique, a nonlinear spatial analysis method for noise removal from data, and then fuzzy inference for detecting early faults in bearings of rotating machinery. Vibration signal data were provided.

In (25), the proposed method is based on two steps: first a nonlinear reduction of the features extracted from data, then these new features are used to train a support vector machine (SVM) regression model. The model learns to asses and predict the level of wear. In (132), a bearing fault detection method is presented and applied to a three-phase induction motor. Also in this approach, SVM is trained with features extracted from data. As signal-processing tool, a continuous wavelet transform is used for analysing vibration data.

Statistical Approaches These approaches use the collected data together with a probabilistic model of the system to estimate the RUL. The model is fitted to the data and this is then used to forecast degradation trends. This kind of approach requires a good amount of condition monitoring data which can create erroneous behaviours if not sufficient; or if the models used are of the wrong nature, whereby the condition is not accurately reflected.

In (223) a review on statistical approaches for prognosis is presented. Some common statistical prognostics approaches are: regression based methods, stochastic filtering, state estimation methods as Kalman filters, particle filters, Hidden Markov Models etc.

Regression-based models are commonly used in industry and academia for RUL estimation mainly due to their simplicity (151). The principle at the basis of these methods is that the degradation state of the components is related to some condition monitoring variables, then the RUL is predicted by monitoring these variables and predicting when a certain failure threshold is reached.

In (160) the authors were among the first to introduce a general nonlinear regression model for estimating a degradation trend. In (14) the authors presented a degradation model for light displays using nonlinear random coefficients models, allowing for non-monotonic degradation paths. In (85) Bayesian methods that use real-time sensor based condition monitoring information were used to update the parameters of a random coefficient model. This model assumed a Brownian motion (Wiener) error process. Then, they validate these models to degradation signals from accelerated testing of bearings. However, one disadvantage of random coefficient regression models is that they cannot model the temporal variability in the prediction of the RUL (188). Brownian motion with drift (Wiener processes) is used in (153, 243) to determine the RUL of contact image scanners. One disadvantage of Wiener processes is that only the current data from degradation is used to build the model, and all the historical information is ignored. This issue can be solved through an integrated model which considers the cumulative degradation path of the component as seen in (242). Gamma processes are usually used when the degradation process is monotonic and evolving in one single direction, in the case for example of wear or fatigue cracks. Their advantage is their relative simplicity, also they are able to take into account temporal variability (188). A good review for Gamma processes in the context of maintenance is presented in (248). These processes have been proven to be useful in maintenance applications (178, 53). In one recent example (155), a degradation analysis for characterising the health and quality of systems with monotonic and bounded degradation is carried out using a Gamma process. The method is applied on the light intensity of LEDs.

Markovian-based models have been extensively applied in prognostics. In (126) the authors evaluated the failure time for a single-unit system considering the wear as a Markovian process, while in (127) a semi-Markov-chain based model in used to partially solve the issue of independent or memoryless assumption of the system.

Hidden Markov model (HMM) based methods have been applied for RUL estimation problems thanks to their clarity (38, 21, 40). For example, in (261) a prognosis model based on a HMM and stochastic filtering is presented. The extension of this approach, the hidden semi-Markov model (HSMM) has been used as it overcomes the modelling limitation of HMM, as HSMMs do not follow the Markov chain assumption (69). For example, in (70) HSMMs are applied to fault classification of UH-60A Blackhawk's main transmission planetary carriers and prognosis of a hydraulic pump health monitoring application. In (166), a type of Gaussians Hidden Markov Models (MoG-HMMs), represented by Dynamic Bayesian Networks (DBNs), is used as a modelling tool. This allows the use of both temporal and frequency features from the raw signals. The proposed method is applied to real data form accelerated life tests of bearings.

Stochastic filtering-based models are one of the first methods to be used for prognostics. For instance, Kalman filtering was used in 1979 for health monitoring of aero-engines (213). Nowadays, some works on stochastic filtering-based models can be found in (122, 149), they usually use a particle filter approach in non-linear stochastic systems for fault detection. In (42), a probabilistic filtering approach is applied to the estimation of the RUL using oil-based wear information.

Covariate based hazard models are used when there is the need to assume that one or more parameters are affecting the component's wearing out process. These parameters are called covariates. Some works are found in (250), where RUL estimates are based on historic failure data from roller bearings, and (87), where the authors address the problem of imperfect observations.

2.3.4.3 Hybrid Prognostics

Hybrid approaches use both physics-based and data-driven methods and combine their advantages, and leverage their strengths to improve prediction performance. This is because here the data-driven methods are used to update the physics-based model parameters, which then maximises their prediction capability.

The knowledge on physical behaviour could in fact be used for creating the mathematical model in data-driven methods. More information about these approaches is found in (154).

In (39) a methodology for predicting the RUL in aircraft actuator components is presented using a fuzzy logic process for quantifying a system's damage index. This was implemented using rules derived from the knowledge of the system. Then, Kalman filtering was applied to forecast the progression of the degradation. Another example in (45), where the authors presented a prognosis approach for power devices under thermal stress using data-driven and physics-based methodologies.

2.3.5 Maintenance Management

This part of PHM builds upon the advancements already achieved in optimal maintenance decision making in the field of CBM. The optimality of these decision relies heavily on accurate prognostics and thus the RUL estimations that are provided. The idea here is to perform life-cycle decision-making or lifetime-extension of the systems at hand, aiming to achieve near-zero downtime and maximising the reliability and performance of the equipment.

However difficulties are faced which allow for further research to be conducted.

One of the main challenges in this area is transforming RUL estimates and PHM insights into decision making support.

In (239) the authors conduct a study investigating the adoption of certain prognostic approaches. They then provide a framework for implementing prognostics with decision making. They also identify many challenges that companies face when implementing prognostics of which the transformation of the actual monitoring data into insightful maintenance decision support. In (206), aeronautical systems are considered. The authors combine the RUL estimates with information on system architecture and therefore aim for the estimation of a system-level RUL. Then a methodology for decision support regarding maintenance planning is proposed.

2.4 Multi-Component Dependencies

The existing literature on CBM and PHM focuses on single component systems. This is because the stochastic analysis, and the composition of optimal CBM policies are much more difficult for multi-component systems (4).

According to the following literature reviews that mainly deal with the maintenance of multi-component systems (237, 57, 177), dependencies between components can be grouped into three types, namely structural dependence, economic dependence and stochastic dependence. However some reviews on the topic even consider other types, such as resource dependence in (125).

Multi-component dependencies are fundamental to this thesis. Stochastic dependence is specifically considered throughout this whole work. An overview of the different multi-component dependencies is depicted in Figure 2.5

2.4.1 Structural Dependence

Structural dependence concerns on the physical structure of the system, whereby inspecting or replacing one component may require intervention of other components, for example their dismantlement. Also, it considers the manner in which the components affect the system's availability given the configuration of the components themselves. We can mainly distinguish between two configurations, series and parallel (30, 125). In a series system, if one component fails the whole system fails. In a parallel system, the failure of one component is assumed to have no effect on the operation of the whole system. Structural reliability is itself

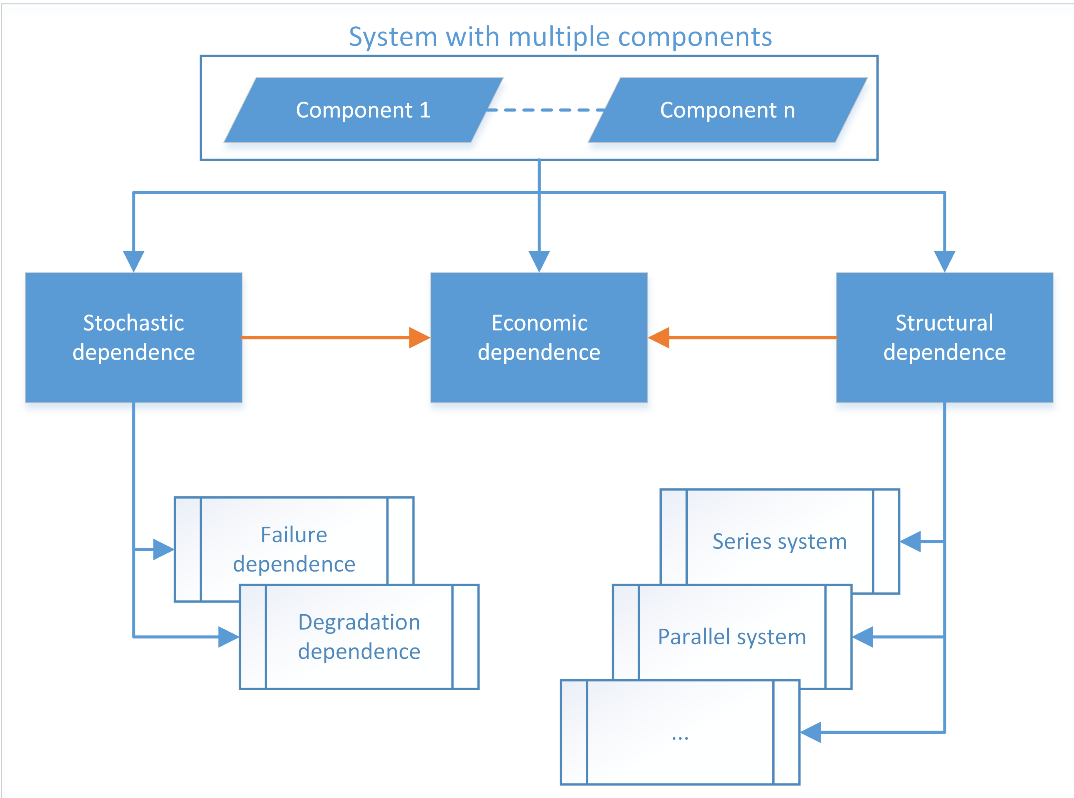

Figure 2.5: An overview of multi-component dependencies

a distinct field (167).

For example, in (147) three parallel multi-component systems cases were studied and the effects of stochastic and economic dependencies were considered. There, a maintenance policy is proposed that capitalises on the present dependencies especially the economic one and provide flexible replacement dates for maintenance grouping.

2.4.2 Economic Dependence

Economic dependence, which we consider in Chapter 7, deals with the cost effectiveness of certain maintenance policies, whereby combining the maintenance of multiple components can lead to reduced costs, positive economic dependence; or increased costs, negative economic dependence.

For an example of positive economic dependence we can consider offshore wind turbines. Travelling to turbines individually to perform a maintenance intervention could be costly. Therefore combining the maintenance of multiple components is beneficial. In (222), multi-bladed offshore wind turbines are considered, and an optimal opportunistic condition-based maintenance policy is proposed. They attempt to optimise the alert threshold aiming to reduce the average long-run maintenance cost, finally showing that the proposed policy shows benefits in both servicing costs and in greenhouse gas emissions.

Therefore economic dependence should be studied when maintenance policies are considered.

2.4.3 Stochastic Dependence

Often the deterioration processes of components are assumed to be independent, see (35, 176, 248). But real world systems are usually complex and include multiple interacting components. This brings about dependencies between the components which potentially affect the overall system availability. This then renders the single component CBM strategies sub-optimal. Recently however, CBM research has been showing a growing interest in multi-component systems (288, 125, 177).

Stochastic dependence is also known as probabilistic interaction (238), failure interaction and probabilistic dependence (177).

This type of dependence means that the degradation of one component can affect the degradation process of other components, usually accelerating the other components' degradation leading to unexpected faults and failures.

Referring to the literature we see that we can distinguish between different types of stochastic dependence. In (171), the authors were the first to propose a distinction. Stochastic dependence was divided into three types; these are failure interactions of types I, II and III. These failure interactions first considered a two component system in (171), but some were later generalised into multi-component systems in (172).

- A type I failure interaction occurs when a natural failure of one component i can induce a failure in other components $j \neq i$. Considering $0 \leq p, q \leq 1$, the failure of component i can induce failure in a component j with probability p, and is said to have no effect on the failure of a component j with probability $1 - p$. A component j can induce failure in a component i with probability q, and is said to have no effect on the failure of a component i with probability $1 - q$. If components i and j are independently subjected to degradation then $p, q = 0$.

- A type II failure interaction is considered as a shock that affects the failure rate of the remaining components of a system after one component fails. Therefore we can consider the failure rate of components by not just the age, but the number of received shocks as well.

- A type III failure interaction is described as being the combination of failure interactions of Type I and II.

In (125) stochastic dependencies are classified into three types as well: failure induced damage, load sharing and common-mode deterioration.

- Failure induced damage is described as the damage that can affect other components when a certain component fails. The damage caused can be origin of a major increase in degradation level of other components or even cause them to fail.

- Load sharing is described as an increase in load on certain components when one component fails. This is so the system keeps generating the same output, and thus a load of work that was divided on an n number of components is now considered to be divided on $n-1$ number of components. This in turn leads to accelerated degradation of the other components that are still functional.

- Common-mode deterioration is described as the concurrent degradation of multiple components. This is when multiple components are subjected to the same working conditions and therefore an increase in degradation of one component implies a proportional increase of degradation for other components.

In (177, 204) the previously mentioned type I and type II failure interactions are only considered, with the latter now called failure rate interactions, and more recently this failure rate interaction is being referred to as degradation interaction where a failure does not have to occur in order to affect the degradation behaviour of other components.

In (152) degradation interactions are further divided into induced and inherent dependence.

- Induced dependence is described as the damage caused to other components when one component does not function properly. They refer to the shock damage model as the typical model used in such cases.

- Inherent dependence is described as the degradation interaction that takes place between multiple components caused by operational and load sharing circumstances. Typically this is modelled using multivariate distributions or couplas as used in (105).

The Two Main Groups of Stochastic Dependence In the case of all of the previously mentioned stochastic dependencies we are really concerned about one question, how does the health state of one component affect the health state of other components. The health state itself can be defined in many ways. These are typically whether a component is failed or not; or the age of the component; or its performance; or other indices derived from condition monitoring; or a combination of all these.

We can therefore organise the stochastic dependencies discussed earlier into two main groups: failure interactions and degradation interaction.

Failure interactions are considered when we wait until the failure of one component to see the effect on other components. This effect can be causing failure of other components as considered in (219), where age based replacement and opportunistic age based maintenance policies are considered to capitalise on the economic dependence that exists in such a system. Or, the effect can be causing a change in the degradation behaviour of other components, usually accelerating the degradation process. This is seen in (215) where a two component system is considered, and where the component 1 is subjected to failure based on a Poisson process, and if component 1 fails a random amount of damage is caused to component 2, this is accumulated until component 2 reaches its failure threshold. A preventive replacement policy is then presented. However it is shown that this does not necessarily lead to lowered expected maintenance cost.

In contrast with failure interactions, degradation interaction is considered when all components are still in operation, i.e. no component has yet failed, and where the health state of one component affects the degradation rates of other components. This is referred to more recently as the state-rate interaction (11, 205), this includes for example common mode deterioration. This kind of interaction is being considered more frequently in recent works. This is not surprising since achieving failure in components is highly unappreciated and since recently there has been an increase in the use of sensors and IOT systems to enable more condition monitoring, and because of the boost in development of prognostics approaches and PHM in general. This kind of stochastic dependence is covered in more detail in Chapter 5 and the results are a fundamental aspect of this thesis.

2.5 Discussion

In this chapter we presented an overview of the field of PHM, starting from the different maintenance strategies that can be considered and showing the importance of CBM. We then moved into the topic of PHM where all its aspects were detailed, namely data acquisition, signal processing, diagnostics, prognostics and health management. We also presented an extensive literature review on prognostics and specifically on data driven approaches, which are considered state-of-the-art. We finally covered multi-component system dependencies which are of great interest to this thesis, and specifically detailed stochastic dependence.

We therefore conclude this chapter with the following note. Although degradation interactions are being considered lately, rarely do these studies focus on the PHM aspect. And as was the case in 2008 (178) for failure interactions, these stochastic dependencies are merely being used as a stepping stone to maintenance optimisation, and the modelling of these dependencies itself is still scarce in the literature. Therefore the work presented in this thesis aims at tackling this issue by presenting a study on multi-component systems from a PHM point of view. Hence we present degradation modelling for such systems within the PHM framework. This is the broad thrust of this thesis.

Chapter 3

A Gearbox Accelerated Life Testing Platform

Contents

3.1 Chapter Summary

Experimentation is a requirement to understand the dynamics of real world systems. This is especially true for complex behaviours such as degradation processes of physical systems.

In this chapter we present a gearbox accelerated life testing platform. We design it with the purpose of generating data that is suitable for studying multi-component degradation; and for providing more intuition towards modelling such complex behaviours. We therefore detail the design and development of the gearbox, and analyse and discuss the data that we obtain.

3.2 Experimental Platforms

Although modelling and simulation of mechanical system dynamics and even cyber-physical systems can be achieved, some well established scientific references on these topics still consider it challenging to conceive of a demonstration which can include all properties of real systems (124, 61). Therefore experiments of real physical systems are still essential in order to investigate their real world behaviour. This is even more crucial when considering complex behaviours such as system degradation where physics based models have had their shortcomings highlighted in the recent literature, see Section 2.3.4.1. For this reason many experimental platforms have been developed for the sake of collecting monitoring data which can later be analysed in order to improve our understanding of such behaviours. By doing so, this makes it possible to more accurately perform diagnostics and prognostics for these systems.

When it comes to the condition monitoring literature, two experimental platforms along with their data sets stand out, and this is because of how often they are considered and cited in the prognostics and health management (PHM) and condition based maintenance (CBM) literature. They are: the NASA bearing data set or IMS bearing data (113, 197) which consists of three tests to failure of four bearings placed on a single shaft as seen in Figure 3.1; and PRONOSTIA (173), which is an experimental platform for bearings accelerated life test and that can be seen in Figure 3.2. This platform has been used to generated the data set adopted in the IEEE ICPHM 2012 data challenge.

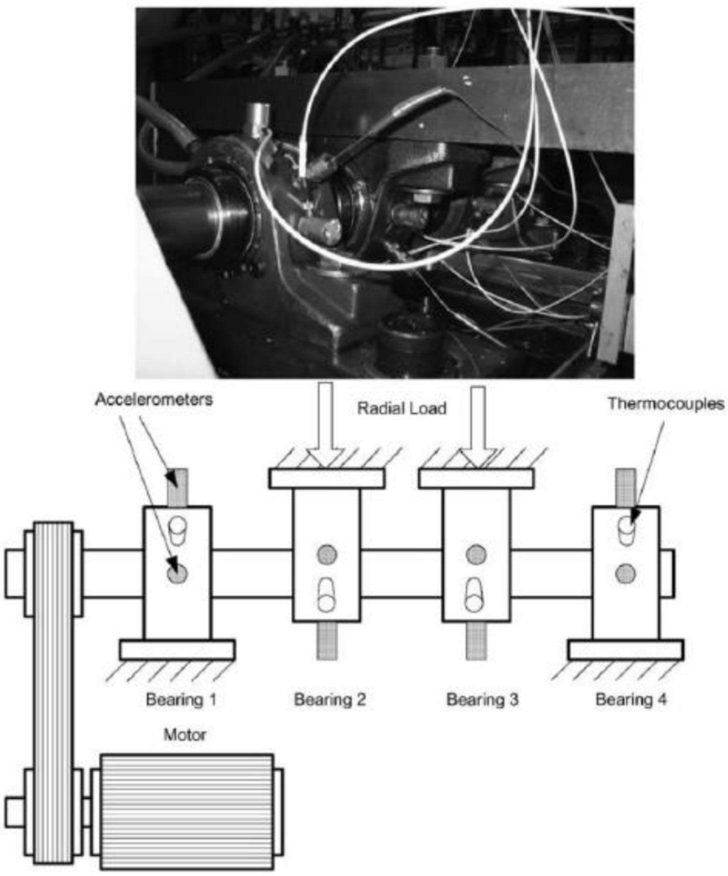

Figure 3.1: NASA bearing test rig (113)

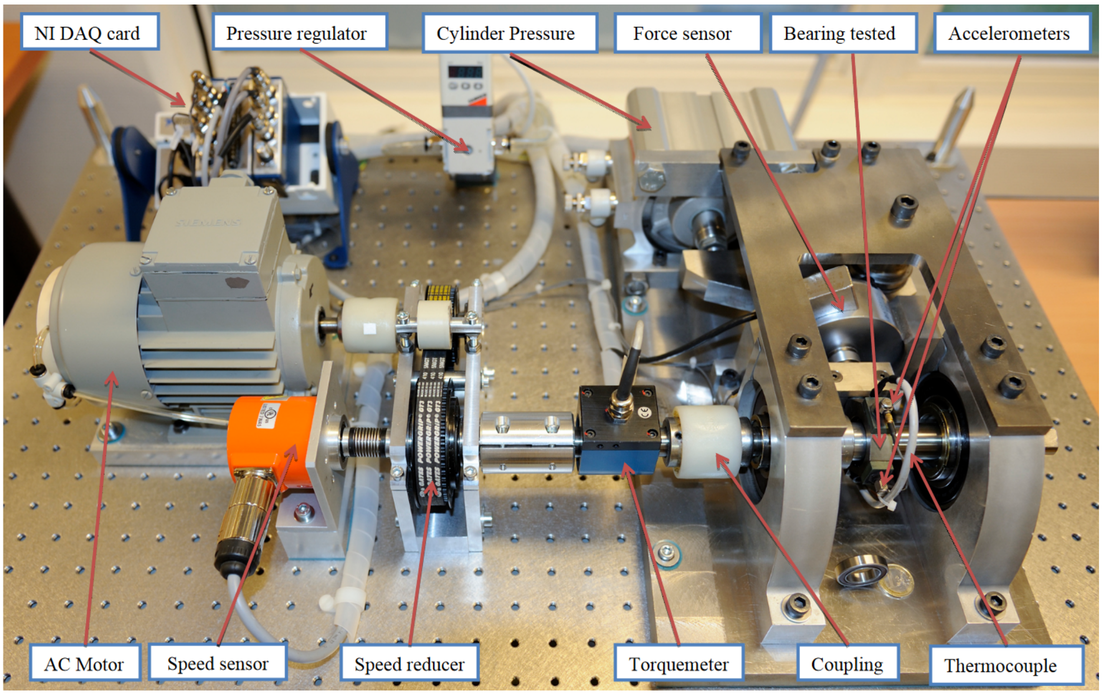

Figure 3.2: PRONOSTIA an experimental platform for bearing accelerated life test (173)

3.3 Gearbox Experimental Platforms

Many means of mechanical power transmission exist, mainly including gears, belts and chains. However gears are usually considered to be the most durable and adapt, specifically when considering their power transmission efficiency that is around 98 % (120). Accordingly, gearboxes are present in virtually any mechanical system, playing the essential role of torque and speed conversion, therefore unforeseen faults can lead to lowered machine up time and decreased plant efficiency.

One example is the amount of attention that the condition monitoring literature pays to gearbox systems in the application of wind turbines (159, 180, 81). This becomes very clear when referring to a study made on condition monitoring and prognostics of utility scale wind turbines (111). This study shows that 30 % of maintenance costs can be attributed to the gearbox and generator. The gearbox accounting for 12 % of all failures, coming second by just 1 % after the electric control system, which accounts for 13 % of all failures. This can be clearly seen in Figure 3.3.

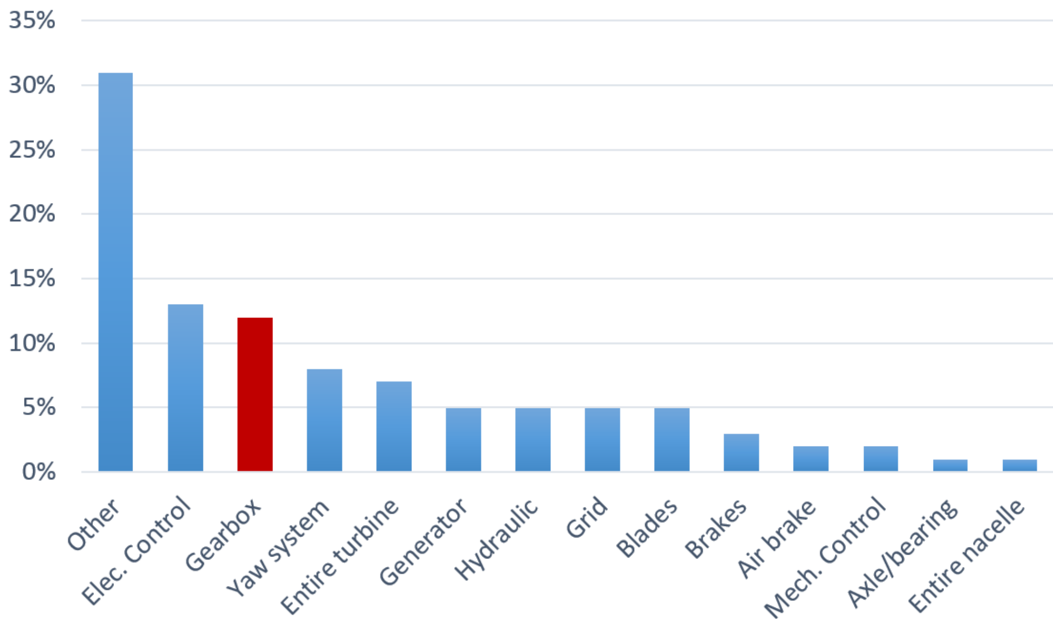

Figure 3.3: Causes of failure for wind turbines

3.3.1 Existing Gearbox Experimental Platforms

Although the experimental platforms mentioned in section 3.2 investigate bearing degradation processes, other less frequently cited platforms have been developed for acquiring condition monitoring data in a gearbox systems setting. Examples include developing an experimental platform for diagnosing gear faults by investigating acoustic emission such as in the case of a vehicle gearbox in (5, 148); and for a dynamic gear transmission test rig in (247). In (185) a gearbox platform was used for investigating gear pitting via analysis of vibration signals emitted by the gearbox. A gearbox developed to test different polymer base gears at different rotational speeds and torques has been developed in (221), gear surface temperature measurements were used for conducting failure analysis.

3.3.2 Gearbox Condition Monitoring Approaches

Gearbox condition monitoring can be achieved in many ways, the main approaches along with some of their main advantages are summarised in Table 3.1. For further detail, the following references can be consulted (292, 159).

In our work we specifically consider gear degradation. And among the different sensing approaches we consider vibration since it has been extensively studied and successfully applied for the purposes of diagnosing industrial rotating machinery

Sensing	Monitored Components	Advantages
Vibration	Gear Bearing Shaft	Reliable Standardised (ISO10816)
Torque	Rotor Gear	Direct measurement of rotor load
Oil and Debris Analysis	Gear Bearing	Direct characterisation of bearing condition
Temperature	Bearing	Standardised (IEEE 841)
Acoustic Emission	Gear Bearing	Able to detect early-stage fault Good for low-speed operation
Current and Power	Gear Bearing	No additional sensor needed Inexpensive Non-intrusive

Table 3.1: Gearbox condition monitoring approaches

(41). Vibration analysis is mainly used for components such as gears (137), bearings (8, 235) and induction motors (23).

The principle of vibration analysis is that all machinery generates vibration. This becomes more apparent when machines contain rotating components. These vibrations have characteristics that differ for different components and for different component conditions. Therefore when a machine is in a healthy operating condition, its vibration signature is distinct from that of a faulty machine (117).

3.4 Description of the Developed Experimental Platform

The experimental platforms presented in Sections 3.2 and 3.3 do not provide data sets that are fit for studying multi-component system degradation. Still, some do acquire data from multiple components such as in the NASA bearing data set (113), where the test rig consists of degradation data acquired from four accelerometers mounted on top of four shaft support bearings. However, this data still can't be used since once we investigate the experimentation design, we see that when a fault is identified, all components are then replaced in order to initiate a new test run. This eliminates the possibility of validating the effects of stochastic dependency on component lifespan.

For this reason we have developed a gearbox experimental platform (Figures 3.4 and 3.5) which aims to validate degradation models for multiple component systems, giving more insight into the true nature of the degradation processes of interacting parts in a real world environment. The choice of using a gearbox as a testing platform is motivated by i) the fact that in a gearbox contains multiple components (gears) which are interacting to form a gear train. and ii) a gearbox is an integral component of almost all industrial machinery as discussed in Section 3.3. Therefore it represents a desirable case study from the point of view of PHM.

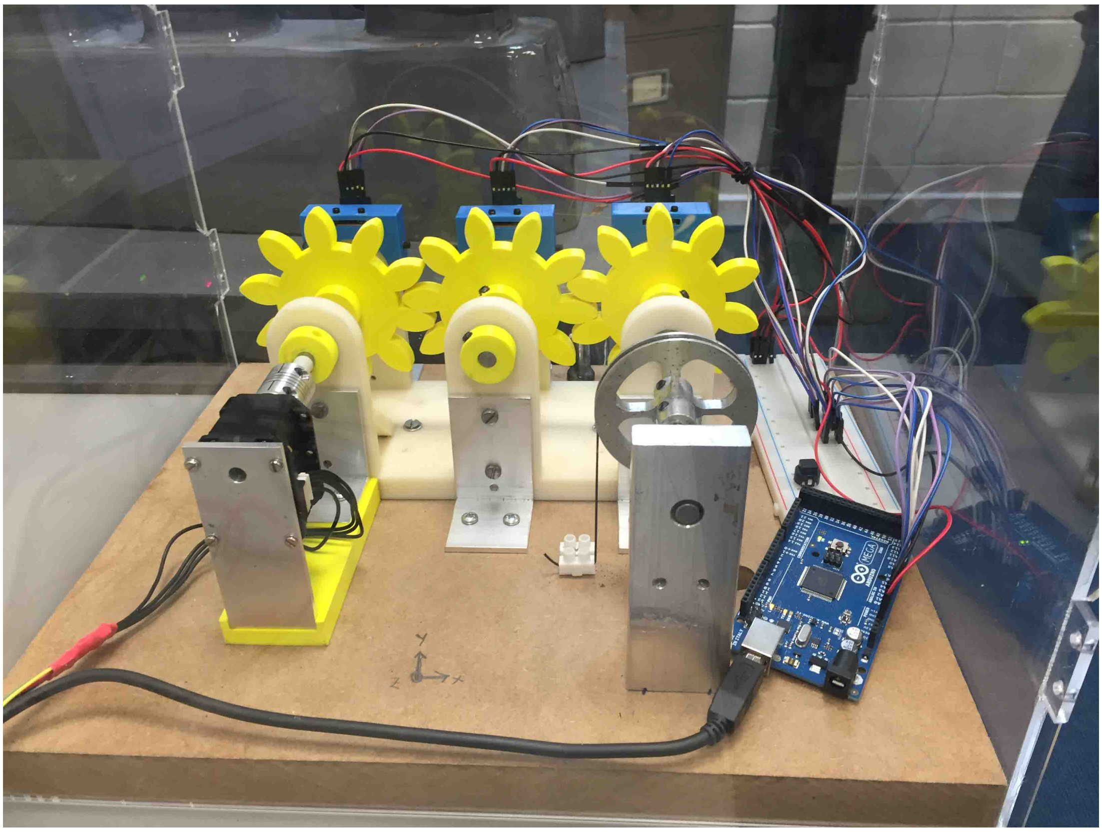

Figure 3.4: The experimental platform, side view

The platform has been designed and developed at the autonomous systems and robotics lab at the University of Salford by using: computer aided design (CAD); 3D printing; the use of a lathe for cutting and preparing the reinforcement parts which are made out of metal for stiffening the platform; and the use of an Arduino

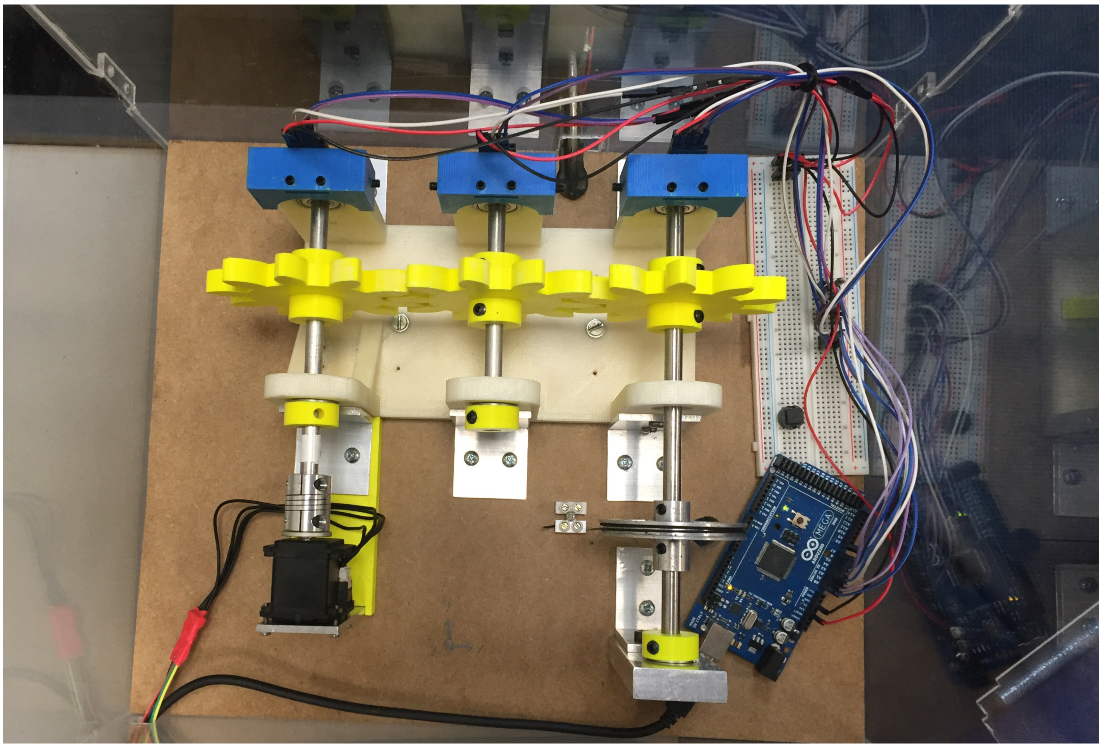

Figure 3.5: The experimental platform, top view

board based data acquisition (DAQ) system. The objective of this platform is to gather experimental data for multi-component degradation models, for the reason that multi-component systems with inter-dependencies follow a highly stochastic degradation process which is inherit property of complex mechanical design.

This experimental platform can be divided into 3 main parts, which are: the driving part, the load part, and the data acquisition part. The numbering of the gears and mounted accelerometers, as well as accelerometer orientation are depicted in Figure 3.6.

The platform can be operated using different configurations. It can test variable gear modules (size and number of teeth), and which are made out of different materials. Other parameters can also be adjusted such as: the driving part, the load part, and the measurements part. This is further presented in Table 3.2.

A detailed description of the experimental platform configuration that is used in this chapter is presented in the next sections.

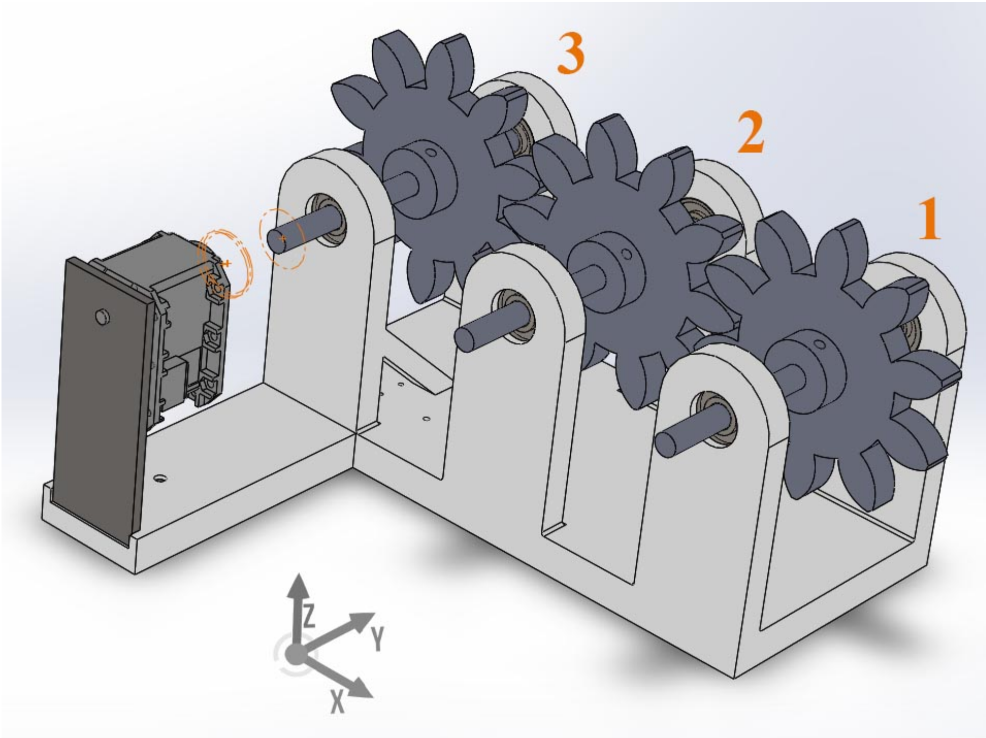

Figure 3.6: Accelerometer/Gear numbering and orientation

Gears	Driving Part	Load Part	Measurements Part
Material	Motor type	Dynamometer brake	Accelerometers
Module	Torque	DC generator	Acoustic
	Speed	Pneumatic brake	DAQ specifications

Table 3.2: Experimental platform adjustable parameters

3.4.1 Driving Part

This part consists of a fixed bracket which holds a continuous turn servo motor, and can be exchanged with different motor sizes and of different power. Feedback is collected on the driving servo motor including motor position which allows rounds per minute (RPM) to be measured, the load on the motor is also acquired along with the motor's instantaneous temperature, which allows for a fail-safe threshold to be set and so prevent overheating. The motor is coupled with a shaft which drives the platform. That shaft is mounted with a gear that meshes with a second gear which in turn meshes with a third gear forming a gear train. The three shafts are held each by two shaft support bearings on each side of the rig. We restrict translational movement of the shafts by using small washers that are in contact with the inner ring of the bearing from one side, and held in place by a 3D printed shaft collar on the other side, creating a frictionless rotation all while keeping the shaft in place, and thus preventing additional noise originating from uncalculated loads as a result of friction.

The installed motor has a stall torque of 7.3 N.m and can reach speeds up to 78 RPM. Both the speed and torque can be set by the user using a software development kit (SDK) or a MATLAB script which in turn saves all commands and records motor operating speed, current usage, torque and load on motor.

3.4.2 Load Part

The load is being applied through a pulley system that resembles a dynamometer mechanism. The pulley is fixed on the last shaft, furthest from the driving shaft, which is in turn held in place by another shaft support bearing. A strong filament is wrapped around the pulley once, the filament is fixed at one side of the medium density foam (MDF) base, and attached on the other side to a load hook where some weights are hanged. This provides an opposing torque on the motor thus creating a load that can be easily changed by adjusting the weights that are being supported.

3.4.3 Measurements Part

Three sensors are used to collect vibration data from this rig, these sensors are 3-axis accelerometers, each mounted as close as possible to the centre line of the shaft support bearing which avoids picking up distorted signals. This is realised by fixing the sensors in place using hex socket screws that pin the sensors to a 3D printed housing, which in turn is fixed on the rig using other hex socket screws.

The 3 accelerometers have each a full sensing range of +/-3g. These accelerometers are connected to an Arduino Mega 2560 board that plays the role of a DAQ system. It converts the analogue signal into a 10bit digital signal with a value range of 0-1023, the sampling frequency of this board can go up to a 10 kHz on 1 single channel. The digital data is then transmitted to MATLAB via serial communication and a 2.0 USB port. Binary communication is used to make sure that the speed of communication over a data channel or baud rate, is never exceeded, and thus no information is lost. The sensor data is also coupled with the Arduino's internal microsecond clock reading. This provides us with the time at which the analogue signal was read, and so gives us robust sampling time and sampling frequency. This can be used to switch from time domain to the frequency domain when processing the data.

3.5 Results

3.5.1 Experimental Setting

The data presented in this section is collected from the experimental platform after running it continuously for 2 hours. This data has a sample size of 1000 corresponding to 2.5159 seconds of accelerometer readings. The load that is being applied to the system corresponds to 1 N.m of opposing torque on the motor. The motor that was used to produce these results is the MX-67 Dynamixel. The motor was running at 14.8V and had an average current intake of 0.66A and thus using 9.768 Watts of Power. The running speed was 64-65 RPM. Using Eq. 3.1 and 3.2 we can calculate that the motor produces a torque of 1.43 N.m.

$$P_{Watt} = V_{Volts} \times I_{Amps} \tag{3.1}$$

$$P_{Watt} = \tau_{N.m} \times \omega_{rad/s} \tag{3.2}$$

The gear train is made out of 3 gears meshing in series, all these gears have the same number of teeth, 10. The gear meshing frequency is calculated to be 10.83 Hz at a running speed of 65 RPM. The sampling frequency for this data is 397 Hz.

3.5.2 Experimental Platform Data

After collecting the raw data, we start with the pre-processing. We centre the readings to 0, and then translate the digital signal that is received from Arduino to a measurement in terms of acceleration of gravity G's. The result can be seen in Figure 3.7. A further discussion on this topic is provided in Chapter 4.

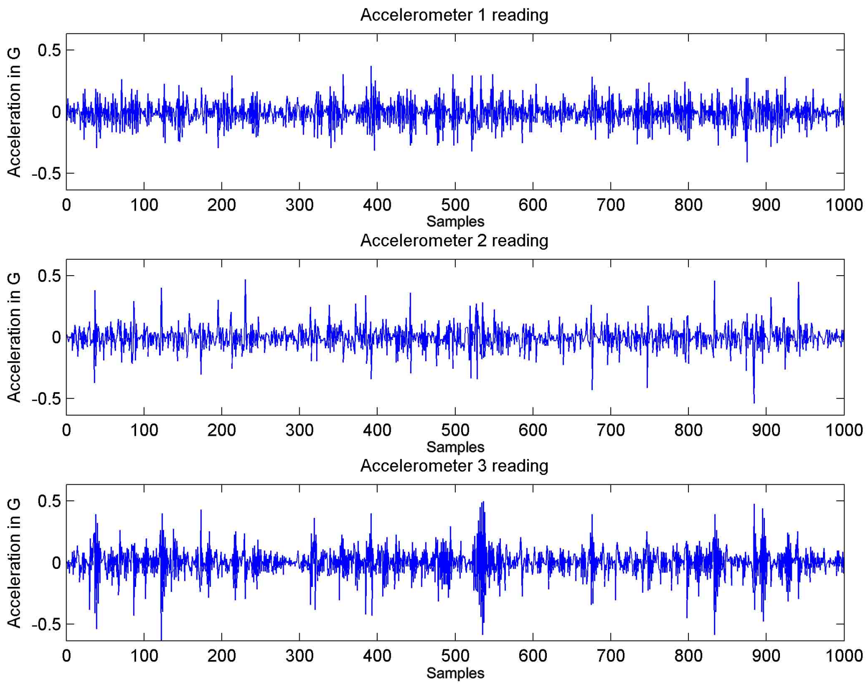

Figure 3.7: Accelerometer signal from the x-axis of the accelerometers

In Figure 3.8 we present the fast Fourier transform (FFT) of all 3 accelerometers in the x orientation, we also highlight the gear meshing frequency at 10.7 Hz and its harmonics, using the dashed red lines.

3.5.3 Multi-Component Interaction

It is important to point out that each accelerometer is mounted over the shaft supporting its respective gear, and so accelerometer 1's signal corresponds mainly to gear 1 and so on. Also, it is important to note that an increase in vibration intensity for a certain gear will lead to a decrease in efficiency when it comes to power transmission. This is caused by the absorption of some of the power that is being transferred. Therefore the vibrating gear would then experience more forces leading to a faster degradation.

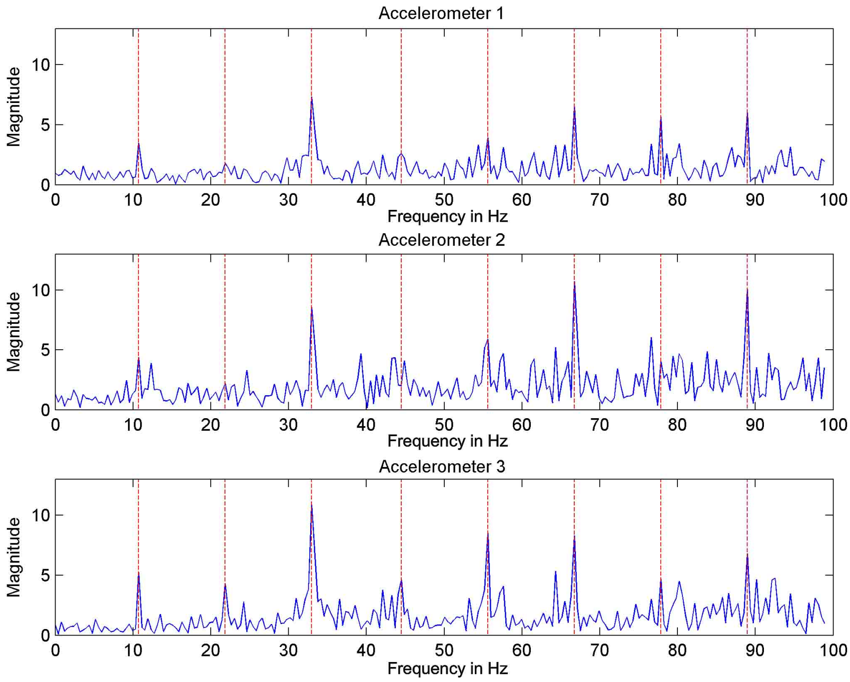

Figure 3.8: FFT of accelerometers' x-axis readings. The gear meshing frequency and its harmonics are highlighted in red

In more detail, this happens because during the gear meshing process, the gears' tooth flanks experience contact load, this is in the form of sliding and rolling motions. The sliding motion causes material removal from the gear's tooth surfaces. This material loss changes the geometry of the tooth profile and renders the dynamic characteristics of a gear to be sub-optimal, consequently vibration increases and leads to more sliding motion and thus accelerated degradation (289).

To measure dependencies between system components the following experiment was carried out:

- Only x-axis accelerometer data were collected, freeing 6 channels from analogue to digital conversion and thus raising the sampling frequency to 655Hz per accelerometer.
- Tooth surface wear was induced into gear number 3 uniformly on all 10 teeth.
- The induced tooth wear on gear number 3 happened in 4 stages, where at each successive stage the wear was slightly increased.

The experimentation platform ran at each stage after inducing wear, and the accelerometers' data were collected. The raw data were then pre-processed and transformed into readings in G's. Then for the signal processing part, the signal's envelope was computed to give a clear indication of the change in vibration intensity, see Figure 3.9. We used the Hilbert transform to get the envelope of our vibration signal. the Hilbert transform takes the FFT of the vibration signal, zeroes out the negative frequencies, and then performs an inverse fast Fourier transform (IFFT). The absolute value of the real and imaginary parts of the transform is the envelope.

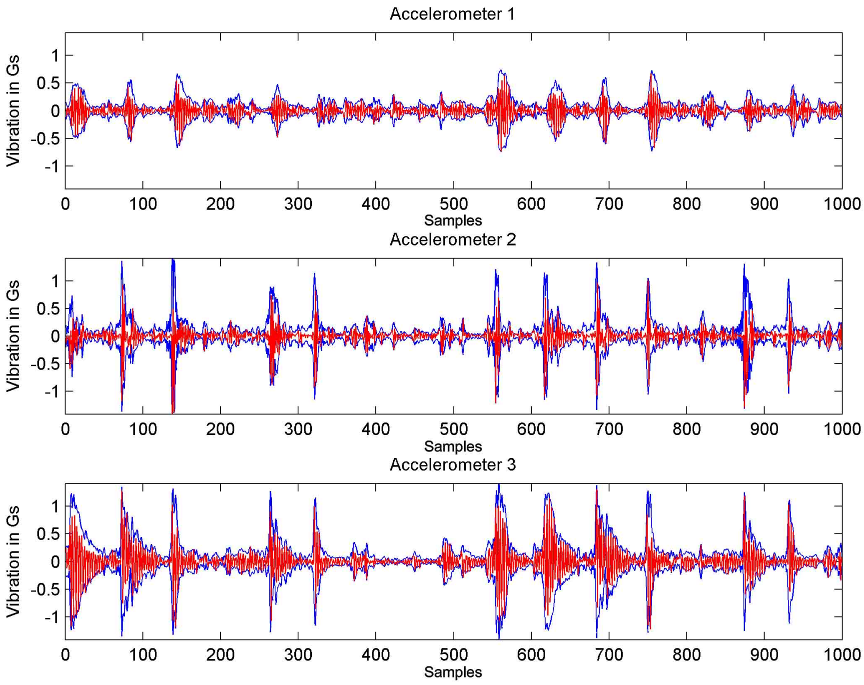

Figure 3.9: Accelerometer data after wear stage 4, acceleration signal in red and envelope in blue

In Figure 3.9, the envelope of the vibration signal shows a greater amplitude in accelerometer 3 when compared to the other 2 accelerometers, and if we sum up the values of the envelopes, we would then have a higher value for accelerometer 3 which is a clear indicator of the now elevated vibration intensity due to wear. Some statistical properties for the accelerometer signals were also computed along with the sum of the signal envelopes. These are presented as wear state indicators of the gears in Figure 3.10.

We can see that after inducing wear on the teeth of gear 3, we get on av-

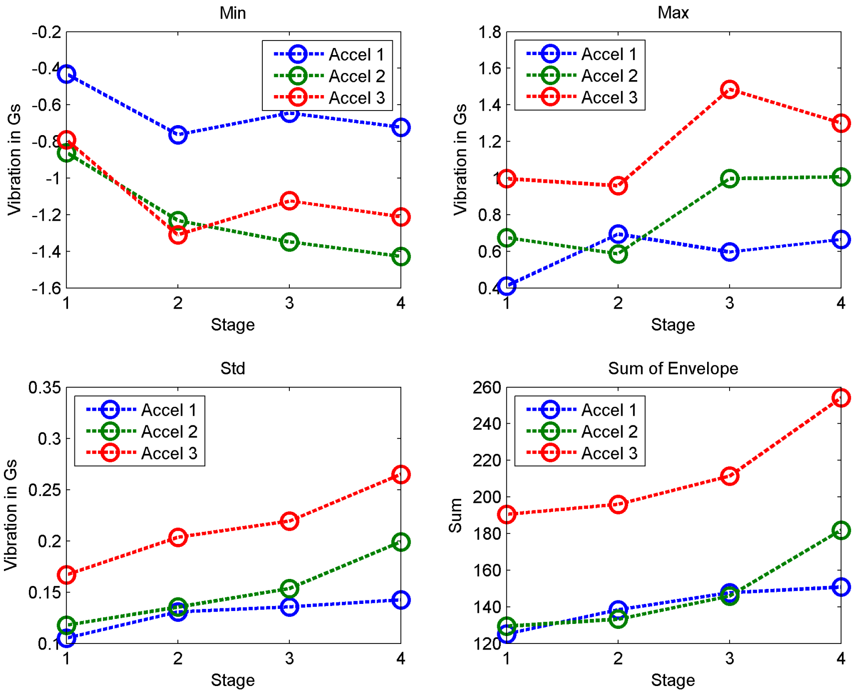

Figure 3.10: Component wear state indicators of all 3 gears at the 4 wear stages

erage a higher maximum value and lower minimum value for gear 3 indicating an increase in the range of vibration, peak to peak vibration, and thus indicates higher vibration intensity. This increase in vibration intensity is represented even clearer when we look at the standard deviation and the sum of the envelope of the signal, we can see that there is always an increase in these 2 indicators the more we induce wear onto gear 3.

From Figure 3.10 we can also observe dependency between the 3 components. This is clear when analysing the standard deviation present in the lower left corner of the figure, and the sum over the envelope values at each stage. We notice that an increase in wear in gear 3 has a strong positive correlation with the increase in vibration in gear 2, with which it is in direct contact and so has direct dependence. A smaller positive correlation with gear 1's vibration level occurs. In this case there is an indirect contact and so an indirect dependence is present. This dependence can be clearly illustrated if we compute the difference between vibration intensity, indicated by the standard deviation or the sum over the signal's envelope, when moving from one experimental stage to another. We then see the positive trend of vibration intensity which is incurred after each stage of inducing wear on gear number 3.

3.6 Discussion

In this chapter we have presented our gearbox accelerated life testing platform, its design and its development. An experimental setting was run and condition monitoring data in the form of vibration was acquired though the use of three accelerometer sensors. The data were then pre-processed and processed in order to extract heath state information about the gears. The results were then displayed showing a clear sign of degradation dependence between the interacting gears. From a physics point of view we notice that when one gear is worn out, the rate of degradation of the other gears is accelerated through an increase in vibration.

A further investigation using another experimental configuration of this gearbox platform is presented in Chapter 4. This considers continuous runs to failure of two gears. In doing so we collect data in a continuous fashion, and consider the gears to have failed when a certain vibration intensity is reached, this is done also according to a threshold on power transmission efficiency, and gear state. After a first run to failure, only one gear is replaced, and now coupled with a worn out gear. Then the test runs again allowing us to estimate the parameters of the degradation model that we discussed in Chapter 5.

Chapter 4

Health Indicator Extraction in Multi-component Systems

Contents

4.1 Chapter Summary

While prognostics and health management (PHM) puts considerable emphasis on prognostics, that would in fact not be possible or result in poor predictive capabilities if the health indicators were not correct, whereby they do not accurately represent the health condition of the system (115, 202, 128). For this reason the first step in PHM is to perform health indicator extraction. This relies on condition monitoring and the proper processing of the acquired data.

In this chapter we begin by providing a detailed background on the three main steps of the health indicator extraction process. These are: data acquisition, pre-processing and processing. We then introduce our methodology for extracting accurate health indicators from vibration signals, and transforming those into a time series which represent the degradation trajectories of the different components of a multi-component system.

4.2 Introduction and Background

Health indicator extraction sits at the heart of PHM, it is responsible for refining the raw condition monitoring data so that diagnostics and prognostics can make efficient use of it.

Monitoring the state of machine condition has received great attention in industrial maintenance (259, 235). Usually, the measured signals from real data contain some noise. Therefore, signal pre-processing and processing are necessary steps that need to be performed after data acquisition. Signal pre-processing aims to eliminate the noise in the signal and increase the signal-to-noise ratio (SNR). The final step of health indicator extraction is signal processing which is used to extract health indicators or fault-related information from machinery (142, 143). These indicators would help us accurately diagnose and predict the future states of the system.

This aspect of PHM is done in three main phases, data acquisition, pre-processing and processing of the data. An overview can be seen in Figure 4.1.

4.2.1 Data Acquisition

Data acquisition is the first step of the health indicator extraction process. This relies on acquiring information about the system, it is usually done directly with

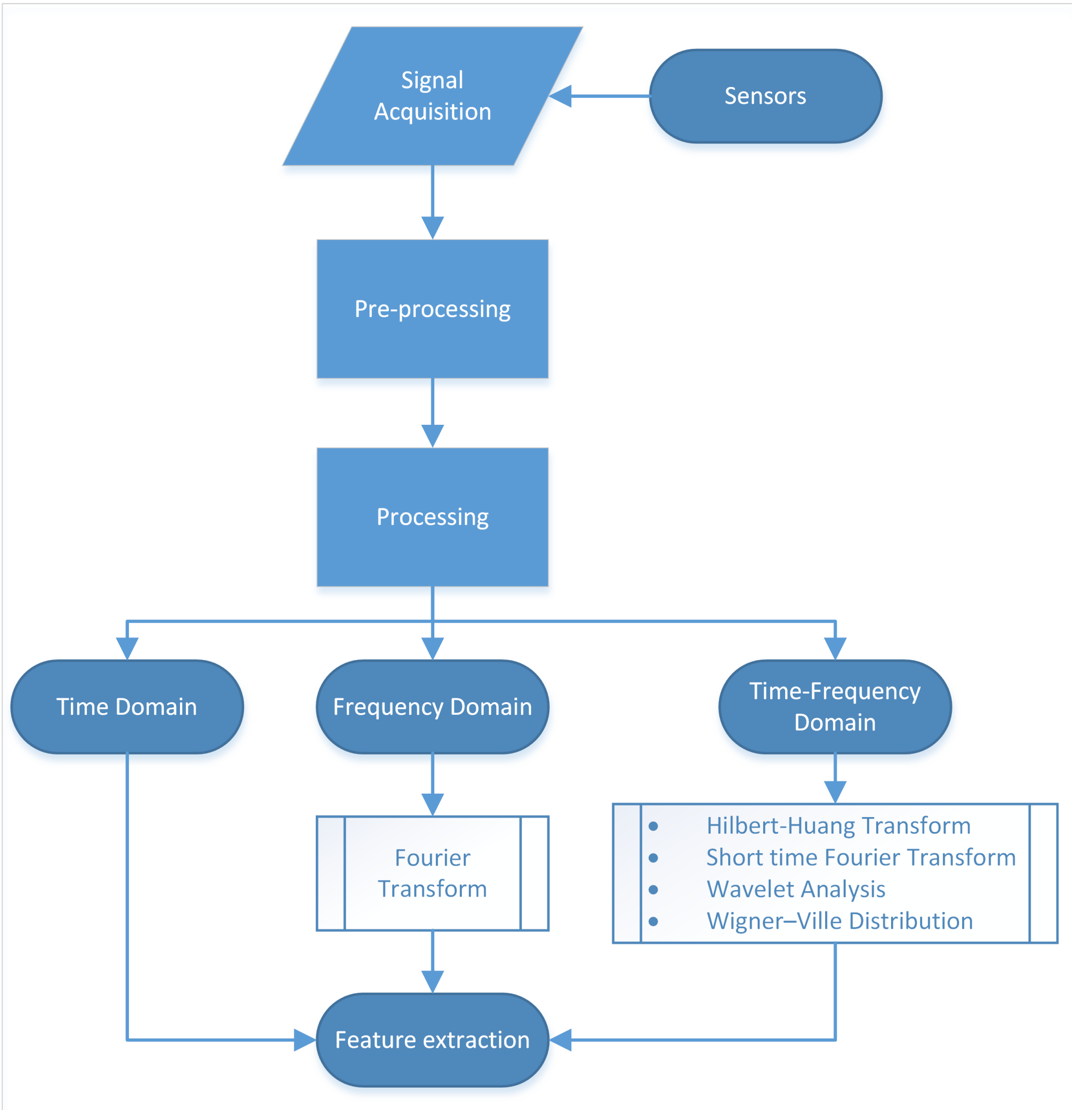

Figure 4.1: Overview of health indicator extraction steps

the use of sensors (49), or indirectly, for example by means of analysing performance or checking the states of lubricants that pass through the system. The key concept here is to extract useful information about the internal state of the system using external means, this is because the system is by preference kept in operation.

The two most popular industrial techniques for achieving this are vibration analysis and lubricant analysis (202). Vibration monitoring allows continuous monitoring of the system, in contrast with lubricant analysis which in most cases stands as intermittent monitoring. We could therefore immediately start to see the advantages of vibration analysis over lubricant analysis. For this reason among others, vibration analysis is considered as the most effective way of conducting monitoring (202, 255). This is especially true when dealing with machinery that contains rotating equipment (41). In this work we use accelerometers to capture vibration data.

An overview on the topic of data acquisition was presented in Section 2.3.1. For a more in depth study of the topic of data acquisition techniques, we suggest consulting the following sources (13, 129) for general applications, and (49) regarding the application in the context of PHM.

After choosing the adequate data acquisition approach, signal pre-processing and signal processing follow.

4.2.2 Signal Pre-Processing

Before processing the data and extracting important indicators, pre-processing signals is a fundamental preparation step to reduce background noise, filter out the errors of measurement systems, and improve the SNR.

Common techniques for data pre-processing are mean-centring, unit variance scaling and signal filtering.

4.2.2.1 Signal Scaling and Centring

A first necessary step is to tailor the raw data so that future calculations are more easily and efficient conducted. There are different ways of approaching this pre-processing step of the signal.

Firstly, the trend component of a signal does not usually contain any important information about mechanical faults. A systematic shift in the data can in fact result from sensor drift. Therefore the trend of a signal should be removed before the signal processing step. This enables us to focus the analysis on the fluctuations in the data. Also, data might have been detected from different sensors, so each stream of data could show different numerical ranges. It is therefore very important to standardise raw data.

These steps can be accomplished by applying unit variance scaling and mean centring. We denote $\mathbf{x}$ to represent a vector containing all the raw data. The mean of a set of data values is represented by x_m and is calculated as:

$$x_m = \frac{\sum_{n=1}^{N} x(n)}{N} \tag{4.1}$$

where N is the number of data points, and n is the n^{th} value. The standard deviation is represented by x_{sd} and is calculated as:

$$x_{sd} = \sqrt{\frac{\sum_{n=1}^{N} (x(n) - x_m)^2}{N-1}} \tag{4.2}$$

And so we centre the data by subtracting the mean value of the signal from all data points as in:

$$\mathbf{x_c} = \mathbf{x} - x_m \tag{4.3}$$

where $\mathbf{x_c}$ represents a vector of the now centred data. Unit variance scaling is then accomplished by dividing all the centred data by the standard deviation s as in:

$$\mathbf{x_{UV}} = \frac{\mathbf{x_c}}{x_{sd}} \tag{4.4}$$

where $\mathbf{x_{UV}}$ denotes the data after unit variance scaling is applied.

4.2.2.2 Signal Filtering

Signal filtering (59) is the process of eliminating unwanted frequencies from the signal and the interference of noise. This can significantly improve the visibility of the signal.

The frequency response can be classified into a number of different bandforms describing which frequency band the filter passes, the passband, and which it rejects, the stopband.

There are many way of classifying filters. According to amplitude–frequency characteristics, describing which frequency bands the filter passes and which it rejects, the filters are commonly divided into four categories: low-pass, high-pass, band-pass, and band-stop filters.

The filters work based on their cutoff frequency w_c. A low-pass filter is a filter that passes signals with a frequency lower than w_c. The frequency range w which occurs $w < w_c$ is called passband while $w > w_c$ is the stopband of the filter. Likewise, the passband of a high-pass filter is $w > w_c$ and stopband $w < w_c$. Instead, for a band-pass filter all the frequencies within the interval $w_{c1} < w < w_{c2}$ will pass, while the stopband is $w < w_{c1}$ and $w_{c2} < w$. Band-stop filters work as the opposite of a band-stop filter.

Referring to the properties of the signal that is being filtered, the filters can also be grouped into analog and digital filters. The latter category is additionally divided into infinite impulse response (IIR) and finite impulse response (FIR) filters.

Let's assume that an acquired signal $x(t) = s(t) + n(t)$ contains a useful component $s(t)$, and a noise component $n(t)$. By using a general filter, we want to compute a filtered signal $y(t)$. The notion of conventional filter is developed on the basis of signal analysis in the frequency domain. The filtered signal in frequency domain will be:

$$Y(w) = X(w)H(w) \tag{4.5}$$

where $Y(w)$ and $X(w)$ are the Fourier Transform (FT) results of the signals $y(t)$ and $x(t)$ respectively, and $H(w)$ is the transfer function of the filter itself.

If noise and useful components are in different frequency bands, the filter can be designed so that it can preferentially remove noise and keep the useful components. The transfer function of the filter will be:

$$\begin{cases} H(w) = 1, S(w) \neq 0 \\ H(w) = 0, S(w) = 0 \end{cases}$$

where $S(w)$ is the FT of the useful components of the acquired signal $x(t)$.

The filter can be also constructed when working in time-domain. However, instead of using the product operator, in time-domain, we have to use the convolution operator $\circledast$.

$$y(t) = x(t) \circledast h(t) \tag{4.6}$$

where $h(t)$ is the inverse FT of $H(w)$, namely unit impulse response function of the filter. In other words, the concept underlying the formulas is the time-domain-frequency-domain duality.

4.2.3 Signal Processing

Signal processing is the step where health features are actually extracted from the signal. This is performed as an analysis in one of the three different domains: the time domain, frequency domain, and time frequency domain. An overview of these analysis approaches is given here. They will be further detailed in the following subsections.

- Time-domain analysis, which is applied directly on the time waveform of the acquired signal. Features are then extracted from the time waveform, usually descriptive statistics such as the peak to peak interval, the mean, standard deviation, crest factor, root mean square, skewness, kurtosis etc.

- Frequency-domain analysis, which is based on the frequency domain of the acquired signal, such as spectrum analysis using the Fourier Transform (FT) and the Fast Fourier Transform (FFT). The advantage in using the frequency domain over the time domain is that certain frequency components of interest could be isolated, and so this can lead to more robust and targeted health indicators.

- Time-Frequency domain analysis, where the analysis is performed on both time and frequency domains. Techniques such as the short time Fourier transform (STFT), which is used in this work, can be applied to perform such analysis. The advantage over the frequency domain analysis is the ability to handle non-stationary waveform signals, which are of great interest when working with machinery faults.

4.2.3.1 Time-Domain Analysis

Time-domain analysis is directly performed on the acquired time–domain waveform to identify specific signatures.

When dealing with a stationary signal, time-domain approaches extract the signal characteristics using statistics like mean, variance, root mean square (RMS), standard deviation, etc.

If the signal is non-stationary, the extracted characteristics could inherit some non-linearities which can then complicate the successive step of prognostics. Non-statistical approaches comprise of auto-regressive moving average (ARMA), or signal processing methods like synchronous averaging or correlation (270).

Commonly extracted statistical features in the time domain are presented in Table 4.1.

Statistical feature	Formula
Signal mean value	$x_m = \frac{\sum_{n=1}^{N} x(n)}{N}$
Standard deviation	$x_{sd} = \sqrt{\frac{\sum_{n=1}^{N}(x(n)-x_m)^2}{N-1}}$
Root amplitude	$x_{root} = \left(\frac{\sum_{n=1}^{N}\sqrt{\lvert x(n)\rvert}}{N}\right)^2$
Root mean square	$x_{rms} = \sqrt{\frac{\sum_{n=1}^{N}(x(n))^2}{N}}$
Peak	$x_{peak} = max\lvert x(n)\rvert$
Skewness	$x_{skewness} = \frac{\sum_{n=1}^{N}\left(x(n)-x_m\right)^3}{(N-1)x_{sd}^3}$
Kurtosis value	$x_{kurtosis} = \frac{\sum_{n=1}^{N}\left(x(n)-x_m\right)^4}{(N-1)x_{sd}^4}$
Crest factor	$x_{crest} = \frac{x_{peak}}{x_{rms}}$
Clearance factor	$x_{clearance} = \frac{x_{peak}}{x_{root}}$
Shape factor	$x_{shape} = \frac{x_{rms}}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$
Impulse factor	$x_{impulse} = \frac{x_{peak}}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$

Table 4.1: Commonly extracted statistical features in the time domain

where N is the number of data points. $x_m, x_{root}, x_{rms}, x_{peak}$ are especially useful when working with vibration signals. This is because in general a fault usually excites mechanical vibration, which is then reflected in the amplitude of the signal. $x_{kurtosis}, x_{crest}, x_{shape}, x_{impulse}$ can instead indicate incipient faults as they are sensitive to sharp impulses in the signal.

For example, in (195), 15 time-based features are initially calculated from bearing vibration data. Then their significance is analysed and reduced using feature dimentionality reduction, namely discriminant analysis. Finally 6 features were used for training a neural network based classifier.

Another well utilised time-domain analysis approach is called time synchronous average (TSA) (55), it uses the ensemble average of a raw signal over time as shown in Eq 4.7 and is used on signals that contain excessive noise with the aim of reducing that.

$$\bar{x}(t) = \frac{1}{N}\sum_{n=1}^{N} x(t+nT)$$
$$0 \leq t < T \tag{4.7}$$

where T is the averaging period and N is the number of samples for averaging. An extensive review on TSA in the field of PHM can be found in (22).

More advanced approaches apply time series models to the signals, from where features are extracted. The autoregressive (AR) model is used to regress the variable on its own past values, it is usually presented as $AR(p)$ where p denotes the number of lags. The AR model can be combined with a moving average (MA) approach which is usually presented as $MA(q)$ where q denotes the number of lags for the MA. The combination or AR and MA is the autoregressive moving average (ARMA) model:

$$x_t = a_1 x_{t-1} + ... + a_p x_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + ... + b_q \epsilon_{t-q} \tag{4.8}$$

The above is an ARMA model of order p, q, where ϵ represents white noise, and a_i, b_i are the parameters of the AR and MA models respectively.

In the literature AR models were applied for extracting features from vibration signals of an induction motor such as in (194), support vector machine (SVM) based classification was then applied on the extracted features. SVM was also applied on features that were extracted by means of frequency domain analysis, namely power spectrum density, cepstrum analysis and signal processing with higher order spectra. It was shown that the AR extracted features allowed SVM to have the best classifiaction performance, with 100 % classification accuracy. However, in this work the classification accuracy is not fully justifiable since metrics such as precision-recall, and f1 score were not reported.

However, in (16), the authors perform a comparison of the AR model, back propagation neural networks and radial basis function where these techniques were applied on different signal lengths. In this comparison, the linear AR model is ranked superior in terms of speed of operation, however back propagation neural networks outperformed the other techniques in terms of accuracy. In fact, the modelling technique based on back propagation neural networks required shorter signal lengths, almost half of the vibration data and performed more accurately when classifying induced rolling element faults. This shows the better noise rejection ability of the back propagation network when compared to traditional linear methods.

4.2.3.2 Frequency-Domain Analysis

Frequency-domain analysis identifies and isolates frequency components of the signals, so it is generally considered more effective than time domain analysis for fault diagnosis application.

The most applied technique in frequency domain analysis, and which allows for transforming time waveform signals into the frequency domain, is the discrete Fourier transform (DFT), and more specifically the FFT which is a more computationally efficient variation of DFT.

The main advantage of using these frequency-domain techniques for fault diagnosis is that they are appropriate for stationary signals (139) and they have already been applied for feature extraction in several categories in the field of PHM, as in (50) for studying the effect of localised changes in stiffness magnitude and phase of gear systems, and in (229) for identifying faulty brushless three-phase synchronous generators by representing their flux-density distribution trough FFT.

The most commonly used tool in the frequency domain is the power spectrum. The power spectrum shows the portion of a signal's power falling within given frequency bins. It is defined as $E[X(f)X^*(f)]$ where $X(f)$ is the Fourier transform of signal, E is the expectation and * denotes the complex conjugate. Other than the power spectrum, other useful spectra for signal processing have been developed in the last years. Other methods are cepstrum, spectral analysis, higher-order spectra and envelop analysis, which are also identified in (115).

Some useful auxiliary tools for spectrum analysis can be found in the literature. In (104) the authors apply an envelope spectrum together with a self-adaptive noise cancellation method in order to remove discrete frequency masking signals. The envelope is achieved by using the Hilbert transform technique. The approach is validated on simulated and actual vibration signals. Also, a band-pass filter and a shift in frequency are applied to the raw signal so that the reduced number of samples could improve the computational costs. In (231), an amplitude modulation (AM) detecting technique is presented for identifying single-point defects in rolling element bearings and applied to simulations and real vibration data from bearings. The approach relies on checking the characteristic fault frequencies in the power spectrum and searching for some peaks of energy to characterise a fault. A bispectrum analysis is also performed, which, together with the AM detector could detect faults in the components. The authors show that by using these three tools they could accurately detect incipient faults.

Other frequency analysis techniques proposed for fault detection and diagnosis are side band structure analysis (32) and Hilbert transform (HT)(279, 200, 72).

However, one of the main limitation of frequency domain analysis is the inefficiency when dealing with non-stationary signals, which are commonly measured from degrading machinery. Also, frequency domain analysis lacks temporal information which is crucial for performing prognostics (139).

4.2.3.3 Time-Frequency-Domain Analysis

Time-frequency domain analysis investigates signals in both time and frequency domains, this is done in order to show fault patterns on specific frequencies in time, and thus allows for more accurate diagnostics.

These techniques are advantageous in the case of non-stationary signals. For example, the short time Fourier transform (STFT) which was first introduced by Gabor et al. in (79), divides the whole signal into short-time windows and then apply Fourier transform to each one of them.

This idea of being able to work with evolving frequencies throughout time has been applied thoroughly and many techniques other than STFT have emerged. For example in (123) vibration signals are acquired from a gearbox operating under several loads, and where artificial defects were introduced in one of the gears. A measure of the current drawn by the motor is also stored. A discrete wavelet transform (DWT) is then applied to these signals where its output is sent to a STFT to recover health features. The authors show how the health features extracted from the frequency windows can predict a change in the energy of the gear mesh frequencies and therefore can be used as a monitoring tool.

However, as STFT works by diving the signal into windows, this brings some limitations as it can only be applied to signals with slow change in their dynamics.

Other approaches in time-frequency-domain are Wigner-Ville Distribution (33). These are bilinear transformations which do not divide the signal into windows as the STFT, thus overcoming this limitation. However, these approaches present a more difficult interpretation of their estimated distribution.

Wavelet transform (WT) (52) is a time-scale representation of a signal, which has also been widely used in signal denoising. It is especially used for vibration content characterisation as in (48). In this work the authors use a variant of the classical WT, an overcomplete rational dilation discrete wavelet transform, as a fault feature extraction technique. This method is applied for showing hidden periodical impulses of low energy in the signals.

In (283) a wavelet packet transform is used for bearing fault detection. A stator current is measured and decomposed into sub-bands at predetermined levels using the wavelet packet algorithm. Then coefficient energies in different frequency bands are calculated. In comparison with a healthy condition, the energy coefficients are higher in the frequency bands related to defects, therefore they are used as a fault index.

Also, a combination of WT and FT was proposed in (254) to enhance feature extraction capability.

Empirical Mode Decomposition (EMD) (108), is among one of the most powerful time–frequency analysis techniques suitable for non-linear and non-stationary processes (234). However, some drawbacks include a high sensitivity to noise.

Other common approaches are the Hilbert-Huang Transform (HHT) (109), spectral kurtosis (9) and cyclostationary analysis (82).

Based on the advantages that are presented by the time-frequency domain analysis; and since we are interested in isolating a specific frequency of interest, namely the mesh frequency and observing its evolution over time, time-frequency-domain analysis will be considered in this thesis, and this should be the case when considering multi-component systems.

4.3 Extracting Heath Indicators for Systems with Multi-Component Wear Interactions

The following sections details our methodology for obtaining component health state data from multi-component systems. This is then evaluated on a new configuration of the gearbox accelerated life testing platform that was presented in Chapter 3. Our methodology for state indicator extraction is presented as a flowchart in Figure 4.2.

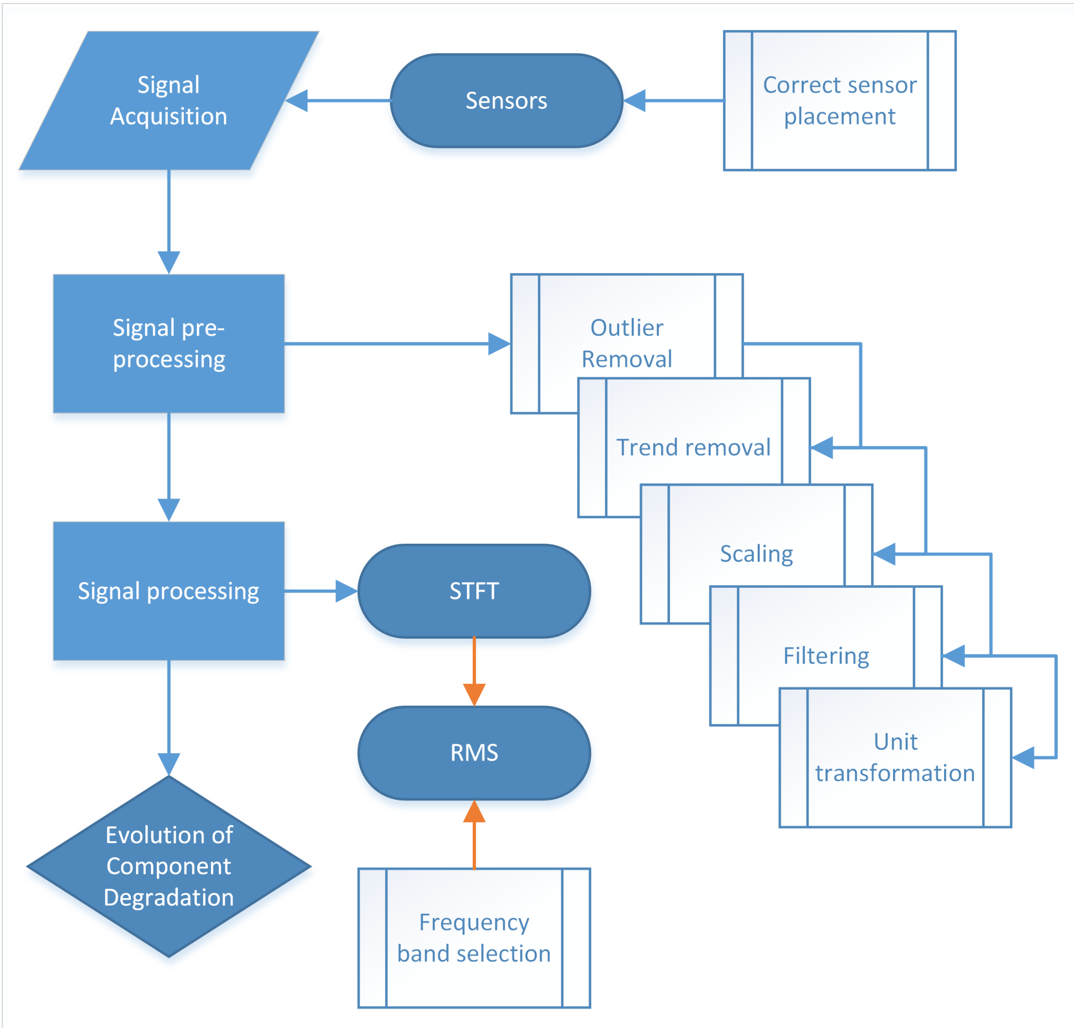

Figure 4.2: Methodology for extracting components' health state indicators in a multi-component system with wear interactions

4.3.1 Methodology

Data Acquisition As shown in Figure 4.2, we start by acquiring data from sensors, specifically using accelerometers to gather vibration data, since it has been extensively studied and successfully used for the purposes of diagnosing industrial rotating machinery, mainly for components such as gears (137), bearings (8, 235) and induction motors (23).

In a multi-component system setting, it is wise to use multiple accelerometers. The placement of these accelerometers should be scattered evenly around the system, or at least the aim should be not to have them too close together. This grants different vantage points for data collection, and should be done so that the different components of the system can be easily differentiated, especially when dealing with similar components that can emit signals around similar frequencies.

Signal Pre-Processing After performing the data acquisition step on a multi-component systems, and since clean data are not often encountered in an industrial setting, a pre-processing step for that data should be considered. It is therefore advised to first clean the data specifically from any outliers that are present. This is because any other pre-processing step that would be performed afterwards such as the ones seen in Section 4.2.2, and which usually heavily rely on the signals' statistical properties, would then have erroneous outcomes.

This cleaning however could be a tedious task to perform manually if not impossible. Therefore we suggest the use of Algorithm 1, which could easily automate such a process. First a window of data points based on the operating profile of the system should be specified and fed to the algorithm. Then the median value or geometric mean of the data and the median absolute deviation (MAD) are computed for that window. The values that exceed the median plus or minus the MAD value, are then filtered by replacing them with a random variable sampled as $X \sim \mathcal{N}(med,\ mad)$, thus preserving as much as possible the true nature of the signal. This is important for diagnostics and prognostics.

Data detrending and centering should then follow. Scaling should be done if necessary, depending on the presence of dissimilar sensors or if different data ranges are used. Filtering the data can follow after this step, this depends on the frequency band of interest and whether other unnecessary frequencies can be rejected without loss of information. This was detailed in Section 4.2.2.2.

Finally the physical meaning of the signal should be obtained, so that an

Algorithm 1: Outlier Removal Algorithm

```
w represents the window length;
input : A signal Sig, a row matrix of size m × w
output: Signal Sig with no outliers
for i ← 1 to m do
    med←ComputeMedian(Sig(i));
    mad←ComputeMAD(Sig(i));
    for j ← 1 to w do
        if Sig(i, j) < (med − mad) or Sig(i, j) > (med + mad) then
            Sig(i, j) = X ∼ N(med, mad)
        end
    end
end
```

engineering perspective can be added. This depends on the specifics of the sensors used. For example if accelerometers are used, this step should be applied and would result with a signal that has its acceleration denoted in acceleration of gravity (G), instead of the generic digital signal value.

Signal Processing At this point we can apply signal processing to the signal with the aim of extracting health state representative features from the signal.

A major challenge for modelling existing stochastic dependency in a multi-component system is the complex nature of the signals acquired. Each signal may represent a mixture of the signals of all components at once, but to varying degrees. Therefore, an accurate way of acquiring component specific degradation state information from multi-component systems is to consider time-frequency domain analysis.

Time-frequency analysis is used in blind source separation (273, 1), in which mixed signals are separated without the aid of information. This is performed by exploiting the difference in the time-frequency signatures of the sources to be separated. Much of the literature in this field focuses on audio applications (196) and machine sound signals (291). Applications of time-frequency analysis for identifying various sources of vibrations data can be found in (253, 58).

A further motivation for this choice over the other data analysis approaches, namely time-domain analysis and frequency domain analysis has been presented in Section 4.2.3.

Consequently, an STFT can be applied on the cleaned signal, and allow for the analysis to be performed in both time and frequency domains, isolating the frequency components of interest all while representing the evolution of their energy through time.

The STFT can be applied over the time-waveform data of a component i in the following manner:

$$s_i^{'} = STFT\{s_i[n]\}(\tau, \omega) = \sum_{n=-\infty}^{+\infty} h[n-\tau]s[n]\exp^{-j\omega n} \tag{4.9}$$

Where $s^{'}$ represents the short-time Fourier transform of the input signal $s(t)$, and $h(t)$ the window function. Optimum window length depends on the application. A high resolution in time and in frequency cannot be accomplished simultaneously. If high resolution in time domain is needed, the size of the window should be reduced. If the application demands frequency domain information to be more specific, then the size of the window should be increased (121, 214). Therefore, if we want to resolve the fundamental and harmonics of a signal, a long window should be used. If it is needed to detect the onset or presence of some events, a short window should be used. Some examples of window functions are Gaussian and a Hamming windows (99, 118).

After the STFT is applied on the signal, the frequency root mean square (FRMS) can be computed over the frequency band of interest. This is done in order to estimate how the magnitude of the frequency band of interest evolves in time. This is applied as such:

$$X_{FRMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} s_i^{'2}} \tag{4.10}$$

where N is the number of data points, and n is the n^{th} value.

In this way, we can study a time series signal that describes the evolution of the health condition of the components over time. This makes the prognostics aspect of PHM easier and more effective.

4.3.2 Case Study

In an industrial setting, gearboxes are present in virtually any mechanical system, playing the essential role of torque and speed conversion. Therefore, unforeseen faults can lead to lowered machine up time and reduced plant efficiency. A gearbox is a good example of a system with multiple components. Therefore with the aim of collecting data on multi-component interactions we carried out our experiments on the gearbox accelerated life testing platform shown in Figure 4.3.

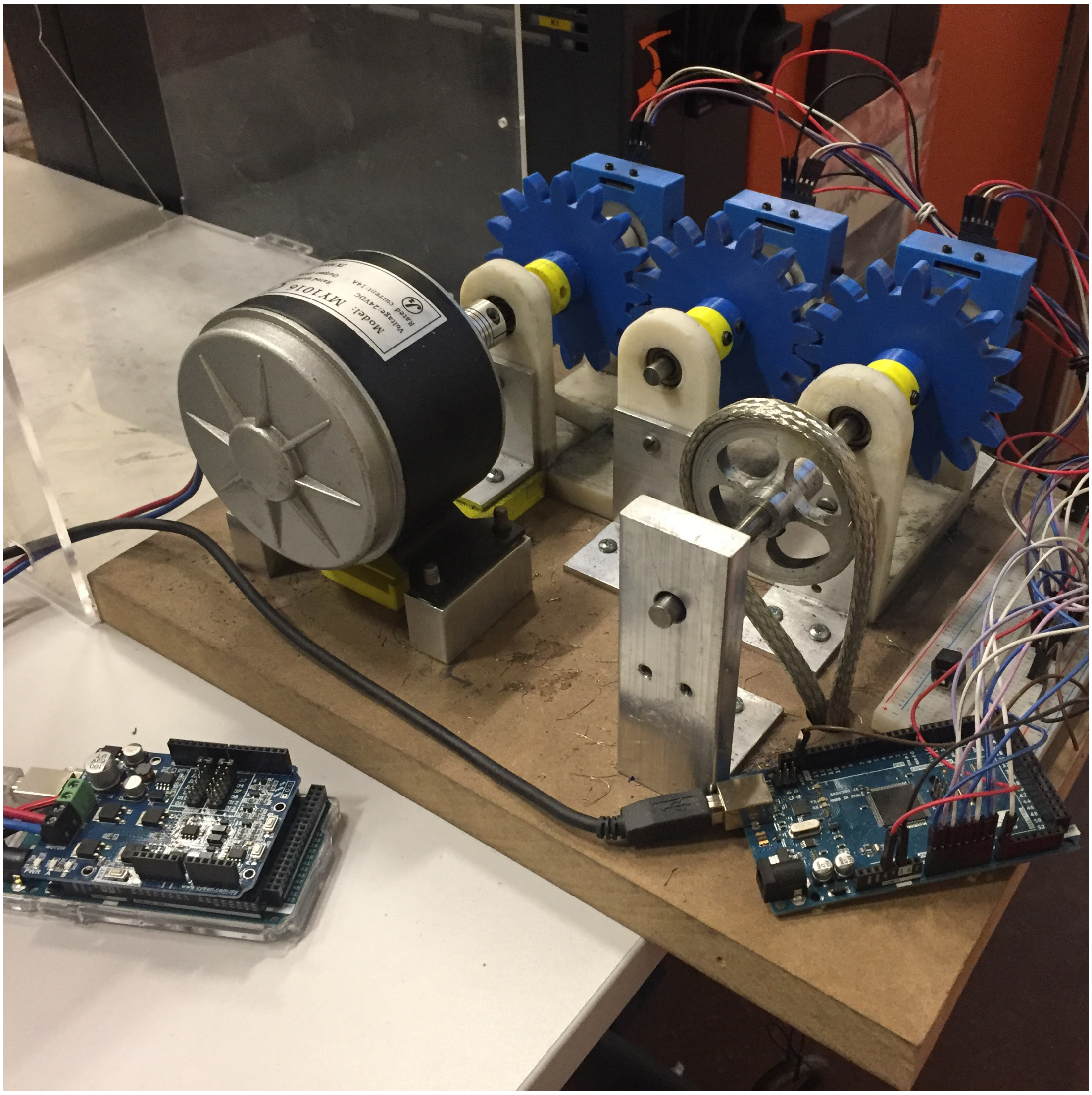

Figure 4.3: Gearbox accelerated life testing platform

The gearbox experimental platform that is considered here is the same as presented in Chapter 3. However it has a different configuration, this is detailed in the following.

The gearbox testing platform is comprised of three gears forming a gear train mounted in series. The gears are arranged as gear 1 (G1) on the left, gear 2 (G2) in the middle and gear 3 (G3) on the right. Each gear is fixed on a shaft. These shafts have restricted translation motion due to a friction-less rotation system.

Such restriction is provided by small washers that are held against the inner ring of the shaft supporting bearings using shaft collars. Friction-less rotation is essential as it prevents additional noise originating from friction and unwanted additional loads. A fixed bracket holds the driving motor. This can be seen on the left of Figure 4.3. This is a 24 Volt, 250 Watt motor that can reach up to 2750 Revolutions per minute (RPM). Feedback from an encoder is collected for extracting exact rotational velocity and steady state behaviour, along with a temperature feedback controller that is used for setting the fail-safe threshold. The gearbox is coupled to a dynamometer system that provides the load.

Vibration analysis has been extensively researched and has become a standard for gearbox system diagnostics and prognostics (137, 211), this is also discussed in Section 3.3.2. Therefore we use three accelerometers, each mounted on one of the three gear supporting shafts, to collect vibration data from the gearbox. This allows the vibration signals of each gear to be distinguished more accurately. The accelerometer signals were transmitted using a data acquisition card (DAQ) to a PC workstation where they are processed. The three accelerometer sensors collect data on three axis and have a full sensing range of $\pm 3Gs$. To avoid distortion of vibration signals, these accelerometers are each mounted over the centerline of the shaft supporting bearing. We do this by fixing these sensors using hex socket screws to a 3D printed housing that lies on top of the frame of the gearbox.

4.3.2.1 Experimental Scenario

To demonstrate the stochastic dependency between components, specifically the degradation state rate interactions, we will only consider a two gear system. Gear 1 and gear 2, referred to as G1 and G2 respectively.

The experimental runs of the gearbox were designed for accelerated life testing, thus achieving failure in a shorter amount of time than it would usually take under normal operating conditions. These runs are an alternating sequence of two types of cycles; the first cycle is a low speed low load cycle, referred to as LSLL; and the second type is a high speed high load cycle, referred to as HSHL.

Due to the nature of the HSHL cycle, a high level of noise is present in the acceleration data. We therefore only use vibration data that are collected in the LSLL cycles in order to improve the signal to noise ratio. These LSLL cycles last for three minutes.

The gearbox platform was run three times, these runs consisted of tests to failure and were conducted in the following manner:

Run 1: The first run consisted of a new G1 and a new G2. The gearbox was run alternating between HSHL and LSLL until high levels of vibration were observed in the gearbox (meshing frequency magnitude exceeding 1800) at which point the experimental run was terminated.

Run 2: After the first run, G1 was replaced with a new gear, while G2 remained unchanged, so the second run consisted of a new G1 and a worn out G2. The gearbox was ran alternating between the HSHL and LSLL cycles until high vibration was observed; on this run high system vibration occurred in a shorter amount of time, and after terminating the run, G2 showed more severe damage on its teeth surface than that observed after the termination of run 1.

Run 3: In the third run, G1 was replaced with a new gear, while G2 remained unchanged, so we find ourselves with a similar condition scenario as in run 2, this time however with a more worn out G2. The gearbox ran alternating between the HSHL and LSLL cycles until high vibration was observed. This run lasted an even shorter amount of time than in run 2, and so the run was terminated earlier than in run 1 and run 2.

The different gear conditions that resulted from these experiments can be seen in Figure 4.4.

4.3.2.2 Component Health State Extraction

Vibration signals were collected from the accelerometers in all three runs. Figure 4.5 shows a two second sample of the raw signals just after the analogue to digital conversion.

The signals were treated following the methodology discussed in section 4.3.1. First, they were input to the outlier removal algorithm so that any outliers can be removed, some might occur due to transmission of the signal between the DAQ and the PC workstation. The data is then centered and filtered using a high pass filter with a cutoff frequency of 180Hz. This is because in this study we are only concerned with the fundamental frequency, and to enable quicker computation. Finally the signal was transformed and represented in G's. The result of this pre-processing is shown over a two second sample in Figure 4.6.

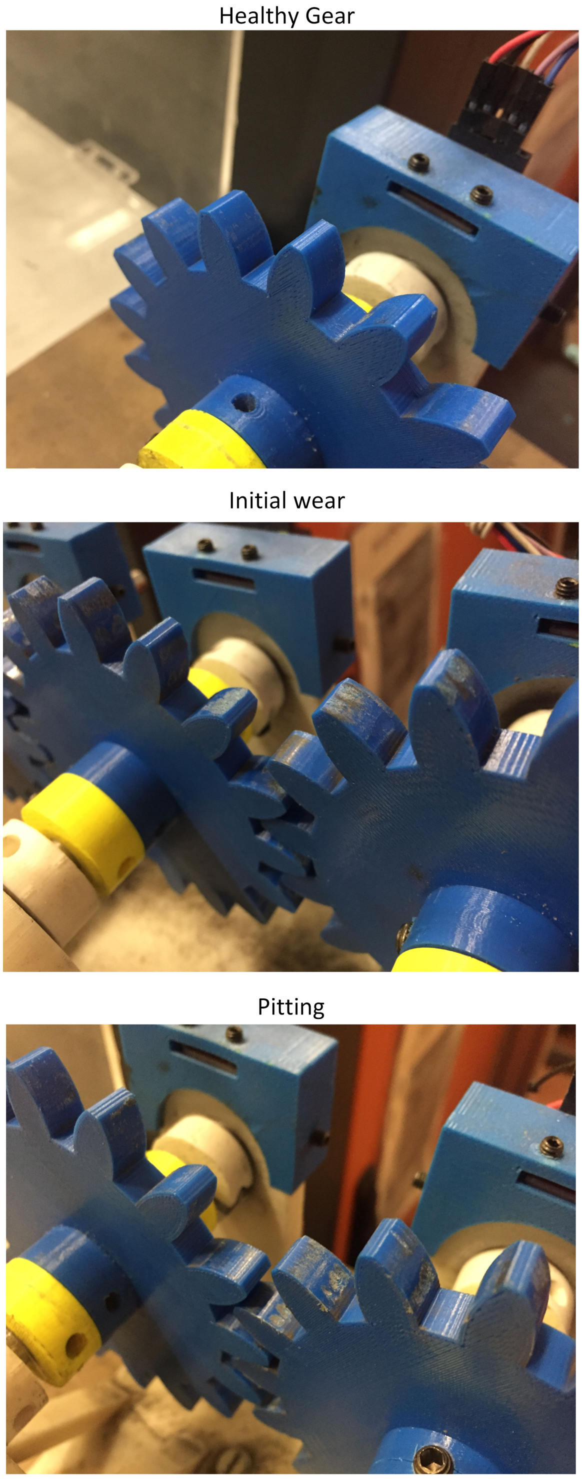

Figure 4.4: The three different gear conditions that were distinguished

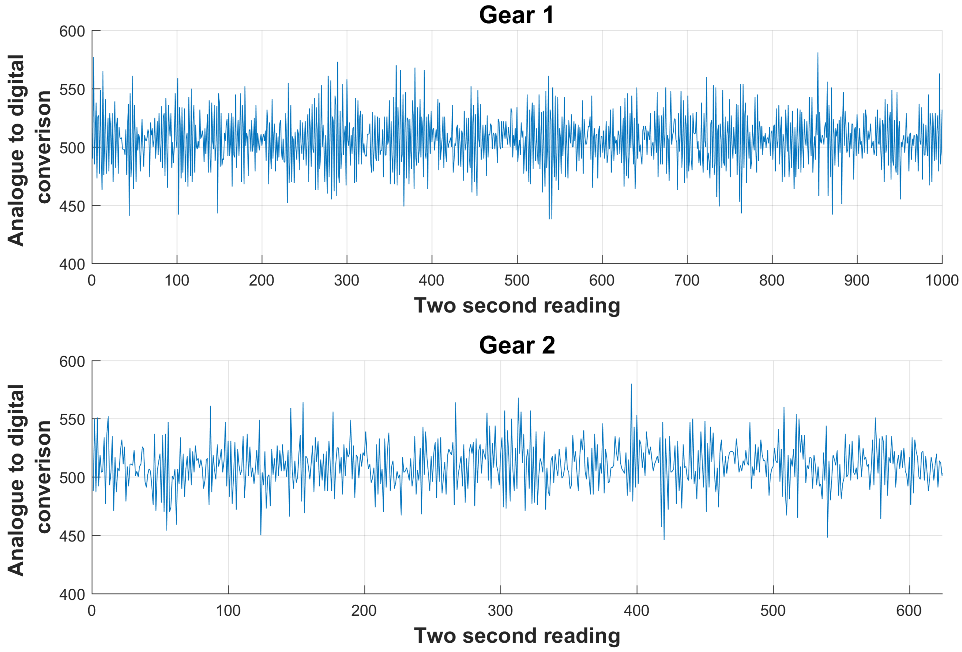

Figure 4.5: Raw accelerometer signals of Gears 1 and 2 after the analogue to digital conversion

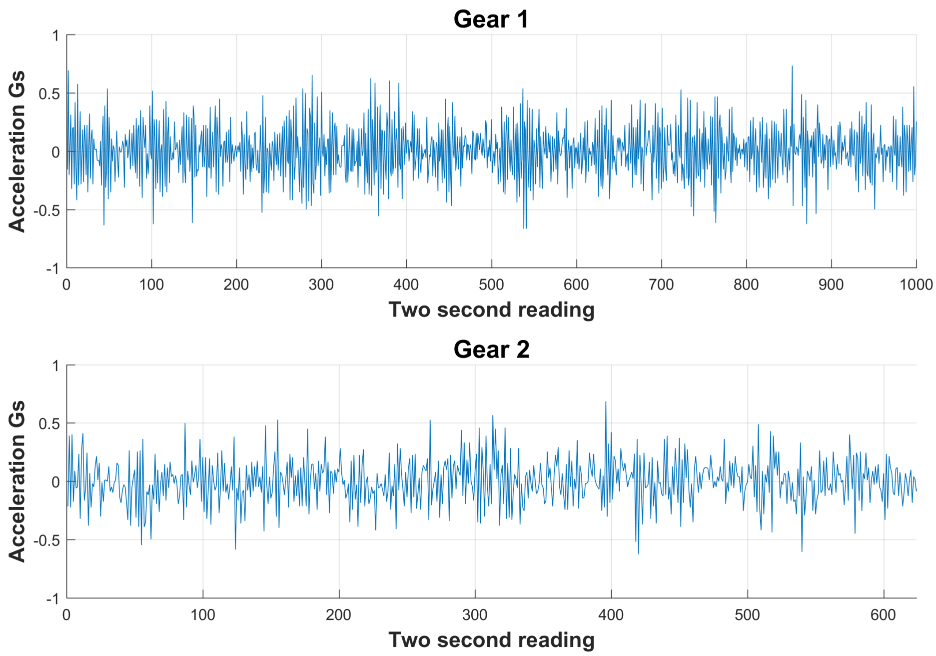

Figure 4.6: Pre-processed accelerometer signals of Gears 1 and 2, represented in Gs.

After the pre-processing steps, the time waveform vibration signals were then turned into time-frequency domain data using STFT. This results in a spectrogram which is shown in Figure 4.7.

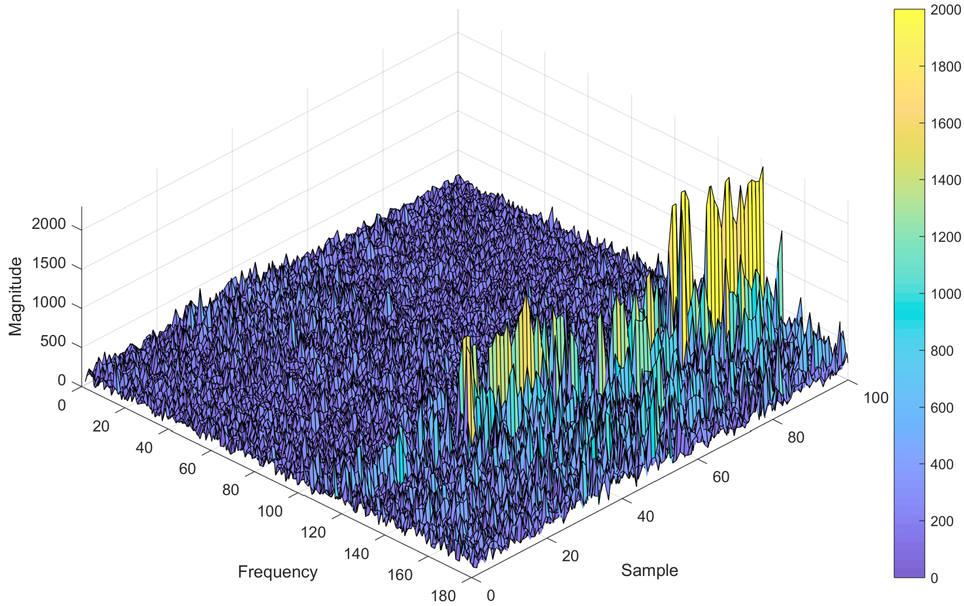

Figure 4.7: Visual representation of the spectrum of frequencies of Gear 1 in Run 1 varying with time

We compute the average SNR of the signal to be 10.6 dB. This is done using the following two equations:

$$SNR = \frac{P_{signal}}{P_{noise}} \tag{4.11}$$

$$SNR_{dB} = 10\log_{10}(SNR) \tag{4.12}$$

Then, we were interested in monitoring the evolution of the gear meshing frequency magnitude f_{mesh} over time. This is computed as:

$$f_{mesh} = RPM \times N \tag{4.13}$$

where RPM stands for revolutions per minute, and N stands for the number of teeth of the gears which is equal to 16 in this case.

Due to the motor speed slightly fluctuating and affecting f_{mesh}, we used a dynamic windowing approach in order to accurately capture the f_{mesh}. We chose a 5Hz frequency band that guarantees to contain the f_{mesh} at its peak magnitude, and so we can then calculate the RMS of that frequency band, and monitor its evolution over time. f_{mesh} is calculated to be around 120Hz.

Here we would like to note that the choice of cutoff frequency requires engineering knowledge of the system at hand. For instance, in this work we use the fundamental meshing frequency to assess the health state of the gears which is computed to be 120Hz; we can therefore remove all frequency elements that are greater that this value without affecting the information that is contained around our frequency of interest. We therefore choose a cutoff frequency of 180Hz for our high pass filter.

We then compute the RMS value for each time step, and the gear mesh frequency is normalised for the range [0 1800]. This results in the degradation time series shown in Figure 4.8. The experimental runs are separated by the black dashed vertical lines, the silver dotted vertical lines represent the start of a new data collection cycle, i.e. an LSLL cycle. Note that a) between every 2 LSLL cycles there exists an HSHL cycle, and b) the HSHL vibration data are not used and is thus not represented in this figure.

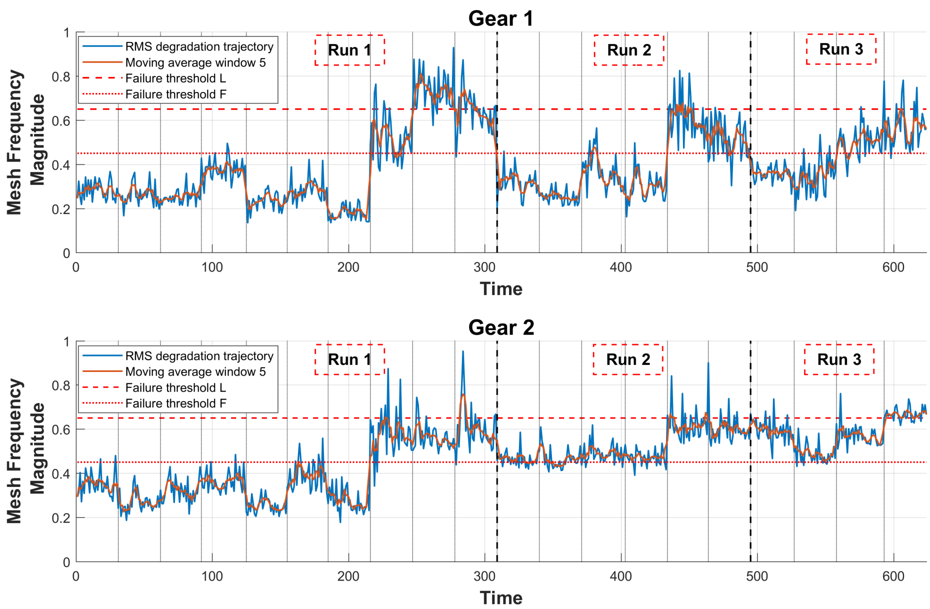

Figure 4.8: RMS degradation trajectories for Gears 1 and 2

4.3.3 Results

Based on the experimental runs, the vibration signals emitted, and the different phases of wear that was observed on the gear tooth surfaces, we can set the two failure thresholds that are seen in Figure 4.8.

Failure threshold F is reached when the gear meshing frequency magnitude reaches 0.45. At this stage the gears show tooth surface wear. This threshold is further motivated by computing the average gear meshing frequency magnitude of each LSLL cycle of G2 in run 2. This does not go below 0.45. At this stage G2 is already considered worn out or faulty, but still operable.

Failure threshold L is reached when the gear meshing frequency magnitude reaches 0.65. At this stage the gearbox platform emits high levels of vibration and is therefore stopped. We consider this to be the hard failure of the system. This threshold is considered to be effectively reached once a moving average of window 5 over the RMS degradation trajectory reaches it.

Now in order to indicate the degradation interactions between the two gears, we compute the average of each LSLL cycle and display the obtained values in Table 4.2. We consider the failure threshold F, since the gears are already faulty at that point.

Note that here the average vibration doesn't necessarily increase at every LSLL cycle, this small fluctuation is due to the change between HSHL and LSLL which can distort the signal acquired by the accelerometers when capturing vibration data. However, we can already see that there is a general trend of increase in vibration with the increase in LSLL cycle count which indicates the degradation of the Gears.

The degradation interactions can already be detected when looking at Table 4.2, however for a better view of the interactions that are taking place, we can look at Table 4.3 which indicates the time to failure of the components. There we can clearly see the accelerated degradation of the Gears that is caused by their interaction.

As shown from Table 4.2, in run 1, it takes seven cycles to reach the G1 failure limit when both gears are new. We can consider this as normal degradation behaviour of the components and so we can say that in this case, the life expectancy of a component when coupled with a new component is 100%. Now looking at run 2, we see that it takes four cycles to reach the G1 failure limit when G1 is new and G2 is worn out. Thus compared to run 1, where both gears were

		LSLL Cycle Number									
Run	Gear	1	2	3	4	5	6	7	8	9	10
1	1	.283	.260	.257	.373	.230	.279	.183	.518	.735	.624
1	2	.347	.267	.325	.360	.253	.388	.283	.580	.560	.602
2	1	.314	.255	.366	.303	.606	.507				
2	2	.464	.465	.489	.477	.626	.604				
3	1	.358	.344	.507	.571						
3	2	.595	.490	.570	.667						

Table 4.2: Average Gear meshing frequency magnitude for each LSLL cycle for both gears in all there runs.

	Gear 1	Gear 2	Gear 1 Cycles to Failure	Gear 1 Life Expectancy (%)
Run 1	new	new	7	100
Run 2	new	worn out	4	57
Run 3	new	severely worn out	2	29

Table 4.3: Effect of deterioration on component interactions

new, we see that having a new component coupled with a worn out component would lead to accelerated wear of the new component and so the life expectancy is reduced to only 57%, in this case, in comparison with normal degradation of the components. Finally in run 3 we see that it only takes two cycles until G1 reaches its failure limit when G1 is new and G2 is severely worn out. This means that in comparison to normal degradation behaviour, G1 has in this case, a life expectancy of 29% relative to that under normal degradation. This is clearly shown in summary in Table 4.3.

These results clearly demonstrate the importance of modelling stochastic dependency between components when performing prognostics on a multi-component system. For if we are to replace a specific component in the system with a new one, ignoring the accelerated degradation effect that results from it being coupled with a now worn out component, there would arise unexpected failures and faults. These would be caused by the reduced lifetime of the new installed components that are not degrading in nominal fashion.

4.4 Discussion

In this chapter we started by providing an introduction and background on health indicator extraction for PHM. We covered data acquisition, signal pre-processing and signal processing approaches, while providing a comprehensive review of the literature in the context of PHM.

We then presented our developed methodology for extracting accurate health indicators for components of a multi-component system. This started with the data collection process and went through the selection of a time-frequency domain analysis for processing the waveform data that should be collected. Finally, this resulted in a time series signal representing the degradation trajectory for each component in the system. We validated this approach and demonstrated our analysis on experimental data collected from a gearbox accelerated life testing platform. Here we showed that when a new gear is coupled with a worn out gear, the life expectancy of the new gear may be reduced to 29% of that of a new gear coupled with another new gear. Through this work we demonstrated the importance of accounting for stochastic dependence, since degradation dependence between old-new component couplings can ultimately lead to accelerated wear of the system.

Chapter 5

Degradation Modelling and Prognostics in Multi-Component Systems

Contents

5.1 Chapter Summary

This chapter is concerned with the degradation modelling of multi-component systems, and their prognostics. We start by presenting a review on the recent literature on stochastic dependence, which motivates the work presented here. This is then followed by the description of our generic degradation model for multi-component systems. This is used to generate simulated degradation trajectories of a three component system. The parameters of the model are fitted to the data generated by the gearbox platform as presented in Chapter 4, this is achieved using a particle filter which is considered as a state-of-the-art technique in the PHM literature. The importance of using this technique is shown through an overview and background, and a literature review of its use within PHM. The results are then presented which show the advantages of modelling stochastic dependence between components.

5.2 Introduction

The ever increasing number of manufacturing requirements is pushing original equipment manufacturers (OEM) to design more complex systems to meet industrial needs. Such machines are becoming increasingly difficult to maintain (276, 277). This is especially due to their degradation processes which are highly stochastic in nature. These degradation processes limit the accuracy of diagnostics and prognostics, which in turn leads to poor remaining useful lifetime (RUL) predictions. This incurs an increase in the number of unforeseen faults and failures, and a reduction in the reliability of multi-component systems in industrial environments.

Consider for example a system with two components, an induction motor with a lifetime, say, up to 5 years that is coupled with bearings that have a lifetime that is but a fraction of this. In many multi-component systems like these, it is almost inevitable that after running the system for a long enough time, old worn out components are then coupled with new healthy components. And since old worn out components may potentially accelerate the degradation process of new components, it is in effect this old-new component coupling that affects the reliability of a system, and leads to a system wearing out in an unforeseen accelerated fashion. Thus, modelling degradation interactions such as old-new component couplings in multi-component systems can play a crucial role in diagnosing and performing prognostics on the health state of a system.

However, the degradation processes of components in complex system are usually assumed to be independent, see (35, 176, 248). But since real world systems typically include multiple interacting components, it is very unlikely that there exists no interactions between the components themselves. And, as mentioned earlier, this can severely jeopardise system reliability and overall availability.

And although condition based maintenance (CBM) research is showing a growing interest in multi-component systems (125), rarely are they considered in a prognostics and health management (PHM) context. And although the issue of stochastic dependence is actually addressed in CBM, this is usually not done with the aim of modelling it (178).

5.2.1 Literature Review

Further to the overview on stochastic dependence present in Section 2.4, we now review the recent and relevant literature on the topic, this serves as motivation to the work that is present in this chapter.

In (157) the authors develop a CBM policy for systems with multiple failure modes. They consider that failures can occur before reaching a maintenance threshold, and that the failure rate of components can be influenced by the age of the system, the overall state of the system or both. This work however does not model the degradation dependence between components and focuses mainly on the CBM policy rather than degradation modelling itself. In (97) the authors present a methodology for mixed signal separation of identical components using independent component analysis (ICA), they specifically consider the case where there exists limiting constraints over sensor placement. Then, they use the separated signals as indicators of degradation severity for each of the components, and validate their approach via a numerical example using simulated degradation signals. Although this work is an essential step for modelling the degradation of such multi-component systems, it does not specifically model the degradation dependence between components. In (29) the authors use stochastic differential equations to model the degradation between components. They specifically study how the degradation rate of one component can be influenced by the degradation state of other components with the aim of predicting the residual lifetime of components. They then evaluate their approach using simulated data and compare it to a benchmark approach which assumed component degradation independence.

Finally they show the importance of capturing degradation dependencies between components. Although this work deals with deterioration dependence between components, it does not account for other interfering factors. In (152) the authors discuss what they refer to as the fault propagation phenomenon. This is described as the co-existence of inherent dependence and induced dependence when considering degradation dependence between components. A continuous time Markov chain approach is developed to capture fault propagation characteristics. However this approach might suffer from the state-space explosion problem and does not consider other factors that may influence degradation. Therefore it does not describe the full underlying mechanism of system degradation. In (205) and (65) the works consider state-rate degradation interactions. Both works study two component systems where they either consider a numerical simulation of a degradation process in (65), or perform degradation modelling for the particular case at hand in (205). They both use the results for optimising the CBM policy. These works mainly deal with the optimisation of the CBM policy rather than developing a general degradation model for interdependent components.

Further to the works mentioned above, and considering the extensive body of literature on degradation modelling, see (258) and (248) for an overview. The literature shows that the degradation process of components may depend on the system operating conditions such as the load on the system, vibration, humidity, temperature etc. see (18, 60, 228) and (66). And it is also shown that a degradation process system may depend on its current state, see (224).

We therefore aim to develop a generic multi-component degradation model whereby the degradation process of a component may be dependent on the operating conditions, the component's own state, and the state of the other components. The model's capability will be first demonstrated through a numerical simulation. We then show how to fit the model to data using particle filter (PF). This method is then used on the degradation trajectory data generated by the gearbox, as seen in Section 4.3.3.

5.3 Generic Multi-Component Degradation Model

In this section we present our generic degradation model. It has the capability of encompassing multiple component stochastic dependencies all while considering the operational effect, and the intrinsic degradation effect.

5.3.1 Degradation Model

Consider a multi-component system with n_c number of components. The degradation state of each component i is represented by an accumulation of wear over time which is assumed to be described by a scalar random variable X_t^i. Component i fails if its degradation state reaches a threshold value L^i. If any of the components fail we consider the system to have failed, and if a component is not operating for whatever reason, no change occurs to its degradation state unless a maintenance intervention is carried out.

We assume the evolution of the degradation state of component i is represented by:

$$X_{t+1}^i = X_t^i + \Delta X_t^i \tag{5.1}$$

where ΔX_t^i represents the degradation increment of component i during one time step.

The degradation of a component i at time step t may depend on the operating conditions, the state of component i, and also the state of other components to a varying degree. Thus we suggest a general stationary model for the increment ΔX_t^i:

$$\Delta X_t^i = \Delta O_t^i + \Delta X_t^{ii} + \sum_{j \neq i} \Delta X_t^{ji} \tag{5.2}$$

where:

- ΔO_t^i represents the degradation increment of component i that is caused by the operating conditions during one time step t. ΔO_t^i can be specified as a deterministic or as a random variable.

- ΔX_t^{ii} represents the degradation increment which is intrinsic to i at time step t. In other words ΔX_t^{ii} depends on the degradation state of component i at time step t. ΔX_t^{ii} can also be specified to be a deterministic or random variable.

- $\sum_{j \neq i} \Delta X_t^{ji}$ represents the sum of all degradation increments which are caused by the interaction of component i with the other components of the system. The degradation interaction between a component i and another component j may be considered to be a deterministic or random variable.

We can now specify different variants of the proposed model:

Case 1: $\Delta O_t^i > 0$, $\Delta X_t^{ii} = 0$ and $\Delta X_t^{ji} = 0$, in this case there is neither an intrinsic nor an interaction effect, and so the proposed model is reduced to a model of homogeneous degradation behaviour of independent components as seen in (248).

Case 2: $\Delta O_t^i > 0$, $\Delta X_t^{ii} > 0$ and $\Delta X_t^{ji} = 0$, in this case there is no degradation interaction between the components, and so the proposed model becomes a model describing non-homogeneous degradation behaviour as seen in (224).

Case 3: $\Delta O_t^i = 0$, $\Delta X_t^{ii} = 0$ and $\Delta X_t^{ji} > 0$, in this case the components have degradation inter-dependencies only, and the proposed model corresponds to the degradation model that was introduced in (205).

Case 4: $\Delta O_t^i > 0$, $\Delta X_t^{ii} = 0$ and $\Delta X_t^{ji} > 0$, in this case the components have degradation inter-dependencies but no intrinsic degradation is present; this case then corresponds to the models presented in (29) and (65).

Case 5: $\Delta X^{ii} > 0$ and $\Delta X^{ji} > 0$, the components are stochastically dependent and the increment in the degradation level of component i may depend not only on the state of component i but also on the state of the other component; this case is considered in (27) for prognostics of a system lifetime, in which a random intrinsic effect is considered and described by a Brownian motion process. The interaction effects were considered deterministic and proportional to the degradation level of other components.

Case 6: $\Delta O_t^i > 0$, $\Delta X_t^{ii} > 0$ and $\Delta X_t^{ji} > 0$, in this case the degradation processes of the components are dependent on the interaction between the components, the intrinsic degradation of the components and the operating conditions of the system.

We will use the following model for the quantification of degradation influence between multiple components:

$$\Delta X_t^{ji} = \mu^{ji} \times (X_t^i)^{\sigma^{ji}} \tag{5.3}$$

where X_t^{ji} represents the degradation impact of component j on component i at time t. And where μ^{ji} and σ^{ji} are non-negative real numbers which are used

Case	Description
$\mu^{ji} = 0$	Component j does not have any influence on the degradation behaviour of component i
$\mu^{ji} = 0$ and $\mu^{ij} = 0$	Component j and i are independently subject to gradual degradation
$\mu^{ji} > 0$ and $\sigma^{ji} = 0$	the impact of component j on the degradation of component i does not depend on the health state of component j

Table 5.1: Degradation influence between multiple components

to quantify component j's influence on component i. This is represented more clearly in Table 5.1.

The effect of component j on component i does not have be similar to the effect of component i on component j. And especially when considering systems with a $n_c > 2$ number of components, the degradation influences μ^{ji} and σ^{ji} between the components of the system can be better represented using square hollow matrices of size $n_c \times n_c$ as seen in Eq. 5.4.

$$\mu^{ji} = \begin{pmatrix} 0 & \mu^{12} & \dots & \mu^{1n} \\ \mu^{21} & 0 & \dots & \mu^{2n} \\ \vdots & \vdots & 0 & \vdots \\ \mu^{n1} & \dots & \dots & 0 \end{pmatrix}, \ \sigma^{ji} = \begin{pmatrix} 0 & \sigma^{12} & \dots & \sigma^{1n} \\ \sigma^{21} & 0 & \dots & \sigma^{2n} \\ \vdots & \vdots & 0 & \vdots \\ \sigma^{n1} & \dots & \dots & 0 \end{pmatrix} \tag{5.4}$$

This can also be extended in the sense where hollow matrices are not used for μ^{ji} and σ^{ji}, but where the diagonals entries of the matrices are occupied by μ^{ii} and σ^{ii} which represent the intrinsic degradation influence of the components on themselves; i.e. the intrinsic degradation rate of the component might depended on the degradation level of the component itself, as would be the case when specific protection coatings of components start to fade (144).

Although the proposed degradation model can encompass as many components as we would like, the number of components to be considered when performing degradation modelling should be kept to the minimum. This is because adding more components would lead to an increase in model complexity and computation, and the composition of optimal CBM policies becomes more difficult and computationally complex. This is an already identified issue in multi-component maintenance literature (4).

5.3.2 Simulation

For the purpose of illustrating the interactions that can influence the degradation process of components in a multi-component system, we will use **Case 5** from the generic degradation model to create a numerical simulation of the degradation process of a three component system.

Since the degradation of most mechanical components accumulates wear over time, we can then use a gamma process to represent this. This is because it is a stochastic process with independent, non-negative increments, see Appendix A.1 for further details on the gamma distribution. It is therefore well suited to model the gradual degradation which accumulates over time as seen in (187, 248). And so for every component i among the n_c components of the system, we assume that the corresponding ΔX_t^{ii} follows a gamma distribution with distinct parameters, shape α^i and scale β^i as shown in:

$$f_{\alpha^i,\beta^i} = \frac{(\beta^i)^{\alpha^i}}{\Gamma(\alpha^i)} x^{\alpha^i - 1} \exp^{-\beta^i x} \tag{5.5}$$

based on the generic degradation model presented in the previous section, we can now model the degradation of the system at hand as follows:

$$X_{t+1}^i = X_t^i + \Delta X_t^i \tag{5.6}$$

$$\Delta X_t^i = \Delta X_t^{ii} + \sum_{j \neq i} \mu^{ji} \times (X_t^j)^{\sigma^{ji}} \tag{5.7}$$

So a simulation can be initiated using the gamma process parameters which are shown in Table 5.2 for ΔX_t^{ii}, and the values for the degradation influence between components μ^{ji} and σ^{ji} shown below. This then generates degradation trajectories as seen in Figure 5.1.

$$\mu^{ji} = \begin{pmatrix} 0 & 0.254 & 0.1080 \\ 0.384 & 0 & 0.346 \\ 0.242 & 0.118 & 0 \end{pmatrix}, \quad \sigma^{ji} = \begin{pmatrix} 0 & 0.54 & 0.7290 \\ 0.785 & 0 & 0.836 \\ 0.838 & 0.555 & 0 \end{pmatrix}$$

From Figure 5.1, we can see the normal degradation trajectories of all 3 components from time step 1 till 40 since the system is considered to have started with all components having a healthy new state. We can now compare the normal degradation trajectory of component 1 with its accelerated degradation after

	Parameter Value		
	Component 1	Component 2	Component 3
Shape α	4.944	4.35	5.193
Scale β	3.919	1.09	2.257

Table 5.2: Simulation parameter values

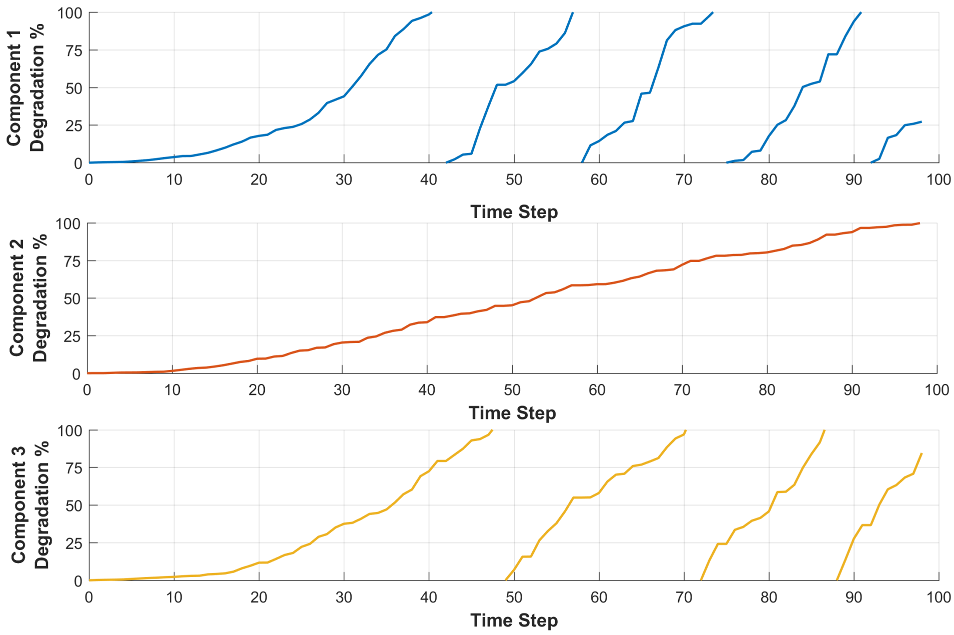

Figure 5.1: Illustration of degradation evolution with rate-state interactions

being replaced at time step 42 and being coupled with the other two worn out components. We can clearly see two phases of accelerated degradation after component 1 has been replaced with a new component. A highly accelerated degradation from time step 42 till 48, then a somewhat less accelerated degradation from time step 49 till 56. The first highly accelerated degradation is due to the fact that a new component 1 was interacting with a worn out component 2 and a severely worn out component 3 that was above 75% degraded; subsequently, the less accelerated degradation is a result of the replacement of component 3 with a new component at time step 49.

This old-new component coupling is clearly influencing the lifetimes of the components after being replaced with new ones. This kind of interaction can lead to accelerated degradation of the components and the system as a whole, thus resulting in unexpected faults and failures.

The degradation trajectories simulated using our generic degradation model are similar to the ones obtained from the experimentation using the gearbox accelerated life testing platform, which were presented and analysed in Section 4.3.3.

5.4 Prognostics for Multi-Component Systems

Prognostics is the process whereby past and present condition monitoring data of a system or component is used to project its health state into the future. This is done with the specific aim of predicting the end of life (EOL) of components, and consequently, the estimation of remaining useful life (RUL) (216).

Many prognostics methods have been developed in the last years. A thorough review has been presented in Chapter 2, additionally literature review articles on the topic can be found in (3, 135).

However, the comparison between different prognostic methods is not an easy task. Generally speaking, the are no metrics that are universally accepted to quantify the benefit of a prognostics method. Nonetheless, methods for quantifying the performance of prognostics algorithms have gained particular attention in the last years (216, 217, 136). In this work, we will use the difference between real EOL and the predicted EOL to assess the prognostics performance of the proposed approach.

The main objectives of prognostics can be summed up with the following:

- Estimate the degradation of components at present time $t = k$

- Estimate the degradation of components at a future time $r > k$
- Estimate the RUL of components $RUL = t_{eol} - k$. where t_{eol} denotes time instant where the degradation prediction of a component crosses the failure threshold.

The generic degradation model that was presented in Section 5.3 can be used for performing predictions of t_{eol} for multi-component systems. The procedure for performing prognostics using any of the variants of the model starts by the proper selection of the variant type, and specifying its different parameters. Parameter identification then follows and can be done using different approaches. We suggest the use of particle filter, and this will be motivated in the next section.

Once the model parameters are identified, the model can be easily used to simulate and predict the health state of a component $X^i_{t_k}$ at a future time $r > k$. This is done until the degradation trajectory hits the failure threshold which indicates t_{eol}. At that point the RUL can easily be extracted, and a maintenance decision can be taken accordingly.

5.4.1 Particle Filter in PHM

There exists an extensive body of literature on the topic of parameter identification, see for example (7, 83, 158). In practice, if the degradation model is not too complex, we can fit the model parameters using maximum likelihood estimation (MLE). However, in the case of multi-component systems this is highly unlikely. Therefore if the model is too complex, or if we are collecting online observation on the health condition of the components and want to achieve real time prognostics, we suggest to use sequential Monte Carlo methods, specifically the PF method which is a very popular approach for parameter estimation (71).

PF allows for an online numerical estimation of the parameter values by means of a recursive Bayesian inference approach. The posterior distribution of the model parameters can be then obtained using a number of particles and their corresponding weights. This method is very flexible and can be used for non-linear models where the noise is not necessarily Gaussian. Such an approach has been successfully used in the field of prognostics for model parameter estimation, and for prognostics as seen in the following works:

In (183) the PF approach has been successfully applied to show faulty axial crack growths in an UH-60 planetary carrier plate. The authors use 2 sequential modules. The first module is the state-space model together with a PF algorithm that is used to evaluate the PDF of the system and the probability of a fault. The other module is used for the prediction of the fault and the RUL, employing unknown time-varying parameters and a PF algorithm which updates the state estimates. In (236) the issue of the number of particles to be used is addressed and proposed as a trade-off between prognosis performance and computational costs. The state model and the noise PDF parameters are adjusted via a feedback loop, and short-term predictions are used to improve the generation of long-term predictions. The approach was validated for predicting crack growth on a test coupon based on different fault mode assumptions. In (293) the proposed PF is implemented together with other estimators to improve the accuracy of the method. This is used for the estimation of the evolution of a nonlinear fatigue crack growth. The authors advise that some challenges remain open:

- The state estimators and predictors should carry a measure of the error, which is important to foresee the confidence on the predictions and thus the actions to take.
- As in general processes change with time, the prognostic approach should account for these changes and work well in different working conditions.
- When considering timely prognostics, the computational costs must be flexible enough depending on the application.
- It is important to consider the existence of multiple fault conditions.

Battery condition effectively degrades with time, and thus entails reduced system performance and economic loss. Therefore predicting battery functionality has attracted many researchers in the field of PHM, see for example the works in (88, 119, 181).

The work presented in (168) introduces an improvement to the PF algorithm where the unscented particle filter (UPF) uses as proposal distribution the results obtained from an unscented Kalman filter (UKF). However, the degradation model is built based on the particular application of lithium-ion batteries. Also in (54), the authors detail the implementation of a particle filter based prognostic in the case of lithium ion batteries. For the state model, a lumped parameter battery model which considers the dynamic characteristics and internal processes of the battery has been used. In more recent work (107) on lithium ion batteries prognostics, it is assumed that degradation trajectories in similar components are unavailable. They use PF and kernel smoothing-based approaches together in order to solve the problem. In (94), the authors compare the capabilities of four different re-sampling algorithms: multinomial re-sampling, residual re-sampling stratified re-sampling and systematic re-sampling. Data sets of lithium-ion batteries from NASA data repository are used to analyse and compare the results. The same author in (95) proposes the F-distribution particle filter approach by dynamically adjusting the particles' weights through the F kernel and demonstrates the feasibility of the approach on real data generated by a hydraulic actuator. In order to update the weights, their approach uses historical observations. Other works on battery prognostics using PF can be found in (100), (267).

The work in (284) relies on using vibration data from rolling element bearings of a helicopter's oil cooler for an on-line parameter adaptation solution. The work is an attempt to approach multi-fault modelling using PF. Also in (182), vibration data from a cracked gearbox plate in a critical aircraft component is used with PF assisted with regularisation algorithms. Vibration data is a good indicator of the failure time of a bearing (198), since its natural frequency and acceleration amplitude can be related to its damage mechanics.

Another interesting approach for PF in PHM uses Rao-Blackwellised PF for estimation of the parameters in the railway vehicle dynamic model (150). In (220) the authors use a hidden Gamma process model for capturing the necessary health features. Then, an approach based on PF is used and applied in the context of semiconductor manufacturing on generated synthetic helium flow signals.

5.4.2 Proposed Degradation Model using Particle Filter

Particle filter draws upon stochastic filtering, Bayesian statistics and Monte Carlo techniques. It is usually also referred to as sequential Monte Carlo, however it should be distinguished from sequential Monte Carlo (SMC), since SMC methods encompass a broader range of algorithms (71), such as the well known Gibbs sampling and Metropolis-Hastings algorithms. PF is also called a Bootstrap filter (93), this is the case of the standard particle filter algorithm which is presented in Algo. 2, and which is implemented in this work.

It is worth noticing that compared to typical works on PF, whereby filtering is mainly considered, prognostics concerns itself with future time horizons, this means that this field tries to go beyond the filtering step. In view of this, PF for prognostics should be used in accordance to the necessity of forecasting the state at future times, mostly without additional observations, adjusting the weights if necessary. Moreover, recent reviews about PF for PHM such as (119) suggest an increasing amount of work on PF in PHM. Therefore this approach is considered as a state-of-the-art technique in PHM.

Some common requirements for performing prognostics are presented in the following:

- Some measures of the degradation status
- A state model that deals with component degradation
- A measurement equation
- A fault detection threshold

In this work measures of the degradation status of multi-component systems can be obtained using the health indicator extraction method described in Chapter 4. We assume measurements are still noisy, accordingly we can add a noise term to the measurement equation. As for the state model, we propose the use of the generic degradation model that is described in section 5.3. Regarding the failure threshold, this is system specific, and needs to be set based on previous runs to failure of a system, or based on domain knowledge.

Consider our goal is to estimate the parameters of a degradation model that corresponds to **Case 5**, and where two components are present. Let's assume the intrinsic effect is stochastic and follows a gamma distribution $\Delta X_t^{ii}\ is\ i.i.d.\ \Gamma(\alpha^i, \beta^i)$. Furthermore, to be coherent with the notation present in the particle filter literature, we denote the future state X_{t+1} as X_t, and the current state X_t as the previous state X_{t-1}.

Then the deterioration model can then be rewritten as:

$$X_t^i = X_{t-1}^i + \Gamma(\alpha^i, \beta^i) + \mu^{ji} \times (X_{t-1}^j)^{\sigma^{ji}} \tag{5.8}$$

In this case there exists two sets of parameters Θ^1 and Θ^2. Where $\Theta^1 = (\alpha^1, \beta^1, \mu^{21}, \sigma^{21}, \epsilon)$ and $\Theta^2 = (\alpha^2, \beta^2, \mu^{12}, \sigma^{12}, \epsilon)$, with ϵ representing the variance of the observation noise which can be assumed to be Gaussian. For each set of parameters we generate a specific n_p number of particles, each having 5 parameter values selected at random from a prior distribution. We then generate a prediction of the next health condition $\tilde{X}_t^{i,n}$ for $n = 1 : n_p$. After observing the next health condition y_t^i we can calculate the importance weight of each particle by computing the likelihood of that observation given the predicted values of each particle $p(y_t|\tilde{X}^{i,n}_{t+1})$. We then normalise the weights and perform bootstrap-importance sampling i.e. we re-sample with replacement n_p particles from the previous set of particles according to their weights. We then repeat the process using the new set of particles. This is shown in Algorithm 2. And a generic PF approach using 10 particles is illustrated in Figure 5.2.

Algorithm 2: Particle Filter Algorithm

---

input : n_p number of particles
Initialisation
$t = 0$
for $i \leftarrow 1$ **to** n_p **do**
    Sample $x^i_0 \sim p(x_0)$
end
for $t \leftarrow 1$ **to** t_{end} **do**
    Importance Sampling
    for $i \leftarrow 1$ **to** n_p **do**
        Sample $\tilde{x}^n_t \sim p(x_t|x^i_{t-1})$
        Set $\tilde{x}^i_{0:t} = (x^i_{0:t-1}, \tilde{x}^i_t)$
    end
    for $n \leftarrow 1$ **to** n_p **do**
        Evaluate importance weights $\tilde{w}^n_t = p(y_t|\tilde{x}^n_t)$
    end
    Normalise importance weights $\tilde{w}^n_t$
    Particle Selection
    for $n \leftarrow 1$ **to** n_p **do**
        Considering $\tilde{w}^n_t$, re-sample with replacement n_p particles
    end
end

5.5 Case Study

Here we fit the generic multi-component degradation model to the data generated from the gearbox accelerated life testing platform presented in Chapter 4. We refer to gear 1 and 2 as component 1 (C1), and component 2 (C2) respectively.

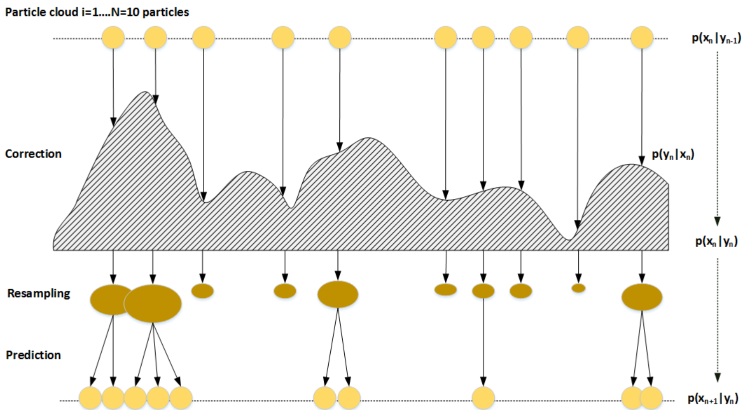

Figure 5.2: A generic illustration of a particle filter

5.5.1 Model Parameter Estimation

After performing the health indicator extraction step in Chapter 4, we acquire the RMS degradation trajectories for C1 and C2. These are presented in Figure 5.3. The RMS values inform us of the vibration energy in the machine that originates from the gears. Therefore they represent the degradation level of the components since the higher the vibration energy, the more the gears are degraded and the more prone the gearbox is to damage. This proness of a gear to damage is the manifestation in reality of the terms ΔX_t^{ii} and ΔX^{ji} in the model Eq. 5.8, the former because the gear itself is worn, and the latter because the other gear is worn. We consider a component to be severely worn out or to have failed once it reaches the threshold vibration magnitude of $L^i = 0.65$ for $i = 1, 2$ as described in Section 4.3.3.

Due to the physical characteristics of the gears, we know that the degradation level of components C1 and C2 increases with time, and that this degradation level cannot decrease without maintenance intervention. Therefore, both components are considered to have inherent degradation that increases with time. Consequently we assume that these degradation increments are gamma-distributed, see Appendix A.1 for more detail. These increments are denoted by ΔX^{11} and ΔX^{22} for C1 and C2 respectively. Thus, $\Delta X^{11} \sim \Gamma(\alpha^1, \beta^1)$ and $\Delta X^{22} \sim \Gamma(\alpha^2, \beta^2)$.

Next, we model the degradation interactions between the two components. From Figure 5.3 it appears that the state of C2 affects the rate of degradation of C1. This can be seen when we observe the time to failure of C1 when coupled with a worn out C2 in both runs 2 and 3, and that in run 3, where C2 was more worn out, the time to failure of C1 was shorter than run 2. Thus the degradation

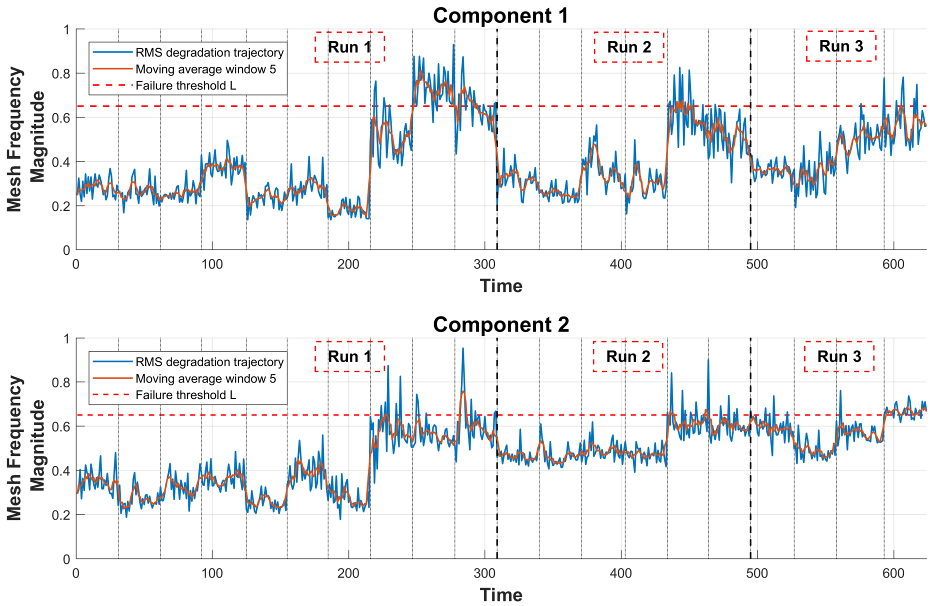

Figure 5.3: Evolution of degradation of the gears in all three runs, represented by the mesh frequency magnitude

rate of C1 appears to be dependant on the degradation level of C2 and vice versa. This has been further analysed in Section 4.3.3.

ΔX^{21} is used to denote the increment in the degradation level of C1 due to C2, and ΔX^{12} the increment in the degradation level of C2 due to C1.

We denote the degradation states for C1 and C2 at time t by X_t^1 and X_t^2 respectively. Further, since in our model $\Delta X^{ii} > 0$ and $\Delta X^{ji} > 0$, the generic degradation model can be seen as **Case 5** of the different variants presented in Section 5.3. Thus, the evolution of degradation for C1 is described as:

$$\begin{aligned} X_t^1 &= X_{t-1}^1 + \Delta X_t^1, \\ \Delta X_t^1 &= \Delta X^{11} + \Delta X^{21}, \\ \Delta X_t^1 &= \Gamma(\alpha^1, \beta^1) + \mu^{21} \times (X_{t-1}^2)^{\sigma^{21}}. \end{aligned} \tag{5.9}$$

and for C2 as:

$$\begin{aligned} X_t^2 &= X_{t-1}^2 + \Delta X_t^2, \\ \Delta X_t^2 &= \Delta X^{22} + \Delta X^{12}, \\ \Delta X_t^2 &= \Gamma(\alpha^2, \beta^2) + \mu^{12} \times (X_{t-1}^1)^{\sigma^{12}}. \end{aligned} \tag{5.10}$$

There are four parameters to be estimated for each component from the data, these sets of parameters are denoted by Θ^1 and Θ^2. Where $\Theta^1 = (\alpha^1, \beta^1, \mu^1, \sigma^1)$ and $\Theta^2 = (\alpha^2, \beta^2, \mu^2, \sigma^2)$. We use the PF method presented in Section 5.4.2 to estimate these parameters. In this case we use $n_p = 1000$

We obtain the mean estimated value of each parameter in Table 5.3. Note that since the degradation level is normalised between 0 and 1, the greater the value of the parameter $sigma^i$ the smaller the impact that is to be considered from the other component on component i.

Component	α^i	β^i	μ^i	σ^i
C1	0.0233	0.0425	0.0995	7.6659
C2	0.0125	0.0914	0.0493	9.7375

Table 5.3: Estimated parameter values

To further validate the parameter values of the degradation model considering the interactions between the 2 components, we compute the R^2 values for the fit of the average estimated degradation trajectory resulting from the particle filter to the real degradation trajectories. For component 1 this is $R_1^2 = 0.792$ and for component 2 it is $R_2^2 = 0.753$. If we were to consider a reduced model whereby no stochastic dependence is considered between the two components and we were left with a gamma process describing the evolution of the degradation level, the average fit of such models would result in a $R_1^2 = 0.671$ and $R_2^2 = 0.575$. The further advantage of considering the interactions between components is motivated in section 7.5.3.

Figure 5.4 shows the particle filter fit to the degradation data of run 1 for both components 1 and 2.

In Figure 5.4, the silver dots represent the estimated degradation level at each time step according the n_p different particles. The yellow dashed line represents the average of these n_p estimates. In green, a moving average of window 5 is applied for both smoothing the observed degradation trajectory and providing effective indication that a component has failed once it reaches the failure threshold L.

5.5.2 Predicting End of Life of Components

Here we use the degradation model with the estimates obtained in Table 5.3 and generate 1000 simulation using the model in order to predict the degrada-

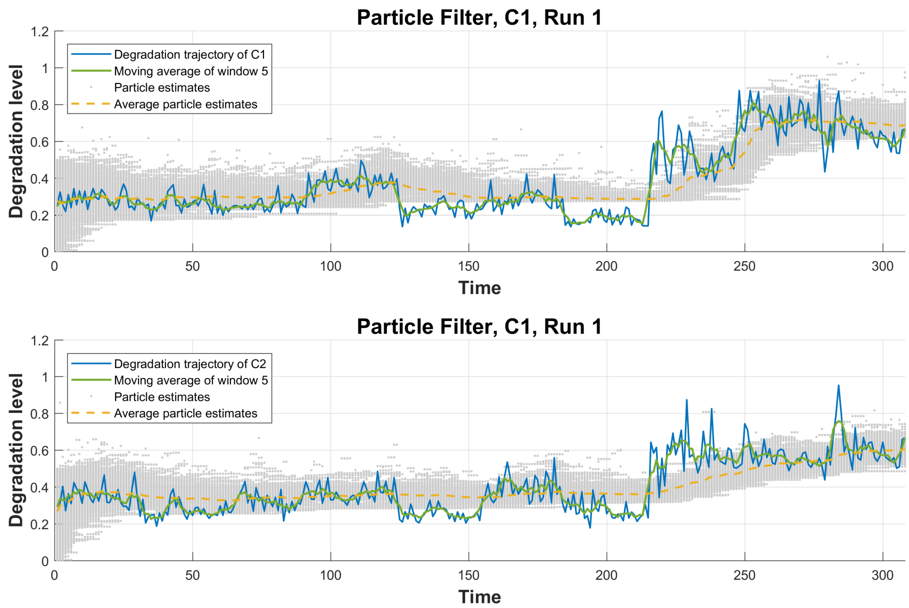

Figure 5.4: Fit of particle filter estimates to degradation data of components 1 and 2

tion trajectories of the components. In the following figures these are referred to as "With Interaction". These simulations are also performed using the reduced model, whereby no stochastic dependence is considered, these are referred to as "No Interaction". This is done so that we can compare the prognostic performance difference between the case where we consider degradation dependence in degradation modelling, and in the case where we do not.

The simulations are performed for C1 in run 1 Figure 5.5, and for C2 in run 1 Figure 5.6. Then, since C2 remains unchanged for run 2, we only simulate the degradation trajectory for C1 in runs 2 and 3 as shown in Figures 5.7 and 5.8, all while considering the state of C2 in those runs.

From the figures it is clear that considering degradation dependencies provides an advantage when attempting to predict the real degradation trajectories of the components. This is clearly seen when considering the time instance where the degradation of a component is supposed to reach the failure threshold t_{eol}. Table 5.4 summarises the different t_{eol} estimates.

From Table 5.4 we see that the difference between the actual observed t_{eol} and the predicted t_{eol} for C1 when not considering stochastic dependence shows a strict growth trend. It starts from 53 in run 1, to 168 in run 2, and then 190 in run 3. This is because the parameters of the models are estimated in run 1 using

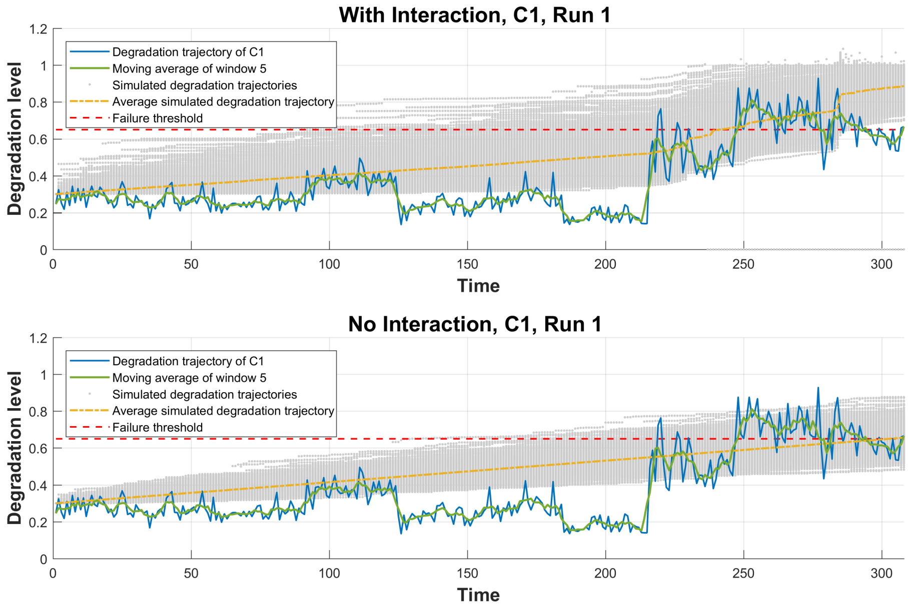

Figure 5.5: Simulated degradation trajectories for component 1 in run 1

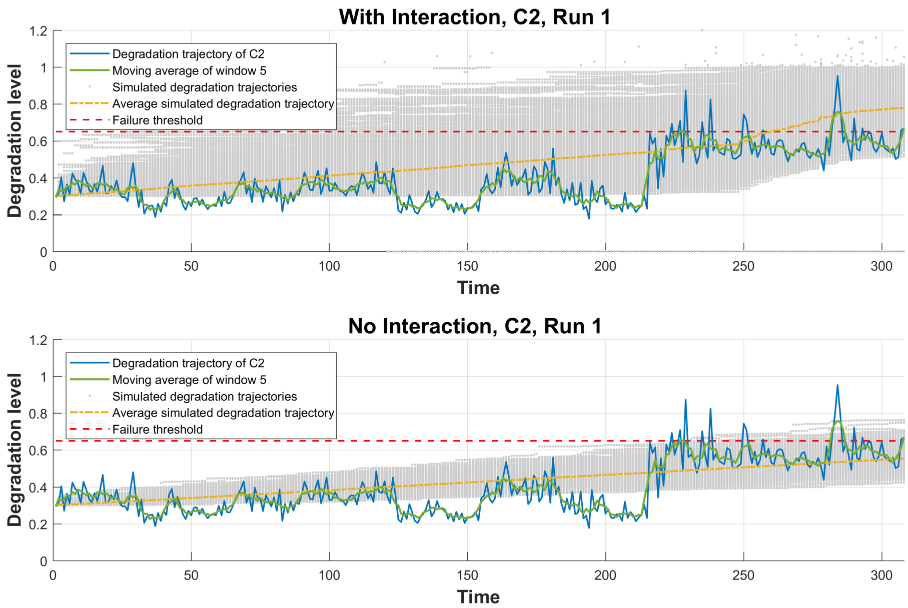

Figure 5.6: Simulated degradation trajectories for component 2 in run 1

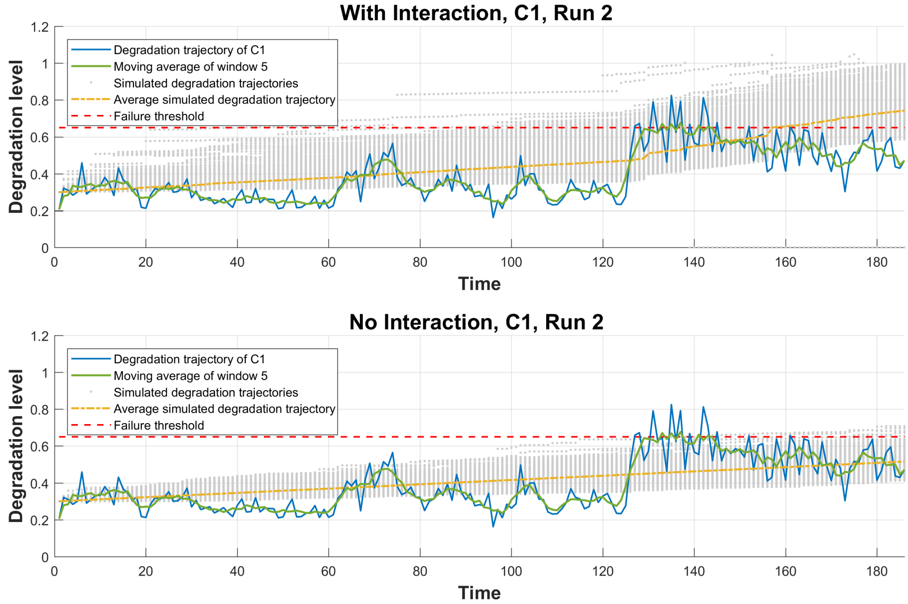

Figure 5.7: Simulated degradation trajectories for component 1 in run 2

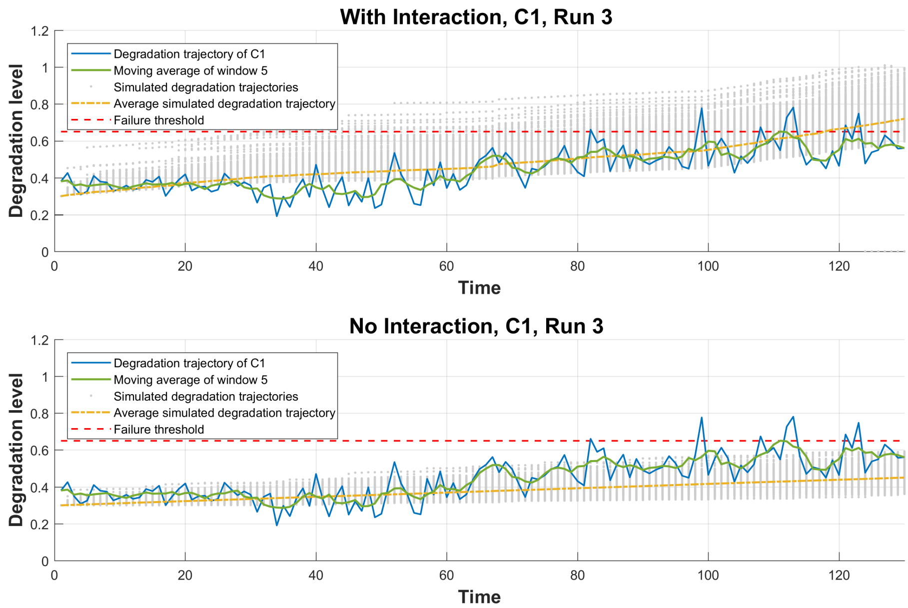

Figure 5.8: Simulated degradation trajectories for component 1 in run 3

	Actual t_{eol}		t_{eol} with interaction		t_{eol} no interaction	
	C1	C2	C1	C2	C1	C2
Run 1	248	227	239	259	301	429
Run 2	133		157		301	
Run 3	111		118		301	

Table 5.4: Actual time of end of life, and average predicted time of end of life for components 1 and 2

PF. Therefore the reduced model cannot account for the accelerated degradation that is due to a new C1 being coupled with a worn out C2. On the other hand this difference does not show this trend when we consider the stochastic dependence. The difference is 9 in run 1, then 24 in run 2, and then just 7 in run 3. This clearly indicates the criticality of modelling stochastic dependencies between components when attempting to do prognostics.

Furthermore, Figures 5.9 and 5.10 represent probability histograms of the estimated degradation distributions at the actual time of failure of the components. Once again it is shown that modelling stochastic dependence has a great impact on predicting the actual degradation trajectory of the components. This is because the average of the estimated degradation level is always closer to the failure threshold value than when not considered.

Finally, a note regarding the prognostics using the generic model provided in this chapter. These prediction of t_{eol} are simulated at $t = 0$ in runs 2 and 3. Therefore, if PF is used for an online update of the parameters after receiving new observations of the component health, we assume that the predictions would then be even more accurate. This would also allow for considering break points in the component's health state, in the likes of shocks that might occur due to environmental effects or sudden excess loading.

5.6 Discussion

This chapter motivated the importance of modelling stochastic dependencies in systems with multiple components. For enabling such a task, we presented a generic degradation model in which the degradation process of a component may be dependent on the operating conditions, the component's own state, and the

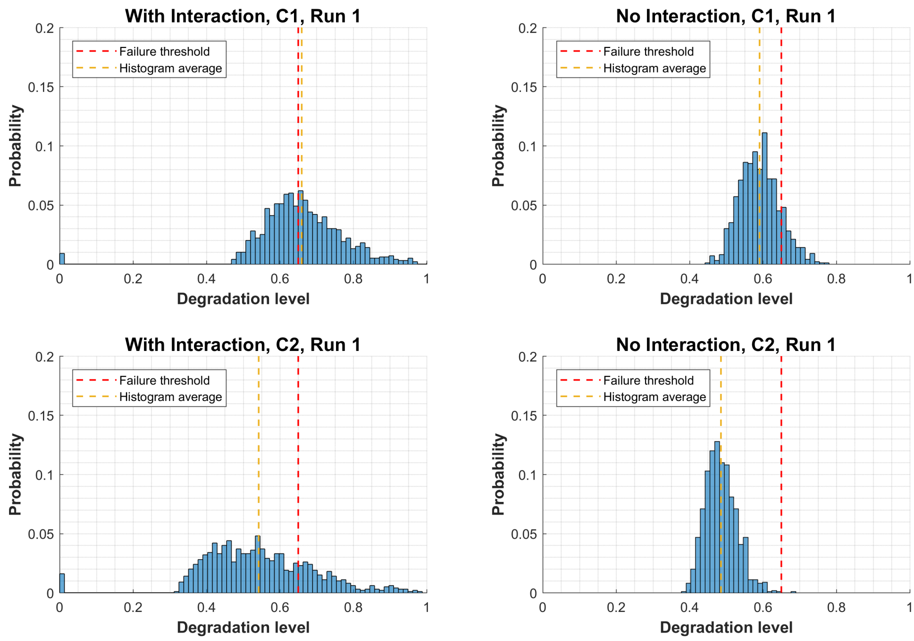

Figure 5.9: Probability histograms of the estimated degradation distribution at the actual failure time of components 1 and 2 in run 1

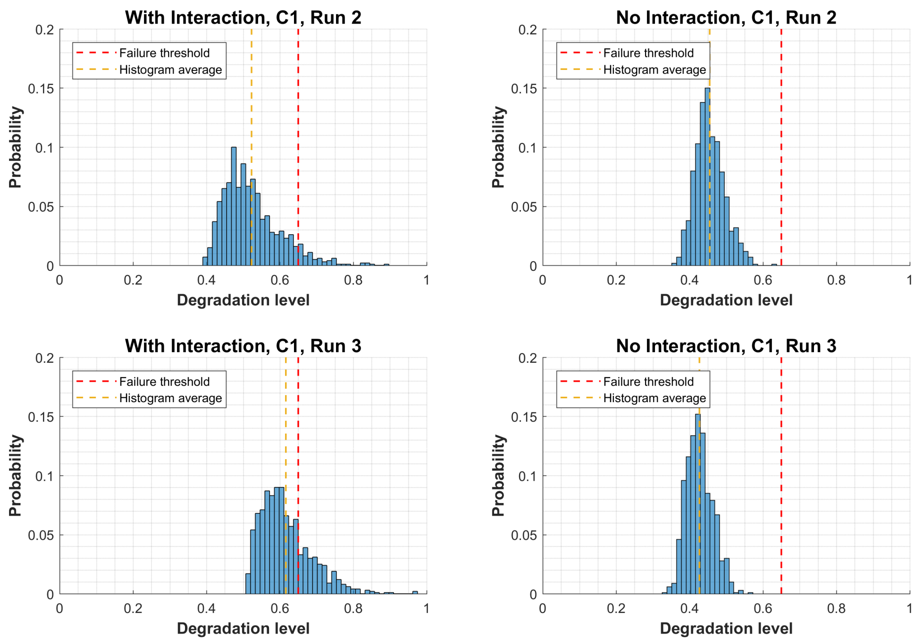

Figure 5.10: Probability histograms of the estimated degradation distribution at the actual failure time of component 1 in runs 2 and 3

state of the other components. We then showed how to fit the models to data using particle filter. This method is then used for the data generated by the gearbox. The results showed a comparative study between the model where stochastic dependence is considered, and a reduced model where this dependence is not considered. From this we conclude that considering stochastic dependence allows for more accurate predictions of t_{eol}. This is essential since this is directly linked to RUL and therefore affects maintenance decision making.

Chapter 6

Unsupervised Learning for Diagnostics of Multi-Component Systems

Contents

6.1 Chapter Summary

Performing diagnostics of a system is not straight forward, the principle is to be able to discriminate novelties that represent a change in system state. In this chapter we start by introducing automatic pattern recognition through unsupervised machine learning. We then review the literature on this topic in the context of PHM. Afterwards, we present our methodology for performing diagnostics of multi-component systems. This is done using Gaussian mixture models for partitioning degradation data, and aims to uncover different degradation phases in such systems. Moreover, this allows discriminating nominal degradation behaviour from accelerated ones.

6.2 Introduction

In light of what is presented in the previous chapters, we notice that stochastic dependence can accelerate the degradation rate of components leading to unexpected faults and failures that jeopardise system reliability. It is therefore important to identify this accelerated degradation behaviour. In this chapter we show our approach to partitioning data collected from the gearbox system with the aim of identifying different degradation rates. Such an approach is further justified when considering the prognostics and health management (PHM) context, since it is not guaranteed that the condition monitoring data are always labelled, whereby monitoring data are associated with the health state of the system being monitored. Furthermore the scarcity or unavailability of such labels make it difficult to infer useful information and thus make meaningful decisions.

Diagnostics is the process of detecting faulty attributes within a condition monitoring signal. See Section 2.3.3 for more detail. However, although humans are capable of recognising and distinguishing a multitude of patterns that occur naturally, we still face many difficulties when trying to understand and discriminate patterns that occur artificially such the ones present in signals and data generated by sensors. This becomes more of an overwhelming challenge when we deal with large amounts of data in the likes of condition monitoring data. Recently however, there has been a surge in the number of applications which employ machine learning techniques for finding patterns in signals and data (51, 114).

Unsupervised learning is the branch of machine learning that can be used to automatically recognise patterns inside the data. It does so by partitioning the unlabelled data into different clusters, this is commonly referred to as clustering. The available algorithms for clustering differ between each other in the way they cluster the data.

The notion of proximity, similarity and dissimilarity is integral to clustering algorithms, since the partitioning of the data is usually performed on data points that are near one another. For example, a measure of proximity like Euclidean distance can be used to show similarity between data points.

There exists two main classes for performing clustering, hierarchical clustering and non-hierarchical clustering (268).

Hierarchical clustering methods work in a way where initially each point is a cluster by itself, and then repeatedly combines the two nearest clusters into one. It is however highly demanding of computational resources with algorithms having a computational complexity of $O(N^3)$ or $O(N^2 \log N)$ at best. Therefore it is not suitable for real time applications and thus unattractive for the field of PHM.

On the other hand, non-hierarchical clustering methods start by assigning data points to clusters based on some randomly allocated cluster centres. They maintain this set of clusters while shifting them slowly over the state space, and then work by adjusting the placement of points into their nearest cluster. They are considerably more efficient than hierarchical clustering methods and therefore have been implemented extensively in the field of PHM. This will be reviewed in the following subsection.

6.2.1 Literature Review

The most commonly implemented techniques for clustering in the PHM context are: K-means, Fuzzy C-means clustering (FCM), self organising maps (SOM), and Gaussian mixture model (GMM).

K-Means K-means (161) is one of the simplest and easy to implement unsupervised clustering method. The method works by initially selecting a k number of means and randomly placing them over the data. The data points are then associated to the most relevant or nearest mean forming a cluster, this is the assignment step. After this the centroid of each of those clusters is computed and becomes the new mean, this is the update step. The assignment and update steps are repeated until convergence, that is, when the centroids no longer change, or until a specific termination criteria is achieved.

A major limitation of this algorithm is that it is very sensitive to the initial starting conditions. Because of this, many works have developed different variations and strategies for choosing the starting point for the cluster centres. An efficient approach would be to have multiple runs of this clustering algorithm and then choose the outcome that is best.

In (272), a vibration model which describes the physics of the dynamic behaviour of defective rolling element bearings is used to simulate data. Then, frequency-domain features and signal envelope are extracted from the simulated data. These features are used as a starting strategy to select the initial cluster centres, giving robustness to the K-means approach. Finally, the method is applied on real data to check whether a bearing fault exists or not and to identify the type of defect. Compared to the general k-means method, the authors reported a substantial improvement of the classification results.

Fuzzy C-means The Fuzzy C-means algorithm (26) is one of the most well-known and used algorithms for clustering, and seen as a an extended version of the K-means algorithm. The K-means algorithm in fact, assigns each data sample to one of the clusters, without any likelihood of that sample belonging to that cluster. FCM overcomes this issue as it provides a fuzzy strategy, by which it assigns the data samples to clusters with a certain grade of membership.

In (141) the authors propose an improved FCM algorithm for fault diagnosis. In fact, this work tries to take into account the different contributions that the features bring to clustering, contrarily to the general FCM approach in which the different importance degrees of the features are not taken into consideration. In this work, time-domain and frequency-domain features are computed from the data. Then, a compensation distance evaluation technique is used to compute the weights for each feature and reflect their different sensitivities to the clustering; then, the improved FCM clustering algorithm is applied to fault diagnosis of locomotive roller bearings. Other works on feature-weight assignments can be found in (262, 47).

In (274) the authors propose a method called Fuzzy positivistic C-means clustering for fault detection and identification in the context of vehicle suspension systems. In fact, it is very complicated to construct physics-based models of suspension systems and these approaches are non satisfactory. Instead data-driven models can be used to extract the meaningful information contained in process measurements. The approach proposed in this work is divided into three steps. 1) data pre-processing, the number of clusters c is roughly identified based on principal component analysis (PCA). 2) the data set is clustered into the c different clusters using fuzzy positivistic C-means clustering. In this step fault identification metrics are defined. Fault lines are structured in the data space, and faults are identified based on the distance from the centre of the cluster to the fault lines. This approach serves for identifying the occurrence and the type of fault. 3) Fisher discriminant analysis is used to isolate the major factors for the faults. Finally, the method's performance is demonstrated on benchmark accelerometer data.

Self Organising Maps Self organising maps is a type of artificial neural networks (ANN) that is trained in an unsupervised manner. A SOM is a collection of neurons represented in a one or two dimensional array. Each neuron has a weight vector that corresponds to a point in the data space or feature space. Each data point is then assigned to one of the neurons according to its proximity to the weight vectors. The idea here is to train the SOM by adjusting the weight vectors so that the vectors for neighbouring neurons are in proximity to each other in the feature space. Accordingly nearby data points are assigned to the same neuron or its neighbours. This brings about clusters by grouping nearby data points together.

for example in (110), vibration data from bearing is used to recover features in time and frequency-domain. Then the data is normalised and fed into a SOM. The SOM is used to cluster the unlabelled data. The output vector contains a minimum quantisation error index which is sent to a back propagation neural network. The MQE indicates how far the new data deviate from the normal operation data sets that were used to train SOM. Finally, in the last stage, the back propagation neural network trained with a degradation database, has its weights chosen according to the failure times method. It can therefore identify how the MQE index predicts the RUL of a bearing. The method is validated experimentally on a bearing vibration database resulting in accurate predictions of RUL.

In (266) A SOM-based radial-basis-function (RBF) neural-network is proposed for early fault detection and preventive maintenance in the context of induction machines, where faults are usually related to rotor faults. In this work, four features are computed from the power spectra of machine vibration data.

These features are used as inputs to a RBF neural network. In order to find the optimal network architecture, an extended SOM comprising of a cell-splitting grid algorithm is used to self-adjust the number of hidden neurons in the network. The advantage of this approach is that the appropriate network architecture modelling does not require many trial tests for the training phase. To verify the effectiveness of the method, the authors apply it on vibration signals from electrical and mechanical faults. Results show that the approach is able to classify different types of machine faults and indicate the extent of the faults.

Gaussian Mixture Model In this thesis we choose to use Gaussian mixture models for performing clustering. This is due to the many advantages it represents over other approaches in the context of PHM. This choice is detailed along with the GMM background in the next section.

In (280) the authors propose an unsupervised learning method for bearing performance degradation assessment. Locality preserving projections (LPP) is used for feature extraction from the original vibration data set which is represented in time-domain, frequency-domain, and time–frequency-domain signals. In fact, in this application, the original data set is generally high dimensional so it needs to be reduced. The authors propose LPP for dimentionality reduction, since compared to PCA, it preserves the local structure of the data set. Then, a GMM-base model is constructed using a healthy data set. This approach guarantees to provide a method for online machine monitoring, every time a new signal is used as input. GMM in fact, performs the log likelihood of each new input and indicates how it follows the probability distribution of the initially GMM trained healthy data set. This is used to evaluate whether a bearing is starting to degrade and shows the degradation propagation.

In (281) GMM is used as a tool to construct baseline clusters only out of healthy data of components. Then, an adaptive GMM with dynamic learning rate is used to update the parameters of the Gaussian components. In this way, the adaptive GMM will learn the changes in the state of the components online. A quantification index is then used to evaluate whether the component is in current degraded state. The quantification index is a Kullback–Leibler divergence (KL) between the adaptive GMM and baseline GMM, which measures the similarity between two probability density distributions of the Gaussian components. Time domain statistical features and wavelet energy are generated from multi-sensor signals, i.e. acoustic emission sensors signal and vibration sensor signals from a milling machine. Then, the final features are extracted by PCA and are used in the GMM and adaptive GMM. Results show that the adaptive GMM can learn to update parameters when the conditions of the components are changed from healthy to degraded. Also, the KL-divergence shows the degradation propagation as early as possible by increasing its value when the degradation occurs.

A highly cited work can be found in (241). A mixture of Gaussian hidden Markov models (MoG-HMM) is used for the estimation of RUL. A Wavelet Packet Decomposition (WPD) technique is used to extract features from raw data. These features are used in the unsupervised learning process of estimating the parameters of a MoG-HMM model. In this case, the MoG-HMM model allows to describe the different states and operating conditions of components. In this learning phase, each raw data history is used to compute the mean duration and standard deviation for which the component has been in each state of the corresponding MoG-HMM model. Also, the number of visits of in the component's history for each state are calculated. The second phase of the process uses the online data to check the component's state and computing its RUL. This is done in three steps. 1) After computing the sequence of the component's state in the MoG-HMM model, the most persistent state in the last observations is used to identify the current state of the component. 2) The probability of the component to find the critical path, from the current state to the failure state, is computed. The probability is computed in the case of short transition between the states, a fast degradation process, and slow transition. 3) The RUL is estimated by using the probabilities of transition. The method is then applied on data from bearings and shows correct estimation of RUL. However, these predictions can only occur once degradation starts.

6.3 Proposed Diagnostics Approach

In this section we first motivate our choice of using Gaussian mixture models for performing clustering on multi-component degradation data, and then present our approach for performing diagnostics for multi-component systems.

Most of the classical clustering algorithms such as hierarchical clustering, self-organising maps, k-means and Fuzzy C-means; are largely heuristically motivated and do not present a rigorous approach for determining an optimal number of clusters (271). In comparison, clustering methods that are based on probability models provide a great advantage since they allow the choice of an optimal number of clusters based on statistical criteria (56, 184), this is the case of the Gaussian mixture model. Successful applications of GMM-based clustering have been reported in the literature (17, 46, 269). Next, we further illustrate the advantages of using GMM for clustering by providing a comparison with the K-means clustering algorithm.

The K-means clustering algorithm is powerful and simple to implement, the clusters are defined by a centre which is a single point in feature space, and we then assign each data point to its nearest cluster. However, when we have groups that overlap in the feature space, then it is hard to know which assignment is right. K-means also best uses Euclidean distance to the centre. So if our clusters are defined in a non-circular shape, for example data points that could be clustered into 2 clusters where one has its spread over one dimension, and the other on the other dimension; and both clusters are centred around the same place; K-means will not perform optimally, and so not discover the right assignment of points to their corresponding groups.

GMM is a very successful clustering technique that is based on probability density estimation, it uses a mixture of Gaussian models and the expectation maximisation (EM) procedure in order to fit the model parameters. Gaussian mixture models can be seen as an extension of the K-means model, in this case clusters are formed using Gaussian distributions. Thus, we now have a mean and covariance which could describe ellipsoidal shapes. We could then fit the model by observing the likelihood of the observed data using the EM algorithm, which will assign the observations to each one of the clusters using a soft probability.

So since GMM assigns each observation in the data with a probability of membership to a specific cluster, this eliminates the need for hard boundaries. Such a thing is essential in the field of fault diagnosis because it would lead to a lowered false alarm rate (75), which is a major challenge for maintenance, and allows for smooth transition from one cluster to the other. In our case the different degradation behaviours which in real life do not usually change suddenly, further justifying the use of soft cluster boundaries versus hard ones.

Also, in GMM, after the clustering takes place, we obtain a generative model for the data, which means that we could use that to sample new examples which are similar to the real observations. This advantage can prove to be crucial in degradation modelling, and could then lead to accurate simulation of the degradation processes leading to more optimal maintenance strategies. Furthermore, this could be used for balancing data sets, since fault and failure data are usually scarce. It would therefore reduce the risk of high bias when performing prognostics.

Therefore we decide to use GMM as our pattern recognition technique due to the many advantages that it presents to the field of PHM.

We aim to cluster degradation behaviour of multiple components within a systems into different phases. GMM is highly appreciated in such a task since it provides all data points with probabilities of belonging to all the clusters, therefore even though a data point is clustered in a healthy degradation phase, it might be belonging to a more critical cluster but to a lesser degree. This can be monitored and allows for incipient fault detection.

When clustering different degradation behaviours we attempt to point out that the degradation behaviour of components can sometimes deviate from what is nominal. This is due to the stochastic dependence that is present between components, and which can present degradation state-rate interaction, whereby the degradation state of one component can affect the degradation rate of other components. If correctly done, this should uncover different phases of degradation rate, and especially point out to accelerated degradation behaviour in multi-component systems. Consequently, appropriate precaution measures can be taken.

6.3.1 Deciding on Cluster Count

A major issue that one might think of in the general case of unsupervised learning is the cluster count, so how can the number of clusters can be chosen beforehand. This is studied when considering cluster analysis, and many approaches have been conceived.

The silhouette method stands out as one of the most successful ways of determining the cluster count (174, 209). The silhouette method is used to determine the optimal number of clusters, which basically iterates between different cluster numbers, and computes what is called a silhouette coefficient for the data, which ranges between $+1$ and -1. Coefficients close to the value $+1$ indicate that the data sample is far from the neighbouring clusters, while coefficients close to -1 indicate that the data sample is in the wrong cluster, and so the approach is to aim for a higher average silhouette coefficient to obtain the optimal number of clusters in that way.

Therefore we use the silhouette method in our approach. Also, we would like to note that since we are aiming at clustering different degradation phases, it is important to cluster the data into more than two clusters. This is because it is highly undesirable to switch from one cluster where degradation behaviour is nominal, to a cluster where degradation behaviour is considered critical without passing through a medium cluster. The reason for this is that such a switch would not allow for the right measures to be taken so that a fault or failure is prevented.

6.3.2 Gaussian Mixture Model Background

For training GMM, we begin with several mixture components, indexed using c, each of which is described by a Gaussian distribution. So each has a mean μ_c, a variance or covariance σ_c, and a mixing coefficient π_c.
We can now see how the joint probability distribution of multiple components c is to be defined by the weighted average of the individual components, such as:

$$p(x) = \sum_{c=1}^{C} \pi_c \mathcal{N}(x; \mu_c, \sigma_c) \tag{6.1}$$

where C is the number of clusters. A way to interpret this joint probability distribution over x in a simple generative manner, is if we were to draw a sample from $p(x)$, we first select one of the components with discrete probability π, thus components with large π are selected more often. And so, select a mixture component with probability π as:

$$p(z = c) = \pi_c \tag{6.2}$$

Then, given the component assignment $z = c$, we could draw a value for x from the corresponding Gaussian distribution as in:

$$p(x|z = c) = \mathcal{N}(x; \mu_c, \sigma_c) \tag{6.3}$$

So together the above mentioned two distribution make a joint model over x and z. Discarding the value of z gives a sample from the marginal $p(x)$ defined in 6.1. Such models are called latent variable models (LVM) (170), the data x are modelled jointly with an additional variable z that we don't get to observe, and so is considered hidden. This presence of the unknown value of z helps explain patterns in the values of x, in this case the clusters.

Since we will be using multivariate features, we will use a multivariate Gaussian, which has same quadratic form, but uses vectors for μ of length n, or the total number of features in a data point x, and a $n \times n$ covariance matrix Σ as in:

$$\hat{\mu}(\underline{x}; \underline{\mu}, \Sigma) = \frac{1}{(2\pi)^{d/2}} |\Sigma|^{1/2} exp(-\frac{1}{2}(\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})) \tag{6.4}$$

where underbars represent vector form.
The EM algorithm basically proceeds iteratively in 2 steps, namely the expectation and maximisation steps. If we were given data from a multivariate Gaussian, the maximum likelihood estimates for the model parameters are as such:

$$\hat{\mu} = \frac{1}{n} \sum_i x^i \tag{6.5}$$

$$\hat{\Sigma} = \frac{1}{n} \sum_i (x^i - \hat{\mu})^T (x^i - \hat{\mu}) \tag{6.6}$$

Where $\hat{\mu}$ represents the first moment of the data, or the mean, and $\hat{\Sigma}$ represents the second moment of the data, or the covariance estimate, which is the mean of the $n \times n$ matrices formed by the outer product of $\underline{x} - \underline{\mu}$ with itself.
Concerning the first step, the expectation or E-Step of the EM algorithm, it considers the Gaussian parameters μ_c, Σ_c, and π_c as fixed. For each data point i and each cluster c, it computes the responsibility value r_{ic}, i.e. the relative probability that data point x_i belongs to a specific cluster c. It does that by computing the probability of x under the model component c, which is a weighted Gaussian, and then normalises by the total of all the values of c:

$$r_{ic} = \frac{\pi_c \mathcal{N}(x_c; \mu_c, \sigma_c)}{\sum_{j=1}^{C} \pi_j \mathcal{N}(x_j; \mu_j, \sigma_j)} \tag{6.7}$$

This yields a $n \times C$ sized matrix, which sums to 1 over the index c. Practically if x_i is very likely under the c^{th} Gaussian, it would get a high weight.
For the second step of EM, the maximisation or M-Step, we fix the assignment responsibilities r_{ic} and update the parameters of the clusters μ_c, Σ_c, and π_c. Then for each cluster c, we update its parameters using an estimate weighted by the probabilities r_{ic} as if it observed a fraction r_{ic} of the data point i. So cluster c sees a total number of data points n_c which is the sum of the soft memberships

or fractional weights assigned to cluster c as in:

$$n_c = \sum_i r_{ic} \tag{6.8}$$

then π_c equals n_c normalised by the total number of data n, so this is seen as the fraction of the data point probabilities that is assigned to cluster c:

$$\pi_c = \frac{n_c}{n} \tag{6.9}$$

The weighted mean μ_c is the weighted average of the data:

$$\mu_c = \frac{1}{n_c} \sum_i r_{ic} x^i \tag{6.10}$$

Similarly the covariance matrix is a weighted average of the $n \times n$ matrices formed by taking the outer product of x^i minus its cluster c's mean:

$$\Sigma_c = \frac{1}{n_c} \sum_i r_{ic} (x^i - \mu_c)^T (x^i - \mu_c) \tag{6.11}$$

These iterations of the EM algorithm increase the log likelihood of the model, and so increase its fit to the data. The log likelihood is then the log probability of the data points under the mixture model. Therefore it is the sum of the data points of log of the probability which is a mixture of Gaussians.

The EM algorithm is considered as a form of coordinate descent, so it is guaranteed to converge. In practice the algorithm is stopped when the parameters or the log likelihood objective start changing slowly. However, convergence is not guaranteed to a global optimum, and so we can start from several different initialisation and then use the log likelihood to select the best. Then for new unseen data, we perform the E–step of the algorithm to assign it to a specific cluster.

6.4 Case Study

We consider the case of the gearbox platform presented in Chapter 4 for applying the diagnostics approach. In the following sections, we demonstrate our use of GMM for clustering the degradation data of the two gear components. This is done in 2D, whereby the degradation state of gear 1 is considered against the degradation state of gear 2.

6.4.1 Clustering of Different Degradation Phases

We use the RMS gear mesh frequency magnitude to represent the degradation level of the components. These degradation data are provided from run 1 to the GMM. They are then clustered after running 10 repetition of the clustering algorithm. This is done since the final outcome might be sensitive to initial starting conditions. Each repetition consists of 200 EM steps. The covariance matrix of the GMM are chosen to be diagonal since this is more computationally efficient than non-diagonal matrices. This then allows this method to be implemented in less time and therefore could more frequently update its clusters in the presence of a new data set. Moreover, the covariance matrices are not shared, this means that they differ from one cluster to another. This adds flexibility to the clusters, since some may span over a different range than others.

After using a silhouette analysis, the optimal cluster count is shown to be 3 clusters. The silhouette plot result is shown in Figure 6.1 and has presented the highest average between different cluster counts. Specifically a cluster count of 3 to 6 were considered, and their silhouette averages measured.

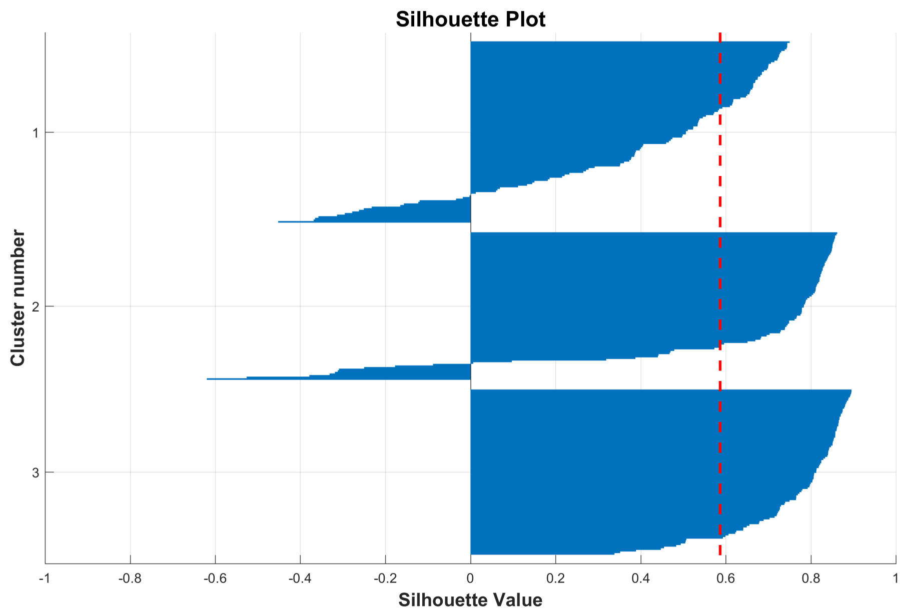

Figure 6.1: Silhouette plot of three clusters. The dashed red line represents the average value

The clustering result of the degradation data from run 1 can be seen in Figure 6.2. We can see three different clusters, each represented by a multivariate Gaussian with distinct mean values and covariance matrices.

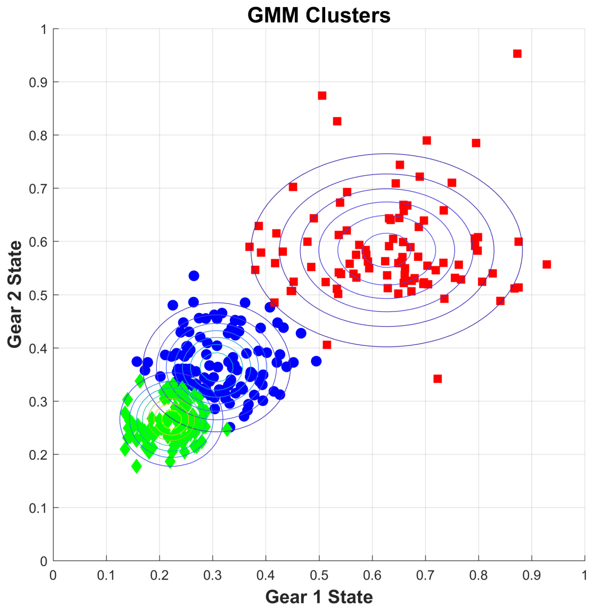

Figure 6.2: Degradation data from run 1 partitioned into three different clusters

From Figure 6.2 we can clearly see the three different clusters. Interpreting these clusters and according to the degradation levels that they span over. We are initially lead to the realisation that the clusters represent different stages of wear of the components. The green cluster represents a healthy state for components; the blue cluster represents a state where components start to wear out; and the red cluster represents a state where components can be considered as severely worn out.

6.4.2 Results

After partitioning the degradation data from run 1 into three different clusters, we can use the clustering model to cluster new incoming data. This can be used as basis for diagnosis of the system.

Next, we show how this can be done by clustering the data from runs 2 and 3. An overlay plot showing the time series degradation trajectories where each time step belongs to a different cluster, can be seen in Figure 6.3.

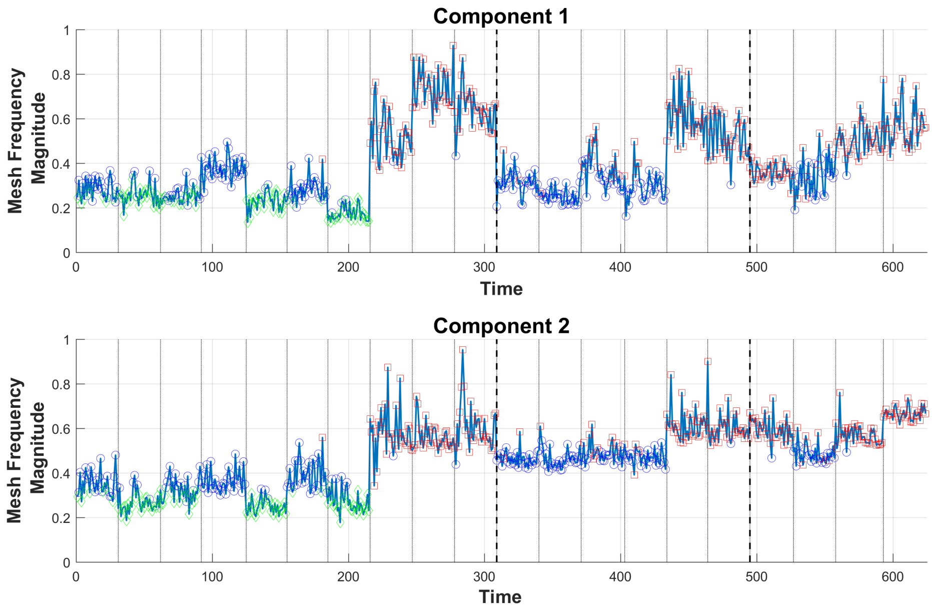

Figure 6.3: Degradation phases overlaid on the time series degradation data of the two gears

From Figure 6.3 we see that straight from the start of run 2 the degradation data points are clustered in the blue cluster. If we look at the start of run 3, we see that the data points are clustered to the red cluster. This was expected since we know that gear 2 was not replaced in runs 2 and 3, but that only gear 1 was replaced. Therefore we can conclude that these three clusters do not only indicate the degradation state, but also the degradation rate of the components. Since we know that that gear 1 reaches failure earlier in run 2, we can thus link that fact to being clustered in the blue cluster. Furthermore, gear 1 is clustered into the red cluster at the start of run 3, and we also know that it reaches failure in that run at a quicker rate than in runs 1 and 2. Thus, we can conclude that these clusters do not only represent the degradation state of the system, but also degradation rate. Accordingly three different degradation phases can be identified in the case of this multi-component system.

Inspired by what we have seen in Section 4.3.3, we can denote the green cluster to represent normal degradation, the blue cluster to represent accelerated degradation and the red cluster represent excessive degradation. Therefore based on the clustering of the incoming degradation data, this can be used as an indication of incipient faults in a multi-component system.

Incipient faults are more clearly shown when we consider the probabilities of belonging to different clusters. This is something that can be achieved when using GMM. Since a probability is associated to each degradation data point linking it to all clusters at variant degrees. This is illustrated in Figures 6.4, 6.5 and 6.6. In these figures, a moving average of window 30 is applied. This corresponds to the loading cycle length in the experimental scenarios discussed in Section 4.3.2.1. This is done in order to smooth out the transitions between the different clusters, and thus grant a better indication of incipient faults.

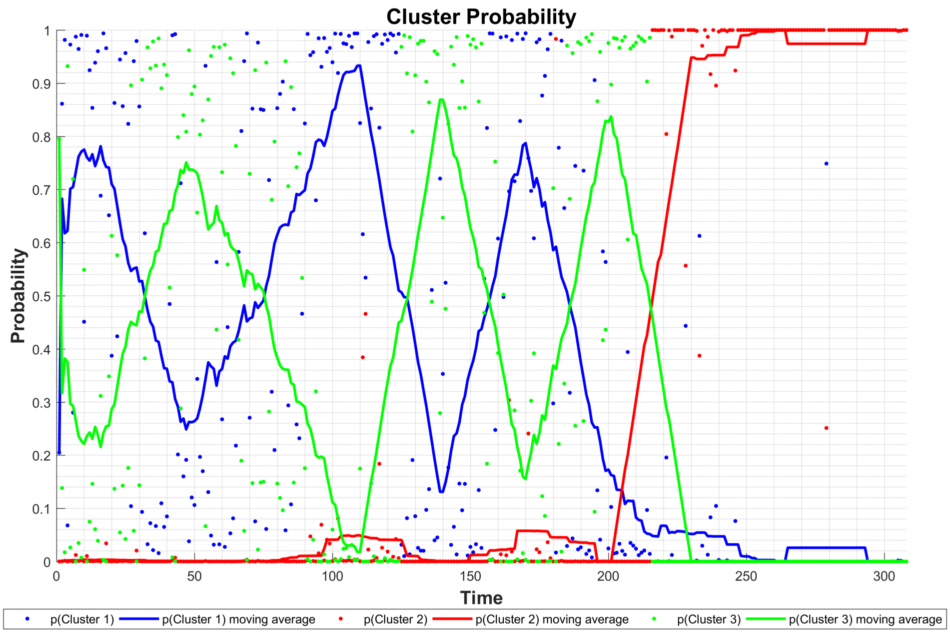

Figure 6.4: Degradation data probability of belonging to the three different clusters in run 1

In Figure 6.4, which corresponds to run 1, we see that the probability of assigning data points to the clusters is oscillating between the clusters green and blue at the start of the run. We can assume this is due to the different loading cycles of the experimental platform. However, although very small, we can see that the degradation data points start acquiring non-zero probabilities of belonging to cluster red between time step 100 and 125, and then again between 150 and 200. At time step 205 the degradation data points are then assigned to the red cluster.

In Figure 6.4, which corresponds to run 2, we see that from the start the dominant cluster is the blue cluster. Also the cluster probabilities show a positive trend of belonging to cluster red. This is until the probabilities are assigned to the red cluster at around time step 130.

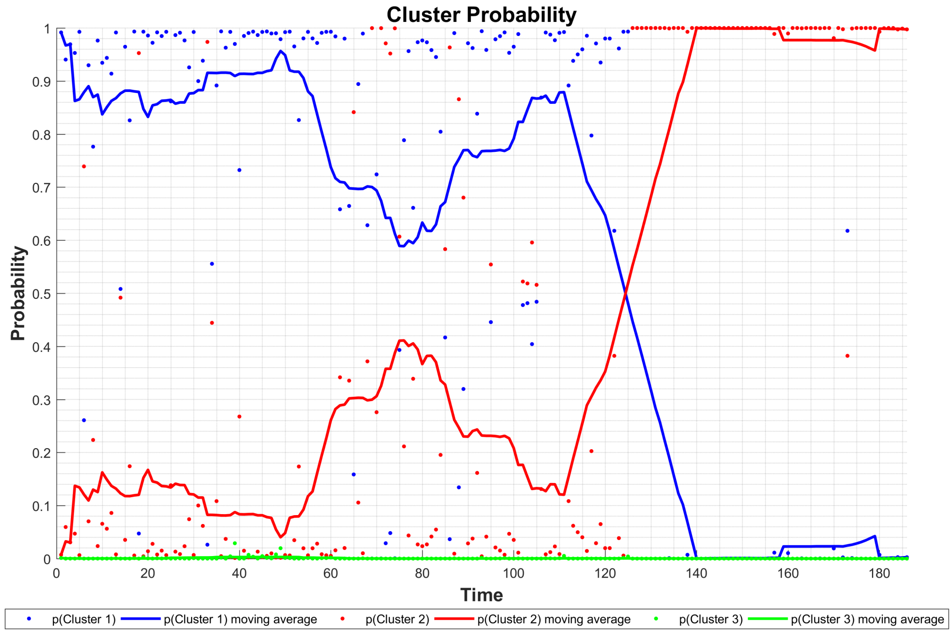

Figure 6.5: Degradation data probability of belonging to the three different clusters in run 2

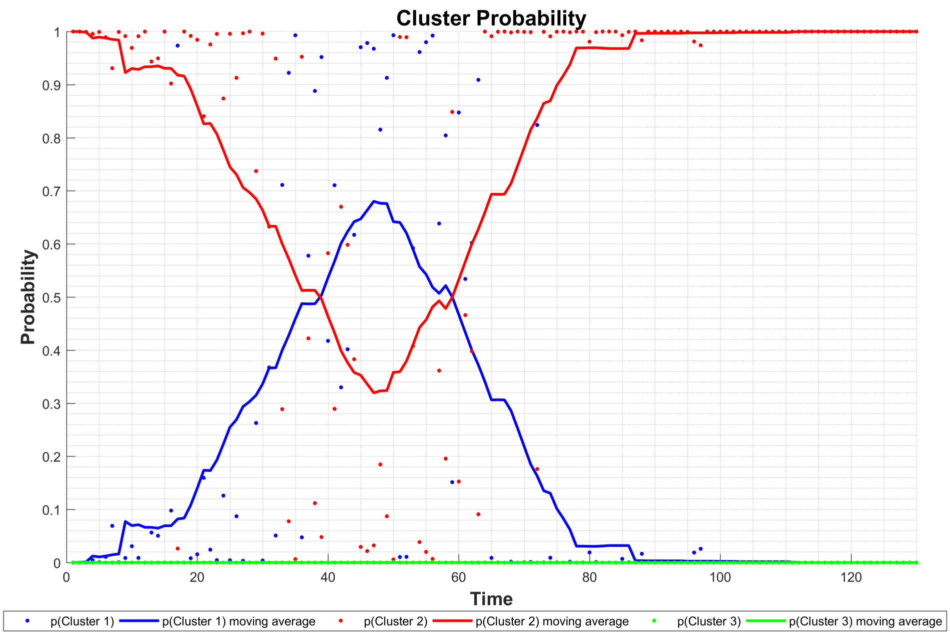

Figure 6.6: Degradation data probability of belonging to the three different clusters in run 3

In Figure 6.6, the dominant cluster is the red cluster, this is exchanged with the blue cluster for a small interval between time steps 40 and 60. In this run gear 1 reaches failure quicker than in runs 2 and 3.

What is shown in Figures 6.4, 6.5 and 6.6, demonstrates one of the strong points of using GMM for clustering degradation data. In some sense these moving averages can be used not only for diagnosing the system degradation phase, but also this shows promise for performing prognostics. However, this aspect still requires further study and validation.

Further to what was presented in this section, we would like to note that the cluster model can be updated once a set of new data is available. This is very useful for real world application, and should be done so that the clusters converge to a more representative position over the degradation data. We do this for the degradation data of runs 1 and 2, and then for runs 1, 2 and 3, these are visualised in Figures 6.7 and 6.8 respectively.

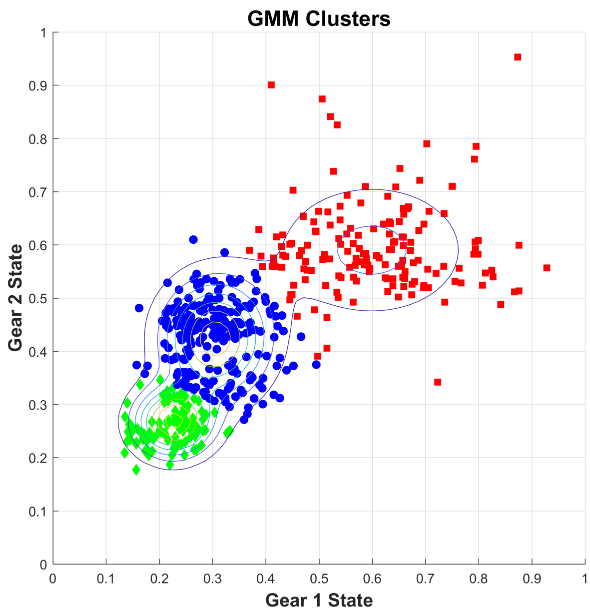

Figure 6.7: Degradation data from runs 1 and 2 partitioned into three different clusters

6.5 Discussion

In this chapter we presented our methodology for performing diagnostics of a multi-component system. This was done through implementing unsupervised

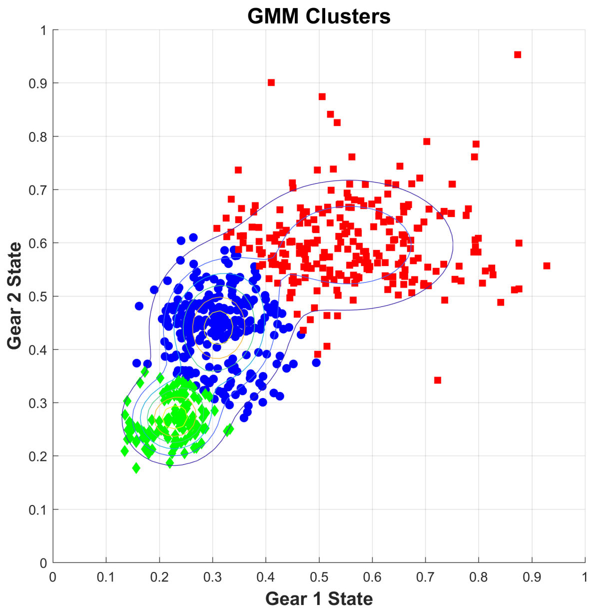

Figure 6.8: Degradation data from all three runs partitioned into three different clusters

learning with Gaussian mixture models. The principle was to cluster the degradation trajectory data of different components into different degradation phases. This is done in order to discriminate accelerated degradation behaviour of a system which arises out of stochastic dependency between the components.

This methodology was applied to the gearbox degradation data. The results showed a successful implementation and therefore partitioning of the data into three different clusters. These clusters indicated three types of degradation: normal degradation; accelerated degradation; and excessive degradation. Whereby the last two types indicate that components tend to wear out faster than usual. Therefore, this methodology allows the detection of incipient faults, even just after performing a maintenance intervention, by which a component can be replaced but starts degrading in an abnormal manner due to stochastic dependence.

Chapter 7

Maintenance optimisation

Contents

7.1 Chapter Summary

After discussing and showing our work on multi-component system health indicator extraction; and prognostics and diagnostics; we now move on to the final aspect of prognostics and health management, health management, and introduce a maintenance policy for such systems.

This chapter is concerned with economic dependence of multi-component systems. We begin by discussing multi-component dependencies. Nonetheless, in contrast to what was presented in Chapter 2, here we put more emphasis on the economic dependency, and present our modelling. We then describe the proposed maintenance policy and the optimisation process. To demonstrate the utility of the proposed maintenance policy we consider the case of the gearbox system and the resulting data from Chapter 4. We then show our results and include sensitivity analyses. We finally present our conclusions and discuss the managerial implications of the work.

7.2 Introduction

Maintenance involves preventive and corrective actions that are carried out to retain a technical system in, or restore it to an operating condition. Maintenance optimisation aims to determine effective and efficient maintenance plans for each component of a system in order to meet operator requirements for safety, reliability and value.

In the literature, many policies have been developed for the maintenance of single-component systems (63, 256). Such maintenance policies may be applied to multi-component systems when the dependencies between components in these systems are neglected. However, for many technical systems it is not reasonable to assume components are independent, and it is necessary to model component dependencies.

As seen in Chapter 2, dependencies can be classified into three types (179): (i) economic dependence, whereby the cost of joint maintenance of a group of components does not equal the sum of individual maintenance costs for these components; (ii) stochastic dependence, whereby the state evolution of a component influences the state evolution of other components; (iii) structural dependence, whereby components structurally form a part, so that maintenance of a failed component implies maintenance or at least the dismantling of other components that have still not failed. A fourth type of dependence is logistical dependency, that exists for example if a single repairman is responsible for the maintenance activities of various units or systems, or if a single stock of spare parts is used for the replacement of multiple units.

Taking into consideration dependencies between components when modelling maintenance of multi-component systems has recently shown an increase in popularity among researchers (28, 67, 89, 112, 179, 218). An overview about recent advances on condition-based maintenance for systems with multiple dependent components is given in (125). In fact, economic dependence has been investigated and integrated in a number of multi-component maintenance models (67, 156, 179, 246). However, in these works, stochastic and structural dependence are not considered. Failure dependence between components, whereby the failure of a component can induce the failure of others has been studied in the context of inspection by (89); and maintenance and warranty optimisation by (218, 286) for two-component systems. In the latter, several block replacement models considering both economic and failure interaction are proposed. Condition-based maintenance (CBM), in which the preventive maintenance decision is based on the observed system condition, has been introduced and has become an important model in maintenance optimisation frameworks. Condition-based maintenance has also been developed for two-component systems, see for example (20, 43, 156). However, in such maintenance models, again only economic dependence is considered. Recently, degradation interaction or state dependence, which implies that the degradation evolution of a component depends on both its degradation level and that of other components, has been introduced in (27, 28) for prognostics of system lifetime, and in (203) for maintenance optimisation. However, this latter work considers neither economic dependence nor intrinsic state dependence whereby state evolution of a component depends on its own state. Thus, there is a need to consider multiple dependencies in CBM.

With this in mind we propose a CBM model for a two-component system with rate-state interaction, whereby the degradation rate of each component depends not only on its own state but also on the state or degradation level of the other component. This dependence phenomenon can be found in a number of industrial systems, e.g., the state or quality of oil may directly impact the degradation process of the crank and vice versa; wear on a pulley may impact the rate of wear of a belt and vice versa; and likewise for chains and gears.

In our model, we suppose that inspections occur at regular time intervals and identify the state of each component. Maintenance actions are then optimally planned based on the current, inspected state of the components, and broadly corresponds to a choice of: do nothing; replace component 1 and not component 2; vice versa; replace both. An interesting consequence of the rate-state interaction that we study is that when one component is replaced but not the other, obviously the system is not perfectly maintained, i.e. it is not renewed, but more interestingly the new component will degrade at a different rate to that when the system was new, because the degradation rate of the new component depends on the state of the old component, for more detail see Section 4.3.2.1.

This partial replacement, or imperfect maintenance, of the system is then an imperfect "repair" that considers imperfect repair in a different way to the existing approaches in the literature, in which age/hazard reduction models predominate (63, 265, 264). It is important to note that when considering state dependence between components, existing CBM models may lead to sub-optimal policies, this can be concluded from the results of Chapter 5. This is because degradation modelling has a significant impact on finding optimal maintenance policy in CBM and, in these existing CBM models, state dependence is not yet considered. Therefore, an important contribution of this work is to propose and develop a CBM policy in which adaptive preventive maintenance and opportunistic maintenance rules to select a component or group of components to be maintained, and in so doing to open a new strand of thinking in the modelling of imperfect maintenance.

This chapter develops a model for a condition-based maintenance policy which is applied on the case study and degradation modelling results that are detailed in Chapters 4 and 5. A cost model is developed to find the optimal maintenance policy. We argue that ignoring stochastic dependence will lead to a maintenance policy that is cost-inefficient. Thus, in our view, our model makes a contribution to the literature that will not only lead to further developments in maintenance optimisation for systems with stochastic dependence but also be useful for practical application. Furthermore a study of the impact of economic dependence is considered via a sensitivity analysis.

7.3 System Description and Dependency Modelling

We refer to our generic multi-component degradation model which is described fully in 5. However, we consider a series system with only two dependent components.

When one or both components fail the system fails. Each component i is subject to a continuous accumulation of degradation in time that is assumed to be described by a scalar random variable X_t^i. Component i is considered as failed if its degradation level reaches the failure threshold L^i, $i = 1, 2$. When a component is not operating for whatever reason, its degradation level remains unchanged during the stoppage period if no maintenance is carried out. We assume that on replacement of a component, the degradation level of the component is reset to zero. Thus, when the two components are replaced together, the system is returned to the "as new" state, renewal.

In our model, we will use the term replacement of a component to denote the maintenance action whereby the degradation level of the replaced component is reset to zero. In reality, such an action may not in fact be a replacement but instead a "repair". Nonetheless, the model will assume a repair and a replacement are synonymous.

7.3.1 Economic Dependence Modelling

All necessary maintenance resources such as spare parts, maintenance tools, repairmen, etc. that are required to execute maintenance actions are assumed always available at a planned inspection time. It is also assumed that maintenance actions, such as replacements and inspections are carried out at discrete times. Replacements may be corrective, that is on failure of the system; or preventive, prior to system failure; and that in the standard manner the costs differ in the two cases.

7.3.1.1 Individual Maintenance Costs

If a preventive replacement is individually carried out, a preventive cost is then incurred. In a general way, the preventive cost of component i, denoted C_p^i, can be divided into two parts: $C_p^i = c_p^i + c_d.d_i$ where $c_d.d_i$ is the downtime cost due to production loss during replacement that takes d_i time units, and c_p^i includes all other costs such as spares, labour, set-up.

In the same manner, the cost of corrective replacement of component i is $C_c^i = c_c^i + c_d.d_i$, $(c_c^i \geq c_p^i)$.

Note, by preventive replacement of a component, we mean the replacement of a component when it has not yet failed, and by corrective replacement of a component, we mean the replacement of a component when it is failed. Full details of the maintenance policy follow in Section 7.4.

7.3.1.2 Economic Dependence and Cost Saving

When two components are simultaneously replaced, total maintenance cost can be reduced (179, 67, 263). In our model, this cost saving arises from the sharing of the replacement set-up cost, and the reduction of replacement duration. In this way, we define the cost-saving of joint replacement as

$$CS_{-,-} = a.(c_-^1 + c_-^2) + b.(d_1 + d_2).c_d, \tag{7.1}$$

where:

- c_-^i $(i = 1, 2)$ could be either c_p^i or c_c^i, i.e. preventive or corrective;
- a $(0 \leq a < min(c_-^1, c_-^2)/(c_-^1 + c_-^2))$ is the cost-saving factor for joint replacement of two components. It is shown in (263), that the cost saving is typically equal to 5% of the total replacement cost of the components $(a = 0.05)$;
- b $(0 \leq b \leq min(d_1, d_2)/(d_1 + d_2))$ is the duration-saving factor for joint replacement.

In this way, a and b express the economic dependence degree between the two components. When $a = 0$ and $b = 0$, the two components are economically independent. The larger the a and b values are, the stronger the economic dependence between the two components. Note, the effect of economic dependence on the availability of a system is studied in (64).

It is important to note that, in this work, the economic dependence is positive $(CS_{-,-} \geq 0)$. However, in parallel or complex structure systems where a failure of a component or a group of group of components may not lead to a failure of the system, the economic dependence may be positive or negative, see (175, 252).

In this work, the elements of the economic dependence are integrated into an opportunistic maintenance model that is described in the next section.

7.4 Maintenance policy

We assume that the degradation level of each component is measured at an inspection that is instantaneous, perfect, and non-destructive. An inspection incurs a cost c_I. A failure of a component is assumed to be instantaneously revealed by a self-announcing mechanism, but that replacement can commence only at the next inspection. In this way, the usual practical requirement to prepare for a replacement is modelled while the system downtime due to failure is known.

7.4.1 Description of the Proposed Maintenance Policy

We assume that the two components of the system are inspected at regular time intervals with inter-inspection interval ΔT. ΔT is a decision variable to be optimised. More precisely, for each component i ($i = 1, 2$), the degradation level at inspection times $T_k = k\,\Delta T$ ($k = 1, 2, ...$) is $X^i_{T_k} = x^i_{T_k}$. The maintenance policy is as follows. For $i = 1, 2$:

- if component i fails between (T_{k-1}, T_k) (when its degradation level reaches the failure threshold L^i), then it is replaced at time T_k;
- if at time T_k, component i is still functioning, it is inspected. Based on the inspection results and the preventive maintenance rules, a decision about whether or not component i should be replaced at time T_k will be taken. We specify rules for individual preventive replacement and for opportunistic preventive replacement.

Individual Preventive Replacement If the degradation level of component i ($i = 1, 2$) at time T_k is greater or equal to a fixed threshold m^i_p ($x^i_{T_k} \geq m^i_p$), component i is immediately replaced. m^i_p, called the preventive threshold of component i, and is a decision variable to be optimised.

Opportunistic Replacement The main idea of the proposed opportunistic replacement model is to capitalise on both the economic dependence and the stochastic dependence between the two components. The economic dependence is manifested in the shared set-up and the cost-saving therein; The stochastic dependence, through the term ΔX^{ji} in Eq 5.2 may also incentivise, depending on the strength of the dependence, joint replacement. To this end, for each component i, an opportunistic threshold, denoted m_o^i $(0 < m_o^i \leq m_p^i)$, is introduced. The opportunistic maintenance decision rule is the following. If component j $(j = 1, 2$ and $j \neq i)$ is correctively replaced or selected to be preventively replaced at time T_k, component i is preventively replaced together with component j if the degradation level of component i is such that $x^i_{T_k} \geq m^i_0$. The latter implies that the system is renewed at time T_k. m_o^i $(i = 1, 2)$ is also a decision variable that must be optimised.

An illustration of the proposed opportunistic maintenance policy is shown in Figure 7.1.

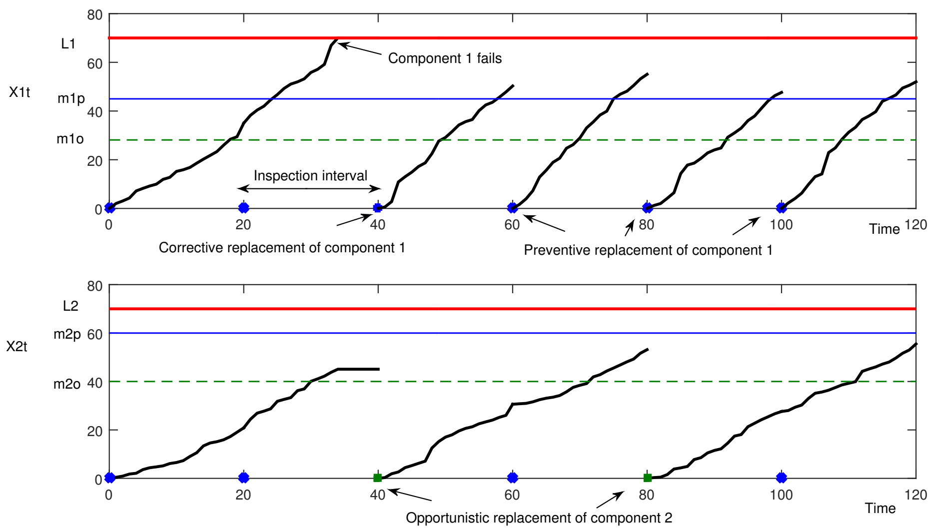

Figure 7.1: Illustration of components' degradation evolution and the proposed maintenance policy

We label this general policy as policy V. To study the impacts of opportunistic replacement, two special cases of this policy are herein considered as follows

- When $m_p^1 = m_o^1$ and $m_p^2 = m_o^2$, there is no opportunistic replacement, the policy becomes a classical condition-based maintenance policy (175) with discrete inspections, which we call policy V1;
- When $m_o^1 = m_o^2 = 0$, two components are jointly replaced together, the proposed policy becomes a joint replacement policy, which we call policy V2.

To investigate the effects of economic and stochastic dependence, we compare the cost-rates of these three policies V, V1 and V2 in Section 7.5.2.

7.4.2 Optimisation of the Proposed Maintenance Policy

As described, ($\Delta T, m_p^1, m_o^1, m_p^2, m_o^2$) are the decision variables of the general opportunistic replacement policy that we study. Their optimal values must be determined, given some suitable criterion. For this purpose, a cost model is developed in this section. In particular, we use the long-run expected cost per unit of time (or cost-rate) including replacement and inspection costs.

The cost-rate is defined generally as:

$$C^{\infty}(\Delta T, m_p^1, m_o^1, m_p^2, m_o^2) = \lim_{t \to \infty} \frac{C^t(\Delta T, m_p^1, m_o^2, m_p^2, m_o^2)}{t}, \tag{7.2}$$

where $C^t(\Delta T, m_p^1, m_o^2, m_p^2, m_o^2)$ is the cumulative total maintenance (replacement and inspection) cost in period $(0\ t]$. According to the renewal theory (208), Eq. (7.2) can be rewritten as follows:

$$C^{\infty}(\Delta T, m_p^1, m_o^2, m_p^2, m_o^2) = \frac{\mathbb{E}[C^{T_{re}}(\Delta T, m_p^1, m_o^2, m_p^2, m_o^2)]}{\mathbb{E}[T_{re}]}, \tag{7.3}$$

where $\mathbb{E}[.]$ is mathematical expectation and T_{re} is the length of the first renewal cycle of the system, i.e., all components of the system are replaced at time T_{re}. Without losses of generality, we assume that that $T_{re} = \Delta T.m$ (m is a positive integer), and so we get:

$$C^{T_{re}}(\Delta T, m_p^1, m_o^2, m_p^2, m_o^2) = \frac{\sum_{k=1}^{m}(C_{ins}^k + C_{main}^k) + T_{down}.c_d}{m \cdot \Delta T},$$

with:

- $C_{ins}^k = u.c_I$ with u ($u = 0, 1, 2$) being the number of components inspected at T_k, noting that failed components are not inspected;
- $C_{main}^k = C_p^1 + C_p^2 - CS_{p,p}$ if the two components are jointly, preventively replaced; $C_{main}^k = C_p^i$ if only component i is preventively replaced; $C_{main}^k = C_p^i + C_c^j - CS_{p,c}$ if component i is preventively replaced and component j ($j \neq i$) is correctively replaced; $C_{main}^k = C_c^i$ if only component i is correctively replaced and $C_{main}^k = 0$ if no replacement is performed at T_k.

Obtaining a closed-form expression for the cost-rate in Equation (7.3) is very difficult or even impossible. In (91), an efficient method based on semi-regenerative processes theory is introduced to obtain a closed-form expression for the cost-rate. However, this analytical method is applicable for single-unit degrading systems with time-homogeneous degradation behaviour. Therefore, in this work, the cost-rate is evaluated, given $\Delta T, m_p^1, m_o^1, m_p^2, m_o^2$, using Monte Carlo simulation. By varying the values of the decision variables and performing an exhaustive search, the minimum cost-rate can be identified.

$$C^{\infty}(\Delta T^*, m_p^{1*}, m_o^{1*}, m_p^{2*}, m_o^{2*}) = \min\{C^{\infty}(.)_{0<\Delta T, 0 \,<\, m_p^1 \,\leq\, L^1, 0 \,<\, m_o^1 \,\leq\, m_p^1, 0 \,<\, m_p^2 \,\leq\, L^2, 0 \,<\, m_o^2 \,\leq\, m_p^2}\}. \quad (7.4)$$

7.5 Case Study

We will consider the case study presented in Chapter 4. With multiple interacting components, we saw that the degradation trajectories of each of the components of a new gearbox, whereby all components are new, to be different to those of a partially gearbox, whereby some components are new.

This stochastic dependence, and the economic dependence arising from shared set-up costs, mean that an opportunistic maintenance policy is appropriate. Therefore, in what follows, we show how the opportunistic replacement policy can be i) optimised and ii) used in practice.

It should be noted that the data are scaled and all parameters are given in arbitrary units, either arbitrary cost unit (acu) or arbitrary time unit (atu).

The inspection cost is 10 acu ($c_I = 10$). When each gear is individually replaced, the replacement cost and the maintenance duration are $c_p^1 = c_c^1 = 500$ acu, $c_p^2 = c_c^2 = 600$ acu and $d_1 = d_2 = 1$ atu. When both gears are replaced together, 5% of the total replacement cost of the components is saved ($a = 0.05$) and the total maintenance duration is reduced by 50% ($b = 0.5$). In addition, when the system fails we have to pay 100 acu per downtime unit ($c_d = 100$). The downtime cost (due to system failure) is taken to be the (negative of the) average of the output performance over the period of observation of the system, although in principle the downtime cost could be specified in other ways.

Regarding the state interactions, we will consider the model presented in Chapter 5, and in particular using the estimated parameters on the gearbox data that were shown in Section 5.5.

Next, the fitted degradation model is integrated with the proposed maintenance model to find the optimum policy.

7.5.1 Optimum Maintenance Policy

To evaluate the cost-rate, a very large number of life cycles of the system were simulated with above data. To find the optimal decision parameters $(\Delta T, m_p^1, m_o^1, m_p^2, m_o^2)$, the cost-rate $C^\infty(\Delta T, m_p^1, m_o^1, m_p^2, m_o^2)$ is evaluated for different values of ΔT $(\Delta T > 0)$, m_p^1 $(0 < m_p^1 \leq L^1)$, m_o^1 $(0 < m_o^1 \leq m_p^1)$, m_p^2 $(0 < m_p^2 \leq L^2)$ and m_o^2 $(0 < m_o^2 \leq m_p^2)$ using Equation (7.3). With a precision of 0.010 specified for the cost-rate, the convergence of the cost-rate is reached from 10000 renewal cycles. The optimum values of the decision parameters are $\Delta T^* = 60$, $m_p^{1*} = 0.55$, $m_o^{1*} = 0.50$, $m_p^{2*} = 0.50$ and $m_o^{2*} = 0.40$ with the minimum cost-rate $C^\infty(\Delta T^*, m_p^{1*}, m_o^{1*}, m_p^{2*}, m_o^{2*}) = 2.90$ acu. It is interesting to note that the probability of individual replacement is 0.31 for C1 and 0.38 for C2. The probability of joint replacement is approximately 0.31.

Figure 7.2 shows the relationships between the minimum cost-rate and the inter-inspection interval ΔT for the proposed opportunistic policy (policy V), non-opportunistic policy (policy V1) and the joint replacement policy (policy V2). Each point represents an optimal policy with a given value of ΔT.

It is shown that the proposed opportunistic maintenance policy (policy V) always provides the lowest cost-rate. We observe that when the inspection interval $\Delta T < \Delta T^*$ the maintenance cost increases rapidly with a decreasing ΔT, and that when ΔT gets smaller, the difference between the three policies gets smaller. However, when $\Delta T > \Delta T^*$, the cost-rate of the non-opportunistic policy (policy V1) increases rapidly with an increasing of ΔT. While the cost-rate of policies V and V2 increases slowly with the increasing of ΔT. These interesting results mean that the opportunistic replacement and the joint replacement can better compensate a sub-optimally large ΔT.

7.5.2 Impact of Economic Dependence on the Cost

We now analyse the impact of economic dependence on the opportunistic replacement maintenance policy. This is carried out by analysing the sensitivity of the minimum cost-rate for three policies V, V1 and V2 to the economic dependence degree (a, b) between the two components.

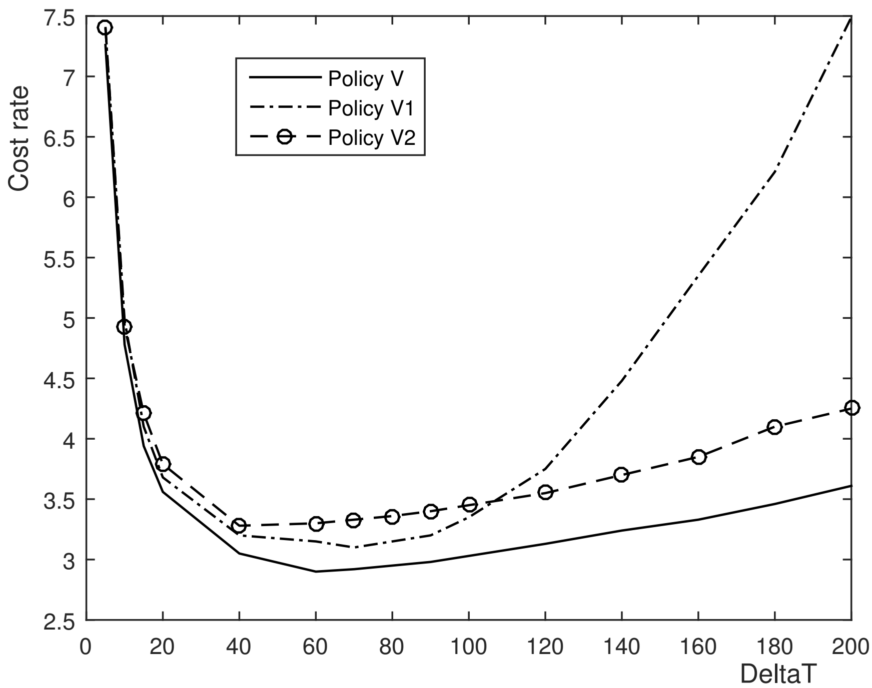

Figure 7.2: Cost-rate as a function of inter-inspection interval ΔT

To study the performance of these three policies, a relative excess-cost in the minimum cost-rate of the proposed opportunistic policy V compared to policy Vi, denoted ΔC_i $(i = 1, 2)$, is used. It is defined as follows:

$$\Delta C_i = \frac{C^{\infty}_{Vi} - C^{\infty}(\Delta T^*, m_p^{1*}, m_o^{1*}, m_p^{2*}, m_o^{2*})}{C^{\infty}_{Vi}} \cdot 100\%$$

where C^{∞}_{Vi} is the minimum cost-rate of policy Vi with $i = 1, 2$. According to the definition, $\Delta C_i > 0$ means that policy V is more effective than policy Vi and less effective in the opposite case.

7.5.2.1 Sensitivity Analysis to a

We vary a from 0 to 20% while the others parameters remain unchanged. For each value of a the minimum cost-rate of each policy is determined and the excess-cost is then evaluated. Summary results are shown in Figure 7.3.

Figure 7.3(a) shows that the cost-rate decreases with the cost-saving factor a. This can be explained by the fact that maintenance costs reduce as a increases. It is not surprising that the proposed opportunistic policy V always provides a

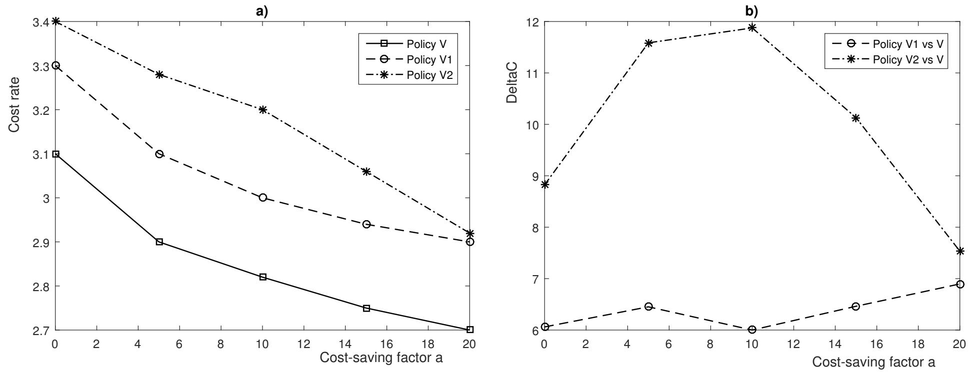

Figure 7.3: Cost-rate (a) and excess-cost (b) as a function of a

lowest cost-rate. This is because policies V1 and V2 are two special cases of policy V.

Figure 7.3(b) shows that when $a < 10\%$ the excess-cost related to policy V2 increases with an increasing of a. This means that the cost-rate of policy V2 decreases more slowly than that of policy V as a increases. However, when $a > 10\%$, the cost-rate of policy V2 decreases more rapidly than the cost-rate of policy V. While the cost-rate of policy V1 decreases more slowly than that of policy V1 with an increasing of a. This can be explained by the fact that when the two components tend to be jointly replaced when the cost-saving factor is high.

To study more the impact of economic dependence degree on the maintenance cost, we consider sensitivity with respect to the duration-saving factor b.

7.5.2.2 Sensitivity Analysis to b

Here we vary b from 0 to 50% while the others parameters remain unchanged. For each value of b the minimum cost-rate of each maintenance policy is determined and the excess-cost is then evaluated. The results obtained are shown in Figure 7.4.

It is not surprising again that an increasing of b (or equivalently a reduction on maintenance duration when two components are replaced together) leads to a decreased cost-rate. However, the effect for both the opportunistic policy (V) and non-opportunistic policy (V1) are broadly the same, in a similar manner to that for a varying a. This suggests that for both policies there is a tendency that replacements of components are simultaneous. This is natural for the opportunistic policy because this is its purpose. However it might have been expected that the

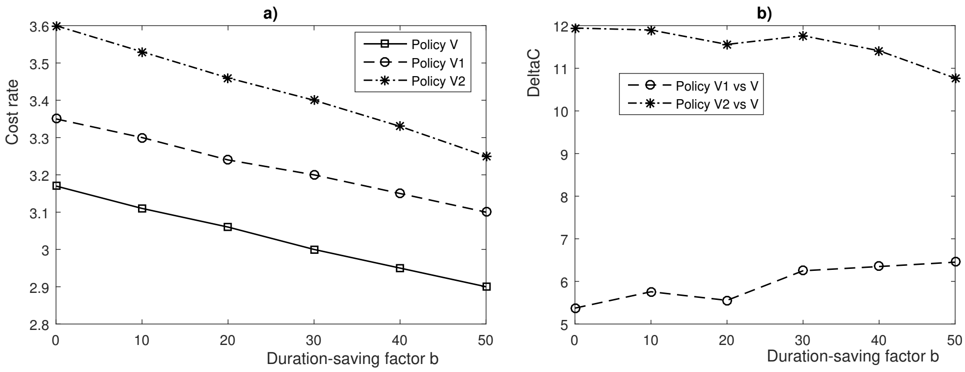

Figure 7.4: Cost-rate (a) and excess-cost (b) as a function of b

non-opportunistic policy to show less dependence on a and b. Our explanation for this is as follows:

When there is no opportunistic replacement, then the threshold for preventive replacement compensates (for component C1 in this case). It is lower (than with the opportunistic policy) so that more often than not, the replacement of components is simultaneous (and set up cost is saved). If it were not the case that replacements are simultaneous then the cost-rate for policy V1 would not depend on a and b in the way it does. This effect only occurs because of the positive stochastic dependence. If there was no positive stochastic dependence then the simultaneous replacement of the components when one reaches a preventive replacement threshold would be inefficient.

Thus, when there is no stochastic dependence between components, opportunistic policies become more effective as the extent of economic dependence increases. This is well known and obvious. However, it would appear that when there is also positive stochastic dependence this phenomenon is much less apparent. This is because a non-opportunistic policy will then compensate for the absence of opportunities for replacement by lowering the threshold for preventive replacement of the components. The positive stochastic dependence ensures that replacements usually remain simultaneous because components will tend to cross their replacement thresholds together. That said, this degrading together phenomenon will tend be more apparent when the lifetimes of the components are broadly similar as is the case for the gearbox system.

7.5.2.3 Optimum Policies When $a = 0$ and $b = 0$

We suppose now that two components are economically independent, i.e., $a = 0$ and $b = 0$. The optimum maintenance policies are given in Table 7.1 where P_{joint} indicates the probability that two components are jointly replaced at each maintenance.

	Optimum decision variables	P_{joint}	Minimum cost-rate
Policy V	$\Delta T^* = 60,$ $m_p^{1*} = 0.6, m_o^{1*} = 0.55,$ $m_p^{2*} = 0.5, m_o^{2*} = 0.45$	0.18	3.22
Policy V1	$\Delta T^* = 60, m_p^{1*} = 0.6, m_p^{2*} = 0.5$	0.12	3.26
Policy V2	$\Delta T^* = 50, m_p^{1*} = 0.55, m_p^{2*} = 0.65$	1	3.69

Table 7.1: Optimum maintenance policies when $a = b = 0$

The obtained results show that when two components are economically independent, the proposed opportunistic policy V is still slightly better than the non-opportunistic policy V1. This is thanks to the opportunistic thresholds which allow policy V to become more flexible and better take into consideration the stochastic dependence between components than the non-opportunistic policy V1. However, the joint replacement (policy V2) leads to a higher cost-rate which means that the joint replacement is not effective for this case.

7.5.3 Impacts of State Dependence on the Cost

To study the impact of state dependence between components on the optimum maintenance policy, we assume now that the degradation process of each component evolves independently. In this way, we could reduce the degradation model to two independent gamma process for which the shape and scale parameters can be estimated using maximum likelihood estimation, or by using the particle filter. The results in the estimates are presented in Table 7.2.

The proposed maintenance policy is then applied. We obtained the optimal decision variables $\Delta T^* = 120$, $m_p^{1*} = 0.60$, $m_o^{1*} = 0.45$, $m_p^{2*} = 0.55$ and $m_o^{2*} = 0.40$. When compared with the results obtained in Section 7.5.1, these optimal values are significantly different. In addition, if we apply these optimal decision variables for the case considering the state dependence between compo-

Component	α^i	β^i
C1	0.1165	0.0100
C2	0.0919	0.0090

Table 7.2: Estimated parameter values without considering stochastic dependence

nents, the cost-rate is then $C^{\infty}(\Delta T^*, m_p^{1*}, m_o^{1*}, m_p^{2*}, m_o^{2*}) = 3.75$ acu which is significantly higher than the one obtained when the state dependence is considered in degradation modelling ((3.75-2.90)/2.90)x100=29.3% higher). This means that not considering the state dependence between two components can lead to a sub-optimal maintenance policy. Of course, the difference is itself dependent on the economic "dependence degree" between the components.

7.6 Discussion

In this chapter, a condition-based maintenance policy for a two-dependent component system was studied. Two kinds of dependency were investigated and integrated in the maintenance modelling: state dependence whereby the degradation rate of each component depends not only on its state but on the state of the other component; and economic dependence whereby set-up cost and duration are shared when components are replaced simultaneously. To select the components to be preventively maintained at each regular time interval, adaptive preventive replacement and opportunistic replacement rules were proposed. A cost model taking into account the economic dependence between components was developed to find the optimal values of the decision variables. The policies were studied in the context of a gearbox system consisting of gears presented in Chapter 4. The results indicated that (i) accounting for the state dependence between components is important, and to ignore it has a significant impact (29.3%) on the cost; (ii) introducing an opportunistic threshold for replacement makes the maintenance policy more flexible and less sensitive to a sub-optimally large inspection interval; and (iii) when there exists positive stochastic dependence between components so that components tend to degrade together, introducing an opportunistic threshold for replacement in order to share set-up costs achieves less when there is positive stochastic dependence between components than when there is not. This is because replacements will tend to be synchronised and this tendency to synchronise arises precisely because of degradation dependence. Thus one might claim a general insight: opportunistic maintenance is less opportune when components tend to degrade together than when they do not.

Chapter 8

Conclusion

8.1 Contributions

This thesis considered multi-component dependencies within prognostics health management (PHM). The emphasis was on accurately modelling the stochastic dependency between components, and predicting their future health state. This is important since all maintenance decision making that follows bases itself on the assessment and predictions of the equipment's health.

This thesis starts with Chapter 2 by providing a literature review and background relevant to the understanding of the works presented later. We presented the different maintenance strategies and gave an overview of the PHM framework. We then discussed the dependencies that can take place between components within a multi-component system. These play an essential role for motivating the work presented in this thesis.

In Chapter 3 we provided an overview on the available experimental platforms and data, highlighting the unavailability of data sets that are suited for studying stochastic dependence between components. This was followed by the development of our gearbox experimental platform which is capable of providing such data. This allows us to conduct a more realistic study on stochastic dependency between components which would lead to more realistic degradation modelling. The first results of the platform clearly showed inter-dependencies between the degrading components. This demonstrates the development of a novel experimental platform capable of studying stochastic dependence between multiple components.

Chapter 4 was dedicated to accurate health indicator extraction from multi-component systems. We presented our methodology for doing so, whereby we made use of different pre-processing steps preparing the vibration signals for the actual processing. We then performed a time-frequency domain analysis via a short time Fourier transform. This resulted in time series data representing the evolution of degradation of the different system components. This was applied after running another experimental configuration of the gearbox experimental platform that was presented in Chapter 3. This experimentation consisted of three runs to failure and was specifically designed so that we could accurately monitor and understand the effect of the degradation interactions that were discovered in Chapter 3. The results demonstrated that a degradation dependence can in fact take place between new and old worn out components, this dependence causes accelerated wear in components which was shown to reduce their lifetime down to 29 %.

In Chapter 5 we presented our development of a novel generic degradation model, which is capable of accounting for the interactions shown in the results of Chapter 4, additionally accounting for the operating condition and intrinsic wear that a component might endure throughout its lifetime. The capability of the model was first shown thorough a numerical simulation which demonstrated similar results to those of the gearbox experimental platform. The model parameters were then estimated using the degradation data of the experimental platform, this was done by via Bayesian inference and sequential Monte Carlo, specifically using the particle filter method. The particle filter was later used to project the degrading state of the components therefore allowing us to perform prognostics using the model, and the observed states of the components.

In Chapter 6, we showed our methodology for performing diagnostics within multi-component system. We performed clustering for uncovering degradation patterns which uses unsupervised machine learning. We then motivated the particular choice of Gaussian mixture models (GMM) in the field of PHM. We used the GMM to cluster the different health states of the gearbox platform and showed that different degradation phases can also be extracted. These phases can indicate accelerated degradation that take place between components that might not be considered otherwise. Therefore the diagnostics methodology presented can be used for detecting incipient faults, and for prior warning of pre-mature system failure.

Finally in Chapter 7, and in light of the work and results that were presented earlier in the thesis, health management for multi-component systems was considered. Different maintenance policies were studied and compared which exposed the importance of accounting for stochastic and economic dependencies within a multi-component system. The maintenance optimisation strategy was validated on the data from the experimental gearbox and using the degradation model presented in Chapter 5. The results showed that ignoring the state dependence between components has a significant impact of 29.3% on the cost. Introducing an opportunistic threshold for replacement makes the maintenance policy more flexible and less sensitive to a sub-optimally large inspection interval. And that when components tend to degrade together introducing an opportunistic threshold for replacement in order to share set-up costs achieves less improvements than when such dependence does not exist.

8.2 Limitations

Although the work presented in this thesis contributes to the PHM community by studying multi-component systems and their dependencies, nevertheless improvements to these works can be conducted.

1. Different configurations of the experimental platform, or a development of more generic platform should be considered for exploring the effects of stochastic dependencies other than the studied degradation dependencies. This could be seen to incorporate different environmental effects into the operation process with varying work loads.

2. Although we proposed the use of a generic data driven degradation model, this can be further improved by integrating the capability of accounting for non numerical data and event data. This could greatly benefit the accuracy of the model when historical data can be provided.

3. Other multi-component system case studies should be considered for studying stochastic dependence between components. This would improve the generality of the findings presented in this thesis. Also, a comparative study of different state of the art prognostics approaches could be conducted in the case of multi-component systems.

4. In this thesis three runs to failure are conducted, where one test is performed in each run. This allow us to determine the stochastic dependence

coefficients which are presented in Chapter 5. However, more robust coefficient values can be identified if more tests are conducted. This would result in a distribution where the average coefficient values can be extracted.

8.3 Future Work

1. Many legacy industrial systems do not have the correct sensor configuration for extracting health indicators for the different components that are present. Therefore works on signal separation can be considered in the context of condition monitoring.

2. As suggested in Chapter 6, Gaussian mixture model clustering results should be further studied since they showed promise for providing means of performing prognostics.

3. Regarding maintenance policies and for a more generic approach, reinforcement machine learning should be considered since degradation simulation can be easily provided using the generic degradation model that we developed. An agent can then be given the objective yo explore different policies under different constrains, ultimately leading to a best policy.

8.4 Main Finding

Throughout this thesis we have investigated prognostics and health management for multi-component systems. The main emphasis was on stochastic dependence, whereby the degradation state of a component can affect the degradation rate of others. Through the provided case study of the gearbox accelerated life testing platform, we have realised that old worn components can influence the degradation rate of new components. This effect is usually neglected when performing PHM, which as shown throughout the thesis can lead to unexpected faults and failures. We have therefore provided a PHM approach that accounts for such issues.

8.5 Closing remarks

Our work on prognostics health management for multi-component systems ultimately contributes to the PHM society by presenting a generic approach for dealing with health indicator extractions; diagnostics and prognosis; and health management of complex systems. This answers a need in PHM since machinery and equipment keeps getting more complex, which is also seeing more integration of sensors and abundance in condition monitoring data.

In addition to this work, we believe that there will be a need for integrating such work within a big data framework, whereby condition monitoring data is transmitted to cloud services and processed. The aggregation of all this condition monitoring data along with event data can eventually lead to more generic and accurate prognostics for different machinery and therefore allowing us to go beyond the development of customised PHM approaches for some specific equipment.

We believe that in the future such a thing might only be accomplished by end to end learning approaches in the likes of what is seen today in deep learning, despite the challenges which were mentioned in Chapter 2. For example the success of such learning techniques has soared in the computer vision field recently, and this is mainly due to its capability of extracting increasingly abstract features from the images (90). Although works on deep learning are being done in PHM, to really reap the power of such techniques, transfer learning should be embraced. This is because the data provided in condition monitoring is mainly unbalanced and failures and fault data are very scarce for some equipment but abundant for others.

Appendix A

A.1 Gamma Distribution

A random variable X which is gamma-distributed with shape α^i and rate β^i is denoted

$$X \sim \Gamma(\alpha^i, \beta^i).$$

The corresponding probability density function (PDF) is

$$f_{\alpha^i,\beta^i}(x) = \frac{1}{\Gamma(\alpha^i)}(\beta^i)^{\alpha^i} x^{\alpha^i-1} e^{-\beta^i x} \mathcal{I}_{\{x\geq 0\}},$$

where:

- $\Gamma(\alpha^i) = \int\limits_0^{+\infty} u^{\alpha^i-1} e^{-u} du$ denotes a complete gamma function;
- $\mathcal{I}_{\{x\geq 0\}}$ is an indicator function. $\mathcal{I}_{\{x\geq 0\}} = 1$ if $x \geq 0$, $\mathcal{I}_{\{x\geq 0\}} = 0$ and otherwise.

References

[1] Frédéric Abrard, Yannick Deville, and Paul White. A new source separation approach based on time-frequency analysis for instantaneous mixtures. *Proc. ECM2S*, pages 259–267, 2001.

[2] Rosmaini Ahmad and Shahrul Kamaruddin. An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63(1):135–149, 2012.

[3] Farzaneh Ahmadzadeh and Jan Lundberg. Remaining useful life estimation. *International Journal of System Assurance Engineering and Management*, 5(4):461–474, 2014.

[4] Suzan Alaswad and Yisha Xiang. A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157:54–63, 2017.

[5] Essam Allam, Ibrahim Ahmed, and Shawki Abouel-Seoud. An experimental investigation of noise emission from a vehicle gearbox system. *Journal of Mechanical Engineering Research*, 3(3):75–84, 2011.

[6] ML Amor-Segan, Ross McMurran, Gunwant Dhadyalla, and RP Jones. Towards the self healing vehicle. In *Automotive Electronics, 2007 3rd Institution of Engineering and Technology Conference on*, pages 1–7. IET, 2007.

[7] Dawn An, Nam H Kim, and Joo-Ho Choi. Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. *Reliability Engineering & System Safety*, 133:223–236, 2015.

[8] Xueli An and Dongxiang Jiang. Bearing fault diagnosis of wind turbine based on intrinsic time-scale decomposition frequency spectrum. *Proceed-*

ings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 228(6):558–566, 2014.

[9] Jérôme Antoni. The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20(2):282–307, 2006.

[10] R Assaf, S Nefti-Meziani, and P Scarf. Unsupervised learning for improving fault detection in complex systems. In *Advanced Intelligent Mechatronics (AIM), 2017 IEEE International Conference on*, pages 1058–1064. IEEE, 2017.

[11] Roy Assaf, Phuc Do, Phil Scarf, and Samia Nefti-Meziani. Wear rate-state interaction modelling for a multi-component system: Models and an experimental platform. *IFAC-PapersOnLine*, 49(28):232–237, 2016.

[12] Roy Assaf, P Do, Philip Scarf, and Samia Nefti-Meziani. Diagnosis for systems with multi-component wear interactions. In *Prognostics and Health Management (ICPHM), 2017 IEEE International Conference on*, pages 96–102. IEEE, 2017.

[13] Howard Austerlitz. *Data acquisition techniques using PCs*. Academic press, 2002.

[14] Suk Joo Bae and Paul H Kvam. A nonlinear random-coefficients model for degradation testing. *Technometrics*, 46(4):460–469, 2004.

[15] Arshdeep Bahga and Vijay Madisetti. *Internet of Things: A hands-on approach*. Vpt, 2014.

[16] DC Baillie and J Mathew. A comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings. *Mechanical Systems and Signal Processing*, 10(1):1–17, 1996.

[17] Jeffrey D Banfield and Adrian E Raftery. Model-based gaussian and non-gaussian clustering. *Biometrics*, pages 803–821, 1993.

[18] Piero Baraldi, Michele Compare, Antoine Despujols, and Enrico Zio. Modelling the effects of maintenance on the degradation of a water-feeding turbo-pump of a nuclear power plant. *Proceedings of the Institution of*

Mechanical Engineers, Part O: Journal of Risk and Reliability, 225(2):169–183, 2011.

[19] Piero Baraldi, Michele Compare, Sergio Sauco, and Enrico Zio. Ensemble neural network-based particle filtering for prognostics. *Mechanical Systems and Signal Processing*, 41(1-2):288–300, 2013.

[20] F. Barbera, H. Schneider, and E. Watson. A condition based maintenance model for a two-unit series system. *European Journal of Operational Research*, 116:281–290, 1999.

[21] Pundarikaksha Baruah and Ratna B Chinnam. Hmms for diagnostics and prognostics in machining processes. *International Journal of Production Research*, 43(6):1275–1293, 2005.

[22] Eric Bechhoefer and Michael Kingsley. A review of time synchronous average algorithms. In *Annual Conference of the Prognostics and Health Management Society, San Diego, CA, Sept*, pages 24–33, 2009.

[23] M El Hachemi Benbouzid. A review of induction motors signature analysis as a medium for faults detection. *IEEE Transactions on Industrial Electronics*, 47(5):984–993, 2000.

[24] Marcus Bengtsson, Erik Olsson, Peter Funk, and Mats Jackson. Design of condition based maintenance system—a case study using sound analysis and case-based reasoning. *Condition Based Maintenance Systems—An Investigation of Technical Constituents and Organizational Aspects; Malardalen University: Eskilstuna, Sweden*, page 57, 2004.

[25] T Benkedjouh, Kamal Medjaher, Noureddine Zerhouni, and Saïd Rechak. Health assessment and life prediction of cutting tools based on support vector regression. *Journal of Intelligent Manufacturing*, 26(2):213–223, 2015.

[26] James C Bezdek. Objective function clustering. In *Pattern recognition with fuzzy objective function algorithms*, pages 43–93. Springer, 1981.

[27] L. Bian and N. Gebraeel. Stochastic framework for partially degradation systems with continuous component degradation-rate-interactions. *Naval Research Logistics*, 61:286–303, 2014.

[28] L. Bian and N. Gebraeel. Stochastic modeling and real-time prognostics for multi-component systems with degradation rate interactions. *IIE Transactions*, 46:470–482, 2014.

[29] Linkan Bian and Nagi Gebraeel. Stochastic framework for partially degradation systems with continuous component degradation-rate-interactions. *Naval Research Logistics (NRL)*, 61(4):286–303, 2014.

[30] Zygmund William Birnbaum. On the importance of different components in a multicomponent system. Technical report, Washington Univ Seattle Lab of Statistical Research, 1968.

[31] Benjamin J Blaiszik, Sharlotte LB Kramer, Solar C Olugebefola, Jeffrey S Moore, Nancy R Sottos, and Scott R White. Self-healing polymers and composites. *Annual Review of Materials Research*, 40:179–211, 2010.

[32] G Wesley Blankenship and Rajendra Singh. Analytical solution for modulation sidebands associated with a class of mechanical oscillators. *Journal of Sound and Vibration*, 179(1):13–36, 1995.

[33] Boualem Boashash and Peter Black. An efficient real-time implementation of the wigner-ville distribution. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(11):1611–1618, 1987.

[34] IP Bond and IR Farrow. Fatigue life prediction under complex loading for xas/914 cfrp incorporating a mechanical fastener. *International Journal of Fatigue*, 22(8):633–644, 2000.

[35] Keomany Bouvard, Samuel Artus, Christophe Bérenguer, and Vincent Cocquempot. Condition-based dynamic maintenance operations planning & grouping. application to commercial heavy vehicles. *Reliability Engineering & System Safety*, 96(6):601–610, 2011.

[36] Edward R Brown, Neal N McCollom, Erin-Elaine Moore, and Andy Hess. Prognostics and health management a data-driven approach to supporting the f-35 lightning ii. In *Aerospace Conference, 2007 IEEE*, pages 1–12. IEEE, 2007.

[37] RJ Bucci. Development of a proposed astm standard test method for near-threshold fatigue crack growth rate measurement. In *Fatigue Crack Growth Measurement and Data Analysis*. ASTM International, 1981.

[38] Carey Bunks, Dan McCarthy, and Tarik Al-Ani. Condition-based maintenance of machines using hidden markov models. *Mechanical Systems and Signal Processing*, 14(4):597–612, 2000.

[39] Carl S Byington, Matthew Watson, and Doug Edwards. Data-driven neural network methodology to remaining life predictions for aircraft actuator components. In *Aerospace Conference, 2004. Proceedings. 2004 IEEE*, volume 6, pages 3581–3589. IEEE, 2004.

[40] Fatih Camci and Ratna Babu Chinnam. Health-state estimation and prognostics in machining processes. *IEEE Transactions on Automation Science and Engineering*, 7(3):581–597, 2010.

[41] E Peter Carden and Paul Fanning. Vibration based condition monitoring: a review. *Structural Health Monitoring*, 3(4):355–377, 2004.

[42] MJ Carr and W Wang. A case comparison of a proportional hazards model and a stochastic filter for condition-based maintenance applications using oil-based condition monitoring information. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 222(1): 47–55, 2008.

[43] B. Castenier, A. Grall, and C. Berenguer. A condition-based maintenance policy with non-periodic inspections for a two-unit series system. *Reliability Engineering and System Safety*, 87:109–120, 2005.

[44] Carlo Cecati. A survey of fault diagnosis and fault-tolerant techniques—part ii: fault diagnosis with knowledge-based and hybrid/active approaches. *IEEE Transactions on Industrial Electronics*, 2015.

[45] Jose Celaya, Abhinav Saxena, Sankalita Saha, and Kai F Goebel. Prognostics of power mosfets under thermal stress accelerated aging using data-driven and model-based methodologies. In *Prognostics and Health Management*, 2011.

[46] Gilles Celeux and Gérard Govaert. Gaussian parsimonious clustering models. *Pattern recognition*, 28(5):781–793, 1995.

[47] Elaine Y Chan, Wai Ki Ching, Michael K Ng, and Joshua Z Huang. An optimization algorithm for clustering using weighted dissimilarity measures. *Pattern Recognition*, 37(5):943–952, 2004.

[48] Binqiang Chen, Zhousuo Zhang, Chuang Sun, Bing Li, Yanyang Zi, and Zhengjia He. Fault feature extraction of gearbox by using overcomplete rational dilation discrete wavelet transform on signals measured from vibration sensors. *Mechanical Systems and Signal Processing*, 33:275–298, 2012.

[49] Shunfeng Cheng, Michael H Azarian, and Michael G Pecht. Sensor systems for prognostics and health management. *Sensors*, 10(6):5774–5797, 2010.

[50] FK Choy, V Polyshchuk, JJ Zakrajsek, RF Handschuh, and DP Townsend. Analysis of the effects of surface pitting and wear on the vibration of a gear transmission system. *Tribology International*, 29(1):77–83, 1996.

[51] M Bishop Christopher. *Pattern recognition and machine learning*. Springer-Verlag New York, 2016.

[52] Charles K Chui. *An introduction to wavelets*. Elsevier, 2016.

[53] Martin Crowder and Jerald Lawless. On a scheme for predictive maintenance. *European Journal of Operational Research*, 176(3):1713–1722, 2007.

[54] M Dalal, J Ma, and D He. Lithium-ion battery life prognostic health management system using particle filtering framework. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 225(1):81–90, 2011.

[55] G Dalpiaz, A Rivola, and R Rubini. Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears. *Mechanical Systems and Signal Processing*, 14(3):387–412, 2000.

[56] Abhijit Dasgupta and Adrian E Raftery. Detecting features in spatial point processes with clutter via model-based clustering. *Journal of the American Statistical Association*, 93(441):294–302, 1998.

[57] Rommert Dekker, Ralph E Wildeman, and Frank A Van der Duyn Schouten. A review of multi-component maintenance models with economic dependence. *Mathematical Methods of Operations Research*, 45 (3):411–435, 1997.

[58] Vladimir Dekys, Peter Kalman, Pavel Hanak, Pavol Novak, and Zuzana Stankovicova. Determination of vibration sources by using stft. *Procedia Engineering*, 177:496–501, 2017.

[59] John R Deller Jr, John G Proakis, and John H Hansen. *Discrete time processing of speech signals*. Prentice Hall PTR, 1993.

[60] Estelle Deloux, Bruno Castanier, and Christophe Bérenguer. Predictive maintenance policy for a gradually deteriorating system subject to stress. *Reliability Engineering & System Safety*, 94(2):418–431, 2009.

[61] Patricia Derler, Edward A Lee, and Alberto Sangiovanni Vincentelli. Modeling cyber–physical systems. *Proceedings of the IEEE*, 100(1):13–28, 2012.

[62] Thomas G Dietterich. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15. Springer, 2000.

[63] P. Do, A. Voisin, E. Levrat, and B. Iung. A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. *Reliability Engineering and System Safety*, 133:22–32, 2015.

[64] P. Do, H.-C. Vu, A. Barros, and C. Berenguer. Maintenance grouping for multi-component systems with availability constraints and limited maintenance teams. *Reliability Engineering and System Safety*, 142:56–67, 2015.

[65] Phuc Do, Phil Scarf, and Benoit Iung. Condition-based maintenance for a two-component system with dependencies. *IFAC-PapersOnLine*, 48(21): 946–951, 2015.

[66] Phuc Do, Alexandre Voisin, Eric Levrat, and Benoit Iung. A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. *Reliability Engineering & System Safety*, 133:22–32, 2015.

[67] P. Do Van, A. Barros, C. Berenguer, K. Bouvard, and F. Brissaud. Dynamic grouping maintenance strategy with time limited opportunities. *Reliability Engineering and System Safety*, 120:51–59, 2013.

[68] Pedro Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87, 2012.

[69] Ming Dong and David He. A segmental hidden semi-markov model (hsmm)-based diagnostics and prognostics framework and methodology. *Mechanical Systems and Signal Processing*, 21(5):2248–2266, 2007.

[70] Ming Dong, David He, Prashant Banerjee, and Jonathan Keller. Equipment health diagnosis and prognosis using hidden semi-markov models. *The International Journal of Advanced Manufacturing Technology*, 30(7-8): 738–749, 2006.

[71] Arnaud Doucet and Adam M Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12 (656-704):3, 2009.

[72] Antonio Garcia Espinosa, Javier A Rosero, Jordi Cusido, Luis Romeral, and Juan Antonio Ortega. Fault detection by means of hilbert–huang transform of the stator current in a pmsm with demagnetization. *IEEE Transactions on Energy Conversion*, 25(2):312–318, 2010.

[73] Peter C Evans and Marco Annunziata. Industrial internet: Pushing the boundaries, 2012.

[74] Yanqiong Fei and Chengyuan Wang. Self-repairing algorithm of lattice-type self-reconfigurable modular robots. *Journal of Intelligent & Robotic Systems*, 75(2):193–203, 2014.

[75] Dimitar P Filev, Ratna Babu Chinnam, Finn Tseng, and Pundarikaksha Baruah. An industrial strength novelty detection framework for autonomous equipment monitoring and diagnostics. *IEEE Transactions on Industrial Informatics*, 6(4):767–779, 2010.

[76] Fiorenzo Filippetti, Giovanni Franceschini, Carla Tassoni, and Peter Vas. Recent developments of induction motor drives fault diagnosis using ai techniques. *IEEE Transactions on Industrial Electronics*, 47(5):994–1004, 2000.

[77] Olga Fink, Enrico Zio, and Ulrich Weidmann. Fuzzy classification with restricted boltzman machines and echo-state networks for predicting potential railway door system failures. *IEEE Transactions on Reliability*, 64 (3):861–868, 2015.

[78] Regina Frei, Richard McWilliam, Benjamin Derrick, Alan Purvis, Ashutosh Tiwari, and Giovanna Di Marzo Serugendo. Self-healing and self-repairing technologies. *The International Journal of Advanced Manufacturing Technology*, 69(5-8):1033–1061, 2013.

[79] Dennis Gabor. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441, 1946.

[80] Zhiwei Gao, Carlo Cecati, and Steven X Ding. A survey of fault diagnosis and fault-tolerant techniques—part i: Fault diagnosis with model-based and signal-based approaches. *IEEE Transactions on Industrial Electronics*, 62(6):3757–3767, 2015.

[81] Mari Cruz Garcia, Miguel A Sanz-Bobi, and Javier del Pico. Simap: Intelligent system for predictive maintenance: Application to the health condition monitoring of a windturbine gearbox. *Computers in Industry*, 57(6):552–568, 2006.

[82] William A Gardner, Antonio Napolitano, and Luigi Paura. Cyclostationarity: Half a century of research. *Signal Processing*, 86(4):639–697, 2006.

[83] Nagi Gebraeel and Jing Pan. Prognostic degradation models for computing and updating residual life distributions in a time-varying environment. *IEEE Transactions on Reliability*, 57(4):539–550, 2008.

[84] Nagi Gebraeel, Mark Lawley, Richard Liu, and Vijay Parmeshwaran. Residual life predictions from vibration-based degradation signals: a neural network approach. *IEEE Transactions on Industrial Electronics*, 51(3):694–700, 2004.

[85] Nagi Z Gebraeel, Mark A Lawley, Rong Li, and Jennifer K Ryan. Residual-life distributions from component degradation signals: A bayesian approach. *IiE Transactions*, 37(6):543–557, 2005.

[86] Omid Geramifard, Jian-Xin Xu, Jun-Hong Zhou, and Xiang Li. A physically segmented hidden markov model approach for continuous tool condition monitoring: Diagnostics and prognostics. *IEEE Transactions on Industrial Informatics*, 8(4):964–973, 2012.

[87] Alireza Ghasemi, Soumaya Yacout, and M-Salah Ouali. Evaluating the reliability function and the mean residual life for equipment with unobservable states. *IEEE Transactions on Reliability*, 59(1):45–54, 2010.

[88] Kai Goebel, Bhaskar Saha, Abhinav Saxena, Jose R Celaya, and Jon P Christophersen. Prognostics in battery health management. *IEEE Instrumentation & Measurement Magazine*, 11(4), 2008.

[89] H. R. Golmakani and H. Moakedi. Periodic inspection optimization model for a two-component repairable system with failure interaction. *Computers and Industrial Engineering*, 63(3):540–545, 2012.

[90] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[91] A. Grall, L. Dieulle, C. Bérenguer, and M. Roussignol. Continuous-time predective-maintenance scheduling for a deteriorating system. *IEEE Transactions on Reliability*, 51:141–150, 2002.

[92] Antoine Grall, Christophe Bérenguer, and Laurence Dieulle. A condition-based maintenance policy for stochastically deteriorating systems. *Reliability Engineering & System Safety*, 76(2):167–180, 2002.

[93] Peter J Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[94] Limeng Guo, Yu Peng, Datong Liu, and Yue Luo. Comparison of resampling algorithms for particle filter based remaining useful life estimation. In *Prognostics and Health Management (PHM), 2014 IEEE Conference on*, pages 1–8. IEEE, 2014.

[95] Runxia Guo and Quan Gan. Prognostics for a leaking hydraulic actuator based on the f-distribution particle filter. *IEEE Access*, 5:22409–22420, 2017.

[96] Yabin Guo, Guannan Li, Huanxin Chen, Jiangyu Wang, Mengru Guo, Shaobo Sun, and Wenju Hu. Optimized neural network-based fault diagnosis strategy for vrf system in heating mode using data mining. *Applied Thermal Engineering*, 125:1402–1413, 2017.

[97] Li Hao, Nagi Gebraeel, and Jianjun Shi. Simultaneous signal separation and prognostics of multi-component systems: the case of identical components. *IIE Transactions*, 47(5):487–504, 2015.

[98] William Hardman. Mechanical and propulsion systems prognostics: Us navy strategy and demonstration. *JOM*, 56(3):21–27, 2004.

[99] Fredric J Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–83, 1978.

[100] Wei He, Nicholas Williard, Michael Osterman, and Michael Pecht. Prognostics of lithium-ion batteries based on dempster–shafer theory and the bayesian monte carlo method. *Journal of Power Sources*, 196(23):10314–10321, 2011.

[101] Aiwina Heng, Andy CC Tan, Joseph Mathew, Neil Montgomery, Dragan Banjevic, and Andrew KS Jardine. Intelligent condition-based prediction of machinery reliability. *Mechanical Systems and Signal Processing*, 23(5): 1600–1614, 2009.

[102] Aiwina Heng, Sheng Zhang, Andy CC Tan, and Joseph Mathew. Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical systems and signal processing*, 23(3):724–739, 2009.

[103] Andrew Hess, Giulio Calvello, and T Dabney. Phm a key enabler for the jsf autonomic logistics support concept. In *Aerospace Conference, 2004. Proceedings. 2004 IEEE*, volume 6, pages 3543–3550. IEEE, 2004.

[104] D Ho and RB Randall. Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mechanical Systems and Signal Processing*, 14(5):763–788, 2000.

[105] HP Hong, Wenxing Zhou, S Zhang, and W Ye. Optimal condition-based maintenance decisions for systems with dependent stochastic degradation of components. *Reliability Engineering & System Safety*, 121:276–288, 2014.

[106] Chao Hu, Byeng D Youn, Pingfeng Wang, and Joung Taek Yoon. Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliability Engineering & System Safety*, 103:120–135, 2012.

[107] Yang Hu, Piero Baraldi, Francesco Di Maio, and Enrico Zio. A particle filtering and kernel smoothing-based approach for new design component prognostics. *Reliability Engineering & System Safety*, 134:19–31, 2015.

[108] Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 454, pages 903–995. The Royal Society, 1998.

[109] Norden Eh Huang. *Hilbert-Huang transform and its applications*, volume 16. World Scientific, 2014.

[110] Runqing Huang, Lifeng Xi, Xinglin Li, C Richard Liu, Hai Qiu, and Jay Lee. Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. *Mechanical Systems and Signal Processing*, 21(1):193–207, 2007.

[111] RW Hyers, JG McGowan, KL Sullivan, JF Manwell, and BC Syrett. Condition monitoring and prognosis of utility scale wind turbines. *Energy Materials*, 1(3):187–203, 2006.

[112] B. Iung, P. Do, E. Levrat, and A. Voisin. Opportunistic maintenance based on multi-dependent components of manufacturing systems. *CIRP Annals Manufacturing Technology*, 65(1):401–404, 2016.

[113] G. Yu J. Lin J. Lee, H. Qiu and Rexnord Technical Services. Bearing data set. NASA Ames Research Center, Moffett Field, CA, 2007.

[114] Anil K Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, 2000.

[115] Andrew KS Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7):1483–1510, 2006.

[116] Kamran Javed, Rafael Gouriveau, and Noureddine Zerhouni. State of the art and taxonomy of prognostics approaches, trends of prognostics appli-

cations and open issues towards maturity at different technology readiness levels. *Mechanical Systems and Signal Processing*, 94:214–236, 2017.

[117] Pratesh Jayaswal, AK Wadhwani, and KB Mulchandani. Machine fault signature analysis. *International Journal of Rotating Machinery*, 2008, 2008.

[118] Douglas L Jones and Richard G Baraniuk. A simple scheme for adapting time-frequency representations. *IEEE Transactions on Signal Processing*, 42(12):3530–3535, 1994.

[119] Marine Jouin, Rafael Gouriveau, Daniel Hissel, Marie-Cécile Péra, and Noureddine Zerhouni. Particle filter-based prognostics: Review, discussion and perspectives. *Mechanical Systems and Signal Processing*, 72:2–31, 2016.

[120] Robert C Juvinall and Kurt M Marshek. *Fundamentals of machine component design*, volume 1. Wiley New York, 2000.

[121] Shubha Kadambe. On the window selection and the cross terms that exist in the magnitude squared distribution of the short time fourier transform. In *Statistical Signal and Array Processing, 1992. Conference Proceedings., IEEE Sixth SP Workshop on*, pages 22–25. IEEE, 1992.

[122] Visakan Kadirkamanathan, Ping Li, Mohamed H Jaward, and Simon G Fabri. Particle filtering-based fault detection in non-linear stochastic systems. *International Journal of Systems Science*, 33(4):259–265, 2002.

[123] Chinmaya Kar and AR Mohanty. Gearbox health monitoring through multiresolution fourier transform of vibration and current signals. *Structural Health Monitoring*, 5(2):195–200, 2006.

[124] Dean C Karnopp, Donald L Margolis, and Ronald C Rosenberg. *System dynamics: modeling, simulation, and control of mechatronic systems*. John Wiley & Sons, 2012.

[125] Minou CA Olde Keizer, Simme Douwe P Flapper, and Ruud H Teunter. Condition-based maintenance policies for systems with multiple dependent components: A review. *European Journal of Operational Research*, 2017.

[126] Jeffrey P Kharoufeh and Jeffrey A Sipe. Evaluating failure time probabilities for a markovian wear process. *Computers & Operations Research*, 32 (5):1131–1145, 2005.

[127] Jeffrey P Kharoufeh, Christopher J Solo, and M Yasin Ulukus. Semi-markov models for degradation-based reliability. *IIE Transactions*, 42(8):599–612, 2010.

[128] Nam-Ho Kim, Dawn An, and Joo-Ho Choi. *Prognostics and Health Management of Engineering Systems*. Springer, 2017.

[129] Nikolay V Kirianaki, Sergey Y Yurish, Nestor O Shpak, and Vadim P Deynega. *Data Acquisition and Signal Processing for Smart Sensors*. Wiley Online Library, 2002.

[130] Khairy Ahmed Helmy Kobbacy and DN Prabhakar Murthy. *Complex system maintenance handbook*. Springer Science & Business Media, 2008.

[131] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[132] P Konar and P Chattopadhyay. Bearing fault detection of induction motor using wavelet and support vector machines (svms). *Applied Soft Computing*, 11(6):4203–4211, 2011.

[133] Ranganath Kothamasu, Samuel H Huang, and William H VerDuin. System health monitoring and prognostics—a review of current paradigms and practices. *The International Journal of Advanced Manufacturing Technology*, 28(9-10):1012–1024, 2006.

[134] Helge Langseth and Luigi Portinale. Bayesian networks in reliability. *Reliability Engineering & System Safety*, 92(1):92–108, 2007.

[135] Khanh Le Son, Mitra Fouladirad, Anne Barros, Eric Levrat, and Benoît Iung. Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety*, 112:165–175, 2013.

[136] Bruno P Leao, Takashi Yoneyama, Guilherme C Rocha, and Kevin T Fitzgibbon. Prognostics performance metrics and their relation to require-

ments, design, verification and cost-benefit. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pages 1–8. IEEE, 2008.

[137] Mitchell Lebold, Katherine McClintic, Robert Campbell, Carl Byington, and Kenneth Maynard. Review of vibration analysis methods for gearbox diagnostics and prognostics. In *Proceedings of the 54th meeting of the Society for Machinery Failure Prevention Technology*, volume 634, page 16, 2000.

[138] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436, 2015.

[139] Jay Lee, Fangji Wu, Wenyu Zhao, Masoud Ghaffari, Linxia Liao, and David Siegel. Prognostics and health management design for rotary machinery systems—reviews, methodology and applications. *Mechanical systems and signal processing*, 42(1-2):314–334, 2014.

[140] Yaguo Lei. *Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery*. Butterworth-Heinemann, 2016.

[141] Yaguo Lei, Zhengjia He, Yanyang Zi, and Xuefeng Chen. New clustering algorithm-based fault diagnosis using compensation distance evaluation technique. *Mechanical Systems and Signal Processing*, 22(2):419–435, 2008.

[142] Yaguo Lei, Naipeng Li, Jing Lin, and Sizhe Wang. Fault diagnosis of rotating machinery based on an adaptive ensemble empirical mode decomposition. *Sensors*, 13(12):16950–16964, 2013.

[143] Yaguo Lei, Jing Lin, Dong Han, and Zhengjia He. An enhanced stochastic resonance method for weak feature extraction from vibration signals in bearing fault detection. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 228(5):815–827, 2014.

[144] A Leyland and A Matthews. On the significance of the h/e ratio in wear control: a nanocomposite coating approach to optimised tribological behaviour. *Wear*, 246(1-2):1–11, 2000.

[145] Bo Li, M-Y Chow, Yodyium Tipsuwan, and James C Hung. Neural-network-based motor rolling bearing fault diagnosis. *IEEE transactions on Industrial Electronics*, 47(5):1060–1069, 2000.

[146] Chuan Li, René-Vinicio Sanchez, Grover Zurita, Mariela Cerrada, Diego Cabrera, and Rafael E Vásquez. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mechanical Systems and Signal Processing*, 76:283–293, 2016.

[147] Heping Li, Estelle Deloux, and Laurence Dieulle. A condition-based maintenance policy for multi-component systems with lévy copulas dependence. *Reliability Engineering & System Safety*, 149:44–55, 2016.

[148] Hui Li, Yuping Zhang, and Haiqi Zheng. Gear fault detection and diagnosis under speed-up condition based on order cepstrum and radial basis function neural network. *Journal of Mechanical Science and Technology*, 23(10): 2780–2789, 2009.

[149] Ping Li and Visakan Kadirkamanathan. Particle filtering based likelihood ratio approach to fault diagnosis in nonlinear stochastic systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 31(3):337–343, 2001.

[150] Ping Li, Roger Goodall, and Visakan Kadirkamanathan. Parameter estimation of railway vehicle dynamic model using rao-blackwellised particle filter. In *European Control Conference (ECC), 2003*, pages 2384–2389. IEEE, 2003.

[151] YG Li and P Nilkitsaranont. Gas turbine performance prognostic for condition-based maintenance. *Applied Energy*, 86(10):2152–2161, 2009.

[152] Zhenglin Liang, Ajith Kumar Parlikad, Rengarajan Srinivasan, and Nipat Rasmekomen. On fault propagation in deterioration of multi-component systems. *Reliability Engineering & System Safety*, 162:72–80, 2017.

[153] Haitao Liao and Elsayed A Elsayed. Reliability inference for field conditions from accelerated degradation testing. *Naval Research Logistics (NRL)*, 53 (6):576–587, 2006.

[154] Linxia Liao and Felix Köttig. Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. *IEEE Transactions on Reliability*, 63(1):191–207, 2014.

[155] Man Ho Ling, Kwok Leung Tsui, and Narayanaswamy Balakrishnan. Accelerated degradation analysis for the quality of a system based on the gamma process. *IEEE Transactions on Reliability*, 64(1):463–472, 2015.

[156] L Liu, M Yu, Y. Maa, and Y. Tu. Economic and economic-statistical designs of an x control chart for two-unit series systems with condition-based maintenance. *European Journal of Operational Research*, 226:491–499, 2013.

[157] Xiao Liu, Jingrui Li, Khalifa N Al-Khalifa, Abdelmagid S Hamouda, David W Coit, and Elsayed A Elsayed. Condition-based maintenance for continuously monitored degrading systems with multiple failure modes. *IIE Transactions*, 45(4):422–435, 2013.

[158] Ariane Lorton, Mitra Fouladirad, and Antoine Grall. A methodology for probabilistic model-based prognosis. *European Journal of Operational Research*, 225(3):443–454, 2013.

[159] Bin Lu, Yaoyu Li, Xin Wu, and Zhongzhou Yang. A review of recent advances in wind turbine condition monitoring and fault diagnosis. In *Power Electronics and Machines in Wind Applications, 2009. PEMWA 2009. IEEE*, pages 1–7. IEEE, 2009.

[160] C Joseph Lu and William O Meeker. Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 35(2):161–174, 1993.

[161] James MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA, 1967.

[162] Abd Kadir Mahamad, Sharifah Saon, and Takashi Hiyama. Predicting remaining useful life of rotating machinery based artificial neural network. *Computers & Mathematics with Applications*, 60(4):1078–1087, 2010.

[163] Arnaz Malhi and Robert X Gao. Pca-based feature selection scheme for machine defect classification. *IEEE Transactions on Instrumentation and Measurement*, 53(6):1517–1525, 2004.

[164] KF Martin. A review by discussion of condition monitoring and fault diagnosis in machine tools. *International Journal of Machine Tools and Manufacture*, 34(4):527–551, 1994.

[165] João F Martins, V Ferno Pires, and AJ Pires. Unsupervised neural-network-based algorithm for an on-line diagnosis of three-phase induction motor stator fault. *IEEE Transactions on Industrial Electronics*, 54(1):259–264, 2007.

[166] Kamal Medjaher, Diego Alejandro Tobon-Mejia, and Noureddine Zerhouni. Remaining useful life estimation of critical components with application to bearings. *IEEE Transactions on Reliability*, 61(2):292–302, 2012.

[167] Robert E Melchers and André T Beck. *Structural reliability analysis and prediction*. John Wiley & Sons, 2017.

[168] Qiang Miao, Lei Xie, Hengjuan Cui, Wei Liang, and Michael Pecht. Remaining useful life prediction of lithium-ion battery with unscented particle filter technique. *Microelectronics Reliability*, 53(6):805–810, 2013.

[169] Satoshi Murata, Eiichi Yoshida, Haruhisa Kurokawa, Kohji Tomita, and Shigeru Kokaji. Self-repairing mechanical systems. *Autonomous Robots*, 10 (1):7–21, 2001.

[170] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[171] DNP Murthy and DG Nguyen. Study of two-component system with failure interaction. *Naval Research Logistics (NRL)*, 32(2):239–247, 1985.

[172] DNP Murthy and DG Nguyen. Study of a multi-component system with failure interaction. *European Journal of Operational Research*, 21(3):330–338, 1985.

[173] Patrick Nectoux, Rafael Gouriveau, Kamal Medjaher, Emmanuel Ramasso, Brigitte Chebel-Morello, Noureddine Zerhouni, and Christophe Varnier.

Pronostia: An experimental platform for bearings accelerated degradation tests. In *IEEE International Conference on Prognostics and Health Management, PHM'12.*, pages 1–8. IEEE, 2012.

[174] Raymond T Ng and Jiawei Han. E cient and e ective clustering methods for spatial data mining. In *Proceedings of VLDB*, pages 144–155. Citeseer, 1994.

[175] K.-A. Nguyen, P. Do, and A. Grall. Condition-based maintenance for multi-component systems using importance measure and predictive information. *International Journal of Systems Science: Operations & Logistics*, 1(4): 228–45, 2014.

[176] Kim-Anh Nguyen, Phuc Do, and Antoine Grall. Condition-based maintenance for multi-component systems using importance measure and predictive information. *International Journal of Systems Science: Operations & Logistics*, 1(4):228–245, 2014.

[177] Robin P Nicolai and Rommert Dekker. Optimal maintenance of multi-component systems: a review. In *Complex System Maintenance Handbook*, pages 263–286. Springer, 2008.

[178] Robin P Nicolai, Johannes Bartholomeus Gerardus Frenk, and Rommert Dekker. Modelling and optimizing imperfect maintenance of coatings on steel structures. *Structural Safety*, 31(3):234–244, 2009.

[179] R.P. Nicolai and R. Dekker. Optimal maintenance of multi-component systems: a review. *Complex System Maintenance Handbook, London: Springer*, pages 263–286, 2008.

[180] Mengyan Nie and Ling Wang. Review of condition monitoring and fault diagnosis technologies for wind turbine gearbox. *Procedia Cirp*, 11:287–290, 2013.

[181] Benjamín E Olivares, Matias A Cerda Munoz, Marcos E Orchard, and Jorge F Silva. Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. *IEEE Transactions on Instrumentation and Measurement*, 62 (2):364–376, 2013.

[182] Marcos Orchard, Gregory Kacprzynski, Kai Goebel, Bhaskar Saha, and George Vachtsevanos. Advances in uncertainty representation and management for particle filtering applied to prognostics. In *Prognostics and Health Management, PHM 2008. International Conference on*, pages 1–6. IEEE, 2008.

[183] Marcos E Orchard and George J Vachtsevanos. A particle-filtering approach for on-line fault diagnosis and failure prognosis. *Transactions of the Institute of Measurement and Control*, 31(3-4):221–246, 2009.

[184] Ming Ouyang, William J Welsh, and Panos Georgopoulos. Gaussian mixture clustering and imputation of microarray data. *Bioinformatics*, 20(6): 917–923, 2004.

[185] Hasan Ozturk, Isa Yesilyurt, and Mustafa Sabuncu. Detection and advancement monitoring of distributed pitting failure in gears. *Journal of Nondestructive Evaluation*, 29(2):63–73, 2010.

[186] Fannia Pacheco, José Valente de Oliveira, René-Vinicio Sánchez, Mariela Cerrada, Diego Cabrera, Chuan Li, Grover Zurita, and Mariano Artés. A statistical comparison of neuroclassifiers and feature selection methods for gearbox fault diagnosis under realistic conditions. *Neurocomputing*, 194: 192–206, 2016.

[187] Zhengqiang Pan and Narayanaswamy Balakrishnan. Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. *Reliability Engineering & System Safety*, 96(8):949–957, 2011.

[188] MD Pandey, X-X Yuan, and JM Van Noortwijk. The influence of temporal uncertainty of deterioration on life-cycle management of structures. *Structure and Infrastructure Engineering*, 5(2):145–156, 2009.

[189] BK Panigrahi and V Ravikumar Pandi. Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm. *IET Feneration, Transmission & Distribution*, 3(3):296–306, 2009.

[190] Michael Pecht. *Prognostics and health management of electronics*. Wiley Online Library, 2008.

[191] Michael Pecht and Jie Gu. Physics-of-failure-based prognostics for electronic products. *Transactions of the Institute of Measurement and Control*, 31(3-4):309–322, 2009.

[192] Michael Pecht and Rubyca Jaai. A prognostics and health management roadmap for information and electronics-rich systems. *Microelectronics Reliability*, 50(3):317–323, 2010.

[193] Ying Peng, Ming Dong, and Ming Jian Zuo. Current status of machine prognostics in condition-based maintenance: a review. *The International Journal of Advanced Manufacturing Technology*, 50(1-4):297–313, 2010.

[194] Sanna Poyhonen, Pedro Jover, and Heikki Hyotyniemi. Signal processing of vibrations for condition monitoring of an induction motor. In *Control, Communications and Signal Processing, 2004. First International Symposium on*, pages 499–502. IEEE, 2004.

[195] Miguel Delgado Prieto, Giansalvo Cirrincione, Antonio Garcia Espinosa, Juan Antonio Ortega, and Humberto Henao. Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. *IEEE Transactions on Industrial Electronics*, 60(8):3398–3407, 2013.

[196] Matthieu Puigt and Yannick Deville. Time–frequency ratio-based blind separation methods for attenuated and time-delayed sources. *Mechanical Systems and Signal Processing*, 19(6):1348–1379, 2005.

[197] Hai Qiu, Jay Lee, Jing Lin, and Gang Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289(4):1066–1090, 2006.

[198] Jing Qiu, Brij B Seth, Steven Y Liang, and Cheng Zhang. Damage mechanics approach for bearing lifetime prognostics. *Mechanical systems and signal processing*, 16(5):817–829, 2002.

[199] Akhand Rai and SH Upadhyay. A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribology International*, 96:289–306, 2016.

[200] VK Rai and AR Mohanty. Bearing fault diagnosis using fft of intrinsic mode functions in hilbert–huang transform. *Mechanical Systems and Signal Processing*, 21(6):2607–2615, 2007.

[201] A Santhana Raj and Nagarajan Murali. Early classification of bearing faults using morphological operators and fuzzy inference. *IEEE Transactions on Industrial Electronics*, 60(2):567–574, 2013.

[202] Robert Bond Randall. *Vibration-based condition monitoring: industrial, aerospace and automotive applications*. John Wiley & Sons, 2011.

[203] N. Rasmekomen and A.K. Parlikad. Condition-based maintenance of multi-component systems with degradation state-rate interactions. *Reliability Engineering and System Safety*, 148:1–10, 2016.

[204] Nipat Rasmekomen and Ajith Kumar Parlikad. Maintenance optimization for asset systems with dependent performance degradation. *IEEE Transactions on Reliability*, 62(2):362–367, 2013.

[205] Nipat Rasmekomen and Ajith Kumar Parlikad. Condition-based maintenance of multi-component systems with degradation state-rate interactions. *Reliability Engineering & System Safety*, 148:1–10, 2016.

[206] Leonardo R Rodrigues, João PP Gomes, Felipe AS Ferri, Ivo P Medeiros, Roberto KH Galvão, and Cairo L Nascimento Júnior. Use of phm information and system architecture for optimized aircraft maintenance planning. *IEEE Systems Journal*, 9(4):1197–1207, 2015.

[207] Michael J Roemer, EO Nwadiogbu, and G Bloor. Development of diagnostic and prognostic technologies for aerospace health management applications. In *Aerospace Conference, 2001, IEEE Proceedings.*, volume 6, pages 3139–3147. IEEE, 2001.

[208] S. Ross. *Stochastic Processes*. Wiley Series in Probability and Statistics. John Wiley and Sons, Inc., 1996.

[209] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[210] Stefania Russo, Roy Assaf, and Samia Nefti-Meziani. Towards a practical implementation of eit-based sensors using artificial neural networks. In *Sensors, 2017 IEEE Conference on*. IEEE, 2017.

[211] Abdulrahman S Sait and Yahya I Sharaf-Eldeen. A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics. *Rotating Machinery, Structural Health Monitoring, Shock and Vibration, Volume 5*, pages 307–324, 2011.

[212] M Sander and HA Richard. Experimental and numerical investigations on the influence of the loading direction on the fatigue crack growth. *International Journal of Fatigue*, 28(5-6):583–591, 2006.

[213] VVS Sarma, KV Kunhikrishnan, and K Ramchand. A decision theory model for health monitoring of aeroengines. *Journal of Aircraft*, 16(3): 222–224, 1979.

[214] L Satish. Short-time fourier and wavelet transforms for fault detection in power transformers during impulse tests. *IEE Proceedings-Science, Measurement and Technology*, 145(2):77–84, 1998.

[215] T Satow and S Osaki. Optimal replacement policies for a two-unit system with shock damage interaction. *Computers & Mathematics with Applications*, 46(7):1129–1138, 2003.

[216] Abhinav Saxena, Jose Celaya, Edward Balaban, Kai Goebel, Bhaskar Saha, Sankalita Saha, and Mark Schwabacher. Metrics for evaluating performance of prognostic techniques. In *Prognostics and Health Management, PHM 2008. International Conference on*, pages 1–17. IEEE, 2008.

[217] Abhinav Saxena, Jose Celaya, Bhaskar Saha, Sankalita Saha, and Kai Goebel. On applying the prognostic performance metrics. In *Annual Conference of the Prognostics and Health Management Society*, 2009.

[218] P. Scarf and M. Deara. Block replacement policies for a two-component system with failure dependence. *Naval Research Logistics*, 50:70–87, 2003.

[219] PA Scarf and M Deara. On the development and application of maintenance policies for a two-component system with failure dependence. *IMA Journal of Management Mathematics*, 9(2):91–107, 1998.

[220] Andrea Schirru, Simone Pampuri, and Giuseppe De Nicolao. Particle filtering of hidden gamma processes for robust predictive maintenance in semiconductor manufacturing. In *Automation Science and Engineering (CASE), 2010 IEEE Conference on*, pages 51–56. IEEE, 2010.

[221] S Senthilvelan and R Gnanamoorthy. Effect of rotational speed on the performance of unreinforced and glass fiber reinforced nylon 6 spur gears. *Materials & design*, 28(3):765–772, 2007.

[222] Mahmood Shafiee, Maxim Finkelstein, and Christophe Bérenguer. An opportunistic condition-based maintenance policy for offshore wind turbine blades subjected to degradation and environmental shocks. *Reliability Engineering & System Safety*, 142:463–471, 2015.

[223] Xiao-Sheng Si, Wenbin Wang, Chang-Hua Hu, and Dong-Hua Zhou. Remaining useful life estimation–a review on the statistical data driven approaches. *European journal of Operational Research*, 213(1):1–14, 2011.

[224] Xiao-Sheng Si, Wenbin Wang, Chang-Hua Hu, Dong-Hua Zhou, and Michael G Pecht. Remaining useful life estimation based on a nonlinear diffusion degradation process. *IEEE Transactions on Reliability*, 61(1):50–67, 2012.

[225] JZ Sikorska, Melinda Hodkiewicz, and Lin Ma. Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing*, 25(5):1803–1836, 2011.

[226] IEEE Reliability society. International conference on prognostics and health management. http://ieeexplore.ieee.org/xpl/conhome.jsp?punumber=1002538, 2008.

[227] PHM society. Conference on prognostics and health management, intenational journal on prognostics and health managment. https://www.phmsociety.org/, 2009.

[228] Sanling Song, David W Coit, and Qianmei Feng. Reliability for systems of degrading components with distinct component shock sets. *Reliability Engineering & System Safety*, 132:115–124, 2014.

[229] Joseph Sottile, Frederick C Trutt, and Aleck W Leedy. Condition monitoring of brushless three-phase synchronous generators with stator winding or rotor circuit deterioration. *IEEE Transactions on Industry Applications*, 42(5):1209–1215, 2006.

[230] Ashok N Srivastava and Jiawei Han. *Machine learning and knowledge discovery for engineering systems health management*. CRC Press, 2011.

[231] Jason R Stack, Ronald G Harley, and Thomas G Habetler. An amplitude modulation detector for fault diagnosis in rolling element bearings. *IEEE Transactions on Industrial Electronics*, 51(5):1097–1102, 2004.

[232] Hua Su and Kil To Chong. Induction machine condition monitoring using neural network modeling. *IEEE Transactions on Industrial Electronics*, 54 (1):241–249, 2007.

[233] Bo Sun, Shengkui Zeng, Rui Kang, and Michael G Pecht. Benefits and challenges of system prognostics. *IEEE Transactions on Reliability*, 61(2): 323–335, 2012.

[234] Hailiang Sun, Zhengjia He, Yanyang Zi, Jing Yuan, Xiaodong Wang, Jinglong Chen, and Shuilong He. Multiwavelet transform and its applications in mechanical fault diagnosis–a review. *Mechanical Systems and Signal Processing*, 43(1-2):1–24, 2014.

[235] N Tandon and A Choudhury. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology international*, 32(8):469–480, 1999.

[236] Liang Tang, Jonathan DeCastro, Greg Kacprzynski, Kai Goebel, and George Vachtsevanos. Filtering and prediction techniques for model-based prognosis and uncertainty management. In *Prognostics and Health Management Conference, PHM 2010.*, pages 1–10. IEEE, 2010.

[237] L. Thomas. A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. *Reliability Engineering*, 16 (4):297–309, 1986.

[238] LC Thomas. A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. *Reliability Engineering*, 16 (4):297–309, 1986.

[239] Wieger Willem Tiddens, Anne Johannes Jan Braaksma, and Tiedo Tinga. The adoption of prognostic technologies in maintenance decision making: a multiple case study. *Procedia CIRP*, 38:171–176, 2015.

[240] Tiedo Tinga and Richard Loendersloot. Aligning phm, shm and cbm by understanding the physical system failure behaviour. In *European Conference on the Prognostics and Health Management Society*, 2014.

[241] Diego Alejandro Tobon-Mejia, Kamal Medjaher, Noureddine Zerhouni, and Gerard Tripot. A data-driven failure prognostics method based on mixture of gaussians hidden markov models. *IEEE Transactions on Reliability*, 61 (2):491–503, 2012.

[242] Sheng-Tsaing Tseng and Chien-Yu Peng. Optimal burn-in policy by using an integrated wiener process. *IIE Transactions*, 36(12):1161–1170, 2004.

[243] Sheng-Tsaing Tseng, Jen Tang, and In-Hong Ku. Determination of burn-in parameters and residual life for highly reliable products. *Naval Research Logistics (NRL)*, 50(1):1–14, 2003.

[244] Serdar Uckun, Kai Goebel, and Peter JF Lucas. Standardizing research methods for prognostics. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pages 1–10. IEEE, 2008.

[245] George J Vachtsevanos, Frank Lewis, Andrew Hess, and Biqing Wu. *Intelligent fault diagnosis and prognosis for engineering systems*. Wiley Online Library, 2006.

[246] F.A. van der Duyn Schouten and S.G. Vanneste. Analysis and computation of(n, n)-strategies for maintenance of a two component system. *European Journal of Operational Research*, 48:260–274, 1990.

[247] Franciscus LJ van der Linden. Gear test rig for health monitoring and quasi static-and dynamic testing; design, construction and first results. In *International Gear Conference 2014 Conference Proceedings Volume II*, pages 976–985. Woodhead Publishing, 2014.

[248] JM Van Noortwijk. A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1):2–21, 2009.

[249] Nikhil Vichare, Peter Rodgers, Valerie Eveloy, and Michael G Pecht. In situ temperature measurement of a notebook computer-a case study in health and usage monitoring of electronics. *IEEE Transactions on Device and Materials Reliability*, 4(4):658–663, 2004.

[250] Pieter-Jan Vlok, Maciej Wnek, and Maciej Zygmunt. Utilising statistical residual life estimates of bearings to quantify the influence of preventive maintenance actions. *Mechanical Systems and Signal Processing*, 18(4): 833–847, 2004.

[251] Gregory W Vogl, Brian A Weiss, and M Alkan Donmez. *Standards related to prognostics and health management (PHM) for manufacturing.* US Department of Commerce, National Institute of Standards and Technology, 2014.

[252] H.-C. Vu, P. Do, A. Barros, and C. Berenguer. Maintenance grouping strategy for multi-component systems with dynamic contexts. *Reliability Engineering and System Safety*, 132:233–249, 2014.

[253] S Vulli, JF Dunne, R Potenza, D Richardson, and P King. Time-frequency analysis of single-point engine-block vibration measurements for multiple excitation-event identification. *Journal of Sound and Vibration*, 321(3-5): 1129–1143, 2009.

[254] Changting Wang and Robert X Gao. Wavelet transform with spectral post-processing for enhanced feature extraction [machine condition monitoring]. *IEEE Transactions on Instrumentation and Measurement*, 52(4): 1296–1301, 2003.

[255] Dong Wang, Kwok-Leung Tsui, and Qiang Miao. Prognostics and health management: A review of vibration based bearing and gear health indicators. *IEEE Access*, 2017.

[256] H. Wang. A survey of maintenance policies of deteriorating systems. *European journal of operational research*, 139(3):469–489, 2002.

[257] Haixia Wang, Jay Lee, Takahiro Ueda, Kondo H Adjallah, and Masoud Ghaffari. Engine health assessment and prediction using the group method of data handling and the method of match matrix: Autoregressive moving

average. In *ASME Turbo Expo 2007: Power for Land, Sea, and Air*, pages 697–702. American Society of Mechanical Engineers, 2007.

[258] Hongzhou Wang. A survey of maintenance policies of deteriorating systems. *European journal of Operational Research*, 139(3):469–489, 2002.

[259] Jianjun Wang, Runfang Li, and Xianghe Peng. Survey of nonlinear vibration of gear transmission systems. *Applied Mechanics Reviews*, 56(3): 309–329, 2003.

[260] Tianyi Wang. *Trajectory similarity based prediction for remaining useful life estimation*. University of Cincinnati, 2010.

[261] Wenbin Wang. A two-stage prognosis model in condition based maintenance. *European Journal of Operational Research*, 182(3):1177–1187, 2007.

[262] Xizhao Wang, Yadong Wang, and Lijuan Wang. Improving fuzzy c-means clustering based on feature-weight learning. *Pattern recognition letters*, 25 (10):1123–1132, 2004.

[263] R.E Wildeman, R. Dekker, and A.C.J.M. Smit. A dynamic policy for grouping maintenance activities. *European Journal of Operational Research*, 99: 530–551, 1997.

[264] S. Wu and J. M. Zuo. Linear and nonlinear preventive maintenance models. *IEEE Transactions On Reliability*, 59(1):242–249, 2010.

[265] Shaomin Wu and Philip Scarf. Two new stochastic models of the failure process of a series system. *European Journal of Operational Research*, 257 (3):763–772, 2017.

[266] Sitao Wu and Tommy WS Chow. Induction machine fault detection using som-based rbf neural networks. *IEEE Transactions on Industrial Electronics*, 51(1):183–194, 2004.

[267] Yinjiao Xing, Eden WM Ma, Kwok-Leung Tsui, and Michael Pecht. An ensemble model for predicting the remaining useful performance of lithium-ion batteries. *Microelectronics Reliability*, 53(6):811–820, 2013.

[268] Rui Xu and Donald Wunsch. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678, 2005.

[269] Ayako Yajima, Hui Wang, Robert Y Liang, and Homero Castaneda. A clustering based method to evaluate soil corrosivity for pipeline external integrity management. *International Journal of Pressure Vessels and Piping*, 126:37–47, 2015.

[270] Weizhong Yan, Hai Qiu, and Naresh Iyer. Feature extraction for bearing prognostics and health management (phm)-a survey (preprint). Technical report, Air Force Research Laboratory Wright-Patterson AFB Ohio Materials and Manufacturing Directorate, 2008.

[271] Ka Yee Yeung, Chris Fraley, Alejandro Murua, Adrian E. Raftery, and Walter L. Ruzzo. Model-based clustering and data transformations for gene expression data. *Bioinformatics*, 17(10):977–987, 2001.

[272] CT Yiakopoulos, Konstantinos C Gryllias, and Ioannis A Antoniadis. Rolling element bearing fault detection in industrial environments based on a k-means clustering approach. *Expert Systems with Applications*, 38 (3):2888–2911, 2011.

[273] Ozgur Yilmaz and Scott Rickard. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on signal processing*, 52(7): 1830–1847, 2004.

[274] Shen Yin and Zenghui Huang. Performance monitoring for vehicle suspension system via fuzzy positivistic c-means clustering based on accelerometer measurements. *IEEE/ASME Transactions On Mechatronics*, 20(5):2613–2620, 2015.

[275] Shen Yin, Steven X Ding, Xiaochen Xie, and Hao Luo. A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 61(11):6418–6428, 2014.

[276] Shen Yin, Steven X Ding, and Donghua Zhou. Diagnosis and prognosis for complicated industrial systems—part i. *IEEE Transactions on Industrial Electronics*, 63(4):2501–2505, 2016.

[277] Shen Yin, Steven X Ding, and Donghua Zhou. Diagnosis and prognosis for complicated industrial systems—part ii. *IEEE Transactions on Industrial Electronics*, 63(5):3201–3204, 2016.

[278] Shen Yin, Steven X Ding, and Donghua Zhou. Diagnosis and prognosis for complicated industrial systems—part 2. *IEEE Transactions on Industrial Electronics*, 63(5):3201–3204, 2016.

[279] Dejie Yu, Junsheng Cheng, and Yu Yang. Application of emd method and hilbert spectrum to the fault diagnosis of roller bearings. *Mechanical Systems and Signal Processing*, 19(2):259–270, 2005.

[280] Jianbo Yu. Bearing performance degradation assessment using locality preserving projections and gaussian mixture models. *Mechanical Systems and Signal Processing*, 25(7):2573–2588, 2011.

[281] Jianbo Yu. Machine tool condition monitoring based on an adaptive gaussian mixture model. *Journal of Manufacturing Science and Engineering*, 134(3):031004, 2012.

[282] Lotfi A Zadeh. Fuzzy sets. In *Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh*, pages 394–432. World Scientific, 1996.

[283] Jafar Zarei and Javad Poshtan. Bearing fault detection using wavelet packet transform of induction motor stator current. *Tribology International*, 40(5): 763–769, 2007.

[284] Bin Zhang, Chris Sconyers, Romano Patrick, and George Vachtsevanos. A multi-fault modeling approach for fault diagnosis and failure prognosis of engineering systems. In *Annual conference of the prognostics and health management society*, 2009.

[285] Guoqiang Zhang, B Eddy Patuwo, and Michael Y Hu. Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, 14(1):35–62, 1998.

[286] Nan Zhang, Mitra Fouladirad, and Anne Barros. Warranty analysis of a two-component system with type i stochastic dependence. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 2017.

[287] Wei Zhang, Chuanhao Li, Gaoliang Peng, Yuanhang Chen, and Zhujun Zhang. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mechanical Systems and Signal Processing*, 100:439–453, 2018.

[288] Xiaohong Zhang and Jianchao Zeng. Joint optimization of condition-based opportunistic maintenance and spare parts provisioning policy in multiunit systems. *European Journal of Operational Research*, 262(2):479–498, 2017.

[289] Fuqiong Zhao, Zhigang Tian, Xihui Liang, and Mingjiang Xie. An integrated prognostics method for failure time prediction of gears subject to the surface wear failure mode. *IEEE Transactions on Reliability*, 2018.

[290] Rui Zhao, Ruqiang Yan, Zhenghua Chen, Kezhi Mao, Peng Wang, and Robert X Gao. Deep learning and its applications to machine health monitoring: A survey. *arXiv preprint arXiv:1612.07640*, 2016.

[291] ZM Zhong, J Chen, P Zhong, and JB Wu. Application of the blind source separation method to feature extraction of machine sound signals. *The International Journal of Advanced Manufacturing Technology*, 28(9-10):855–862, 2006.

[292] Wei Zhou, Thomas G Habetler, and Ronald G Harley. Bearing condition monitoring methods for electric machines: A general review. In *Diagnostics for Electric Machines, Power Electronics and Drives, 2007. SDEMPED 2007. IEEE International Symposium on*, pages 3–6. IEEE, 2007.

[293] Enrico Zio and Giovanni Peloni. Particle filtering prognostic estimation of the remaining useful life of nonlinear components. *Reliability Engineering & System Safety*, 96(3):403–409, 2011.