Improving the Profitability, Availability and Condition Monitoring of FPSO Terminals

Samer Samir A. A. Gowid

A Doctoral Thesis submitted in partial fulfilment of the requirements for the award of

Doctor of Philosophy of Loughborough University

School of Electronic, Electrical and Systems Engineering
Loughborough University
2016

"This page is intentionally left blank."

Abstract

The main focus of this study is to improve the profitability, availability and condition monitoring of Liquefied Natural Gas (LNG) Floating Production Storage and Offloading platforms (FPSOs). Propane pre-cooled, mixed refrigerant (C3MR) liquefaction is the key process in the production of LNG on FPSOs. LNG liquefaction system equipment has the highest failure rates among the other FPSO equipment, and thus the highest maintenance cost. Improvements in the profitability, availability and condition monitoring were made in two ways: firstly, by making recommendations for the use of redundancy in order to improve system reliability (and hence availability); and secondly, by developing an effective condition-monitoring algorithm that can be used as part of a condition-based maintenance system.

C3MR liquefaction system reliability modelling was undertaken using the time-dependent Markov approach. Four different system options were studied, with varying degrees of redundancy. The results of the reliability analysis indicated that the introduction of a standby liquefaction system could be the best option for liquefaction plants in terms of reliability, availability and profitability; this is because the annual profits of medium-sized FPSOs (3MTPA) were estimated to increase by approximately US$296 million, rising from about US$1,190 million to US$1,485.98 million, if redundancy were implemented. The cost-benefit analysis results were based on the average LNG prices (US$500/ton) in 2013 and 2014.

Typically, centrifugal turbines, compressors and blowers are the main items of equipment in LNG liquefaction plants. Because centrifugal equipment tops the FPSO equipment failure list, a Condition Monitoring (CM) system for such equipment was proposed and

tested to reduce maintenance and shutdown costs, and also to reduce flaring. The proposed CM system was based on a novel FFT-based segmentation, feature selection and fault identification algorithm.

A 20 HP industrial air compressor system with a rotational speed of 15,650 RPM was utilised to experimentally emulate five different typical centrifugal equipment machine conditions in the laboratory; this involved training and testing the proposed algorithm with a total of 105 datasets. The fault diagnosis performance of the algorithm was compared with other methods, namely standard FFT classifiers and Neural Network. A sensitivity analysis was performed in order to determine the effect of the time length and position of the signals on the diagnostic performance of the proposed fault identification algorithm. The algorithm was also checked for its ability to identify machine degradation using datasets for which the algorithm was not trained. Moreover, a characterisation table that prioritises the different fault detection techniques and signal features for the diagnosis of centrifugal equipment faults, was introduced to determine the best fault identification technique and signal feature.

The results suggested that the proposed automated feature selection and fault identification algorithm is effective and competitive as it yielded a fault identification performance of 100% in 3.5 seconds only in comparison to 57.2 seconds for NN. The sensitivity analysis showed that the algorithm is robust as its fault identification performance was affected by neither the time length nor the position of signals. The characterisation study demonstrated the effectiveness of the AE spectral feature technique over the fault identification techniques and signal features tested in the course of diagnosing centrifugal equipment faults. Moreover, the algorithm performed well in the identification of machine degradation.

In summary, the results of this study indicate that the proposed two-pronged approach has the potential to yield a highly reliable LNG liquefaction system with significantly improved availability and profitability profiles.

Acknowledgements

I would like to express my sincere appreciation and thanks to my advisors Professor Roger Dixon and Dr Saud Ghani; you have been tremendous mentors for me. I would like to thank you for encouraging and supporting me during the whole PhD journey. Your advice on my research as well as on my career has been priceless. Additionally, I would like to thank the control group members at Loughborough University for their continued support, valuable comments, and sincere advice.

I dedicate this dissertation to my mother, my wife and my sons (Youssef and Adam). Throughout this educational journey, I have had the loving and unconditional support of you. Words cannot express how grateful I am to my mother for all of the sacrifices you have made on my behalf. Your prayers for me are what has sustained me thus far. I would like to thank all of my family members, friends and colleagues who encouraged, supported, and motivated me to strive towards my goal. Finally, I would like to express my sincere appreciation to my wife who has supported me tirelessly, and particularly at the exact times when I needed it most. Without your support, encouragement, and patience, I may not have made it to the end. I really owe a significant debt of gratitude to you all.

Table of contents

List of figures

X

List of tables

List of symbols

λ_o/λ Basic failure rate of component (h^{-1})

λ_E Effective failure rate of component (h^{-1})

μ Repair rate of component (h^{-1})

r Percentage of failure rate that is not influenced by maintenance

α Aging factor of component

τ Length of maintenance time interval

τ_o Actual length of maintenance time interval

γ Switchover rate (h^{-1})

G   Functioning state of the system

C Total maintenance cost

C_{PM} Preventive maintenance cost

C_{CM} Corrective maintenance cost

C_f LNG flaring cost during shutdown

List of acronyms

AE	Acoustic Emission
ANFIS	Adaptive Neuro-Fuzzy Inference System
AWSA	Adaptive Wavelet Stripping Algorithm
AI	Artificial Intelligence
ANN	Artificial Neural Network
AR	Auto-Regressive analysis
BBN	Bayesian Belief Network
BoG	Boil-off Gas
CBM	Condition-Based Maintenance
CL	Confidence Level
C_M	Corrective Maintenance
CM	Condition Monitoring
CF	Crest Factor
DCP	Dynamic Cylinder Pressure
EDM	Electric Discharge Machine
E	Energy
FFT	Fast Fourier Transform
FTA	Fault Tree Analysis
FPSO	Floating Production, Storage and Offloading platform
FDI	Fuzzy Fault Detection and Isolation
FIS	Fuzzy Inference System
GA	Genetic Algorithm

IAS	Instantaneous Angular Speed
IGCC	Integrated Gasification and Combined Cycle
IGCC	Integrated Gasification and Combined Cycle
LNG	Liquefied Natural Gas
MC	Machine Condition
MCHE	Main Cryogenic Heat Exchanger
MA	Markov Analysis
FS2FI	FFT-based Segmentation, Feature Selection and Fault identification algorithm
MFS2FI	Modified FFT-based Segmentation, Feature Selection and Fault identification algorithm
MTTR	Mean Time To Repair
MRC	Mixed Refrigerant Cycle
MLPN	Multilayer Perceptron Network
NG	Natural Gas
Nd	Number of frequency divisions
T	Operational time
P_M	Preventive Maintenance
PCA	Principal Component Analysis
C3MR	Propane Precooled Mixed Refrigerant Cycle
RBD	Reliability Block Diagram
RMS	Root Mean Square
RBC	Rotating Biological Contactor
SPL	Sound Pressure Level

SVM Support Vector Mechanism

Chapter 1. Introduction

1.1 INTRODUCTION

In 2014, the International Energy Agency (IEA) predicted that, by 2040, international demand for energy would increase by 37%, and that demand for natural gas (NG) would be more than 50% higher than the 2014 levels [1]. Global gas demand was 3284 billion cubic meters (bcm) in 2010, and has increased by 2.7% per year over the last decade. Gas has a 21% share of the global primary energy mix, behind oil and coal. The United States was categorised by the IEA as the first consumer and second producer of NG [2]. The U.S. Energy Information Administration (EIA) has estimated that there are 2587 trillion cubic feet (Tcf) of technically recoverable NG in the U.S. This includes undiscovered, unproved, and unconventional natural gas that could significantly change the energy supply and use markets in the U.S. [3]. NG and nuclear power are expected to be the two most promising energy resources in the coming decade. However, the catastrophic earthquake and tsunami that happened in Japan in March, 2011 undoubtedly shook confidence in atomic energy utilisation, thus giving NG utilisation a bigger share of future energy demand. NG is the cleanest fossil fuel and possesses many advantages such as the emission of a high calorific value, and emitting lower levels of potentially harmful by-products.

In 2013, the EIA forecast that worldwide demand for natural gas would dramatically increase over the following 28 years. During this period, NG would overtake coal and become the world's second most widely used fuel after oil. Demand for NG is expected

to grow faster than for both coal and oil. It is predicted that natural gas in its liquid form (LNG) will play a greater role in the overall gas supply. With natural gas production being more than 50% cheaper than diesel on an energy-equivalent basis, the economic logic of converting trucks, buses, locomotives, ships and stationary engines to using cleaner- burning, lower-emission LNG is compelling. Moreover, with growing electricity demand and the retirement of 103 Gigawatts of existing capacity, 340 Gigawatts of new generating capacity is to be added in the AEO2013 Reference case between 2012 and 2040, as shown in Figure 1. Natural gas-fired plants are due to account for 63% of capacity additions from 2012 to 2040 in the Reference case, compared with 31% for renewables, 3% for coal and 3% for nuclear [4].
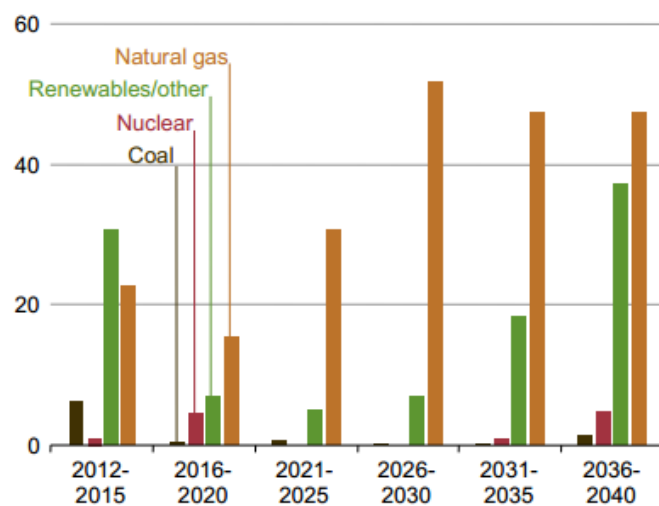
Figure 1: Addition of electricity generation capacity by fuel type [4].

LNG is used for transporting natural gas to markets, where it is regasified and distributed as pipeline natural gas. It can be used in natural gas vehicles, although it is more common to design vehicles so that they can use compressed natural gas. Its relatively high production cost, and the need to store it in expensive cryogenic tanks, have prevented its use in commercial applications from becoming widespread. LNG is produced by taking

natural gas from production fields, removing impurities, and liquefying it. Before liquefaction, the following preliminary processes take place: acid gas removal, dehydration, mercury removal, heavy hydrocarbon removal, and optional feed recompression. The liquefaction process properly starts with the treatment of dry lean natural gas. It is then followed by the three recognised basic steps of precooling the treated gas to about -30 to -40 °C, liquefaction to about -120 to -135 °C, and subcooling the LNG to about -140 to -165 °C, as shown in Figure 2 [5].

At stove burner tips, this condensed liquid form of NG takes up about 1/600th of its initial volume. The LNG is loaded onto double-hulled ships, which are selected for their safety and insulating purposes. Once the ship arrives at the receiving port, the LNG is typically off-loaded into well-insulated storage tanks. Regasification takes place in order to convert the LNG back into its gaseous form; it then enters domestic pipeline distribution systems, and is ultimately delivered to the end-user [6].
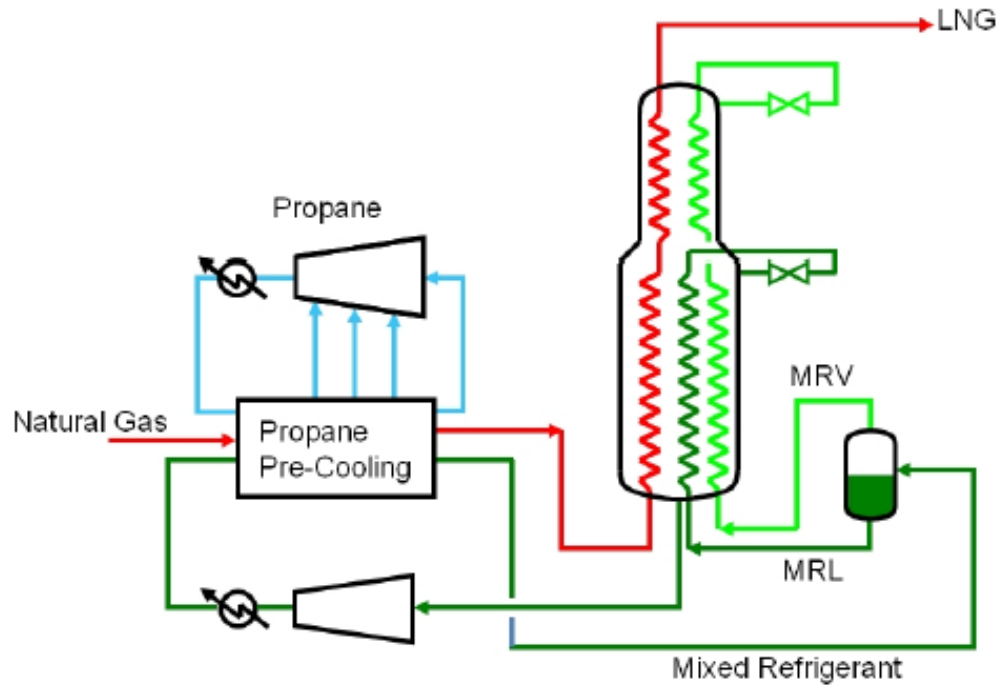
Figure 2: Air Products AP-C3MR™ LNG Process [5].

Due to the increasing demand for LNG and the need to reduce plant costs, the demand for LNG Floating Production, Storage and Offloading (FPSO) platforms has increased. FPSOs are floating plant which can produce and store LNG in the sea, and transfer it to carriers, and then onto the world market. The FPSO has the following benefits over onshore plants: significant cost savings of up to 40% due to the need for more limited investment in related infrastructure; potential further cost savings due to higher construction flexibility when an LNG FPSO is built in a shipyard and towed to site; shorter time-to-market; better flexibility in utilisation; favourable safety features; reasonably-sized less complex projects; early production; and the ability to operate robust cost-effective liquefaction processes [7].

Liquefaction is the key process which takes place on floating LNG platforms, and costs between 30% and 40% of their overall cost. Failures in LNG liquefaction systems may pose serious risks both to LNG ships and to the environment. To ensure optimal operational availability of offshore LNG liquefaction systems at the lowest possible overall cost (which includes capital, running costs, shutdown and maintenance costs), the following factors must be considered: (i) the configuration of the liquefaction system in order to obtain optimal redundancy; (ii) an effective Preventive Maintenance (PM) programme, and (iii) a repair strategy when the ship is at sea, which ensures that spare parts are available onboard [8].

There are three major liquefaction processes used on FPSOs, namely the Mixed Refrigerant Cycle (MRC), the N2 expander cycle, and the Propane Precooled Mixed Refrigerant C3MR process. The majority of FPSO manufacturers select the C3MR process for its high economic performance [9], and hence this process is the focus of this work.

Breakdowns in oil and gas production systems have a significant impact on the profitability of the business as expensive production equipment is left idle, and labour is no longer optimised. The ratio of fixed costs to product output is negatively affected. Quick repair of failed equipment is critical to business success.

However, when equipment breakdowns occur, costs can continue to be incurred well beyond the period of repair. Often, process lines require significant run time after start-up before producing a quality product. The goods being manufactured at the time of breakdown and for some time afterwards may either be unusable or lower in value. Due to negative impacts during and beyond the immediate downtime, businesses have sought

to prevent equipment breakdown by a process known as Condition-Based Maintenance (CBM). With CBM, equipment is continuously monitored in an effort to decrease repair times and prevent breakdown [10]. Unlike preventive maintenance, which is based on servicing machines at scheduled intervals, CBM is based on equipment conditions such as the operating environment and application [11].

It is vital to maintain equipment at the right time, and so in order to do this, CBM uses real-time data to prioritise and optimise maintenance resources. Observing the state of a system is known as "condition monitoring", which determines the equipment's health, and acts only when maintenance is necessary. The instrumentation of equipment has developed and become more extensive in recent years, and with the aid of better tools for analysing condition data, maintenance personnel are able to select the right time to perform maintenance on certain parts of equipment. Condition-based maintenance should minimise the cost of spare parts, system downtime, and maintenance tasks.

The specific advantage of condition monitoring is that potential degradation or failure can be detected; it enables operators to take maximum advantage of the useful life of components, such as bearings, as equipment can remain in service as long as its operational performance meets the desired performance standards. In general, the cost-effectiveness of condition monitoring, either by means of human surveillance or other condition monitoring techniques, should be evaluated against the following criteria [12]:

- The potential Health, Safety and Environment (HSE) consequences if the component/equipment is allowed to run to failure.
- The additional repair costs resulting from potential secondary damage if the component/equipment is allowed to run to failure.

- The expected longer useful life of the component/equipment relative to its scheduled replacement time.

- Efficiency gains in the execution of switching from main to standby systems, planned corrective tasks relative to unplanned ones, and possible additional economic consequences such as production loss.

1.2 PROBLEM IDENTIFICATION AND RESEARCH METHODOLOGY

Profitability is increased both by maximising income and minimising expenses. The profitability of FPSOs can be increased, firstly by improving the availability of plants in order to process as much LNG as possible, and secondly, by reducing operating and maintenance costs.

Hence, this study focuses on improving availability and reducing maintenance costs through the introduction of redundancy, alongside the development of an effective fault/condition monitoring system.

A further aim of this study is to help minimise the environmental pollution attributable to natural gas flaring during shutdowns. Production shutdowns may require the temporary flaring of all the gas stored or arriving at a facility, in order to reduce excessive pressure and avoid catastrophic incidents [13].

The profitability of LNG production plants is directly affected by the maintenance strategy applied and by the availability of the LNG liquefaction plant. LNG production plant availability is strongly related to system redundancy and also to the maintenance strategy applied. Hence, the introduction of redundancy and the choice of maintenance strategy are the two main factors that increase the reliability and operational availability of the C3MR liquefaction process [14] [8].

Figure 3 provides a simple top-level illustration of the general effect of changes to maintenance strategy on the total maintenance cost of the plant. Naturally, in real applications, the total maintenance cost is not as straight forward as this, depending on additional factors such as hazard rate, logistics, machine age, reliability, safety and integrity requirements etc.

In the figure, the total maintenance cost is equal to the sum of shutdown cost, Preventive Maintenance (P_M) cost, and Corrective Maintenance (C_M) cost. The shutdown cost is the cost associated with production stoppages due to maintenance or failure, and the switchover cost is the cost associated with production stoppages when production is switched from a main to a standby system. For large production plants, if the cost-benefits of introducing redundancy are justified, a standby redundancy strategy should be applied together with an effective CBM system. This to ensure that the total maintenance cost is kept to a minimum as the costs of shutdown, switchover and P_M will be reduced. Hence, this study sets out how the profitability of LNG plants can be improved by using the research methodology shown in Figure 4.

Standard maintenance strategy	→	Total maintenance cost= Shutdown cost + P_M cost + C_M cost
Standard maintenance strategy with optimised maintenance time intervals	→	Total maintenance cost= Shutdown cost + Optimised P_M cost + C_M cost
Condition-Based Maintenance (CBM) strategy	→	Total maintenance cost= Shutdown cost + C_M cost
Standby redundancy strategy	→	Total maintenance cost= Switchover cost + C_M cost
Standby redundancy strategy with an effective CBM system	→	Total maintenance cost= C_M cost

Figure 3: Effect on total maintenance cost of maintenance strategy applied.

Figure 4 presents the research methodology employed in this study. The research is divided into two substantive sections: (a) system reliability and (b) system maintenance. In the first section, the redundancy option is introduced, and its cost-effectiveness is investigated by using a newly-developed reliability model.

The second section sets out the development of a fast, automated, robust and easy-to-implement condition monitoring system for the implementation of a CBM strategy in

LNG liquefaction plants. It also characterises the major fault detection techniques and signal features for liquefaction plant rotating equipment with a view to combining fault diagnosis techniques and signal features. Such combinations will help to avoid misdiagnoses which can result in false alarms.
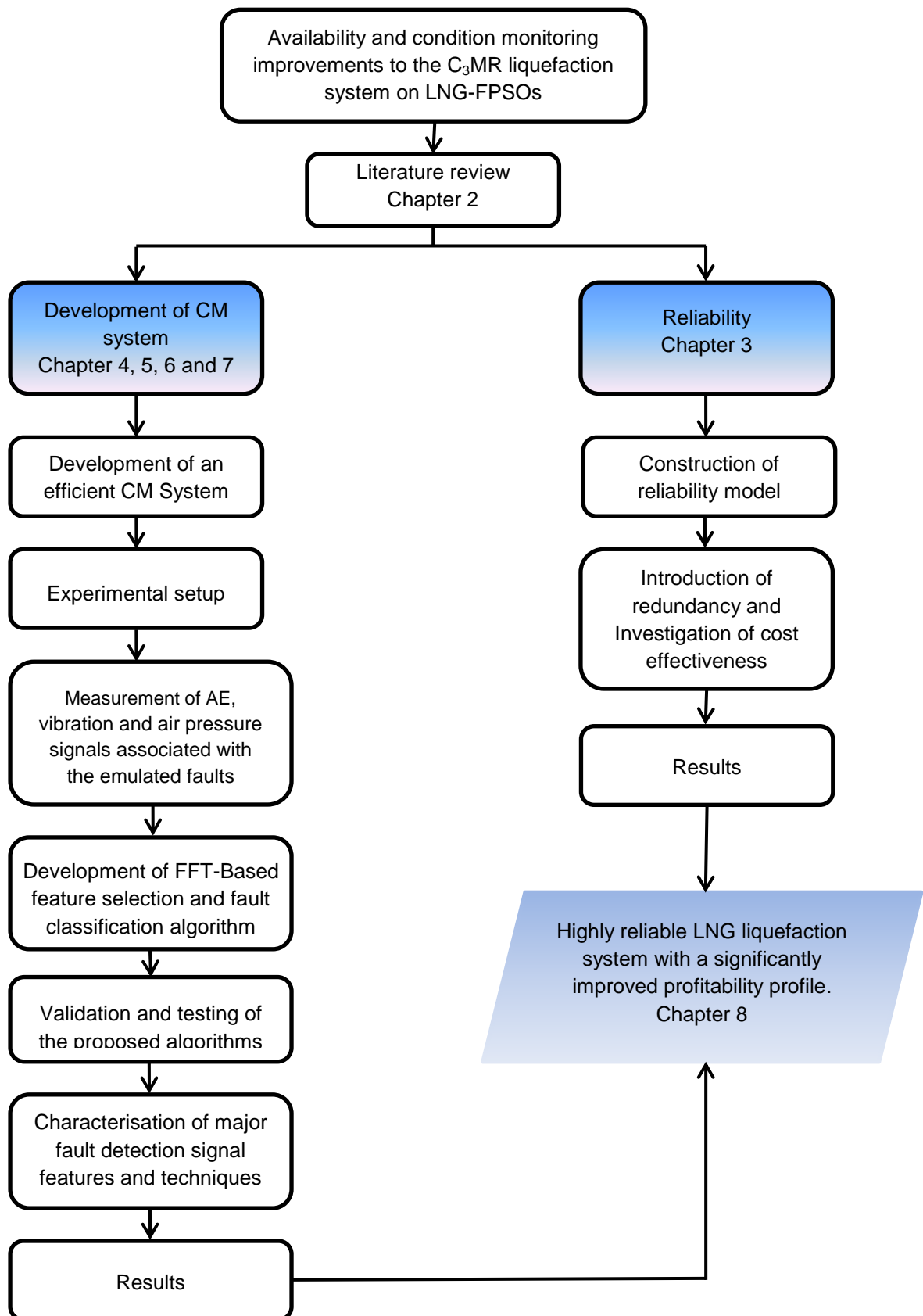
Figure 4: Research methodology.

1.3 NOVEL CONTRIBUTIONS AND PUBLICATIONS

The main aim of this study is to increase the profitability of LNG production plants. The study seeks to achieve this aims through a combination of reliability analysis and the development of new algorithms for fault/condition monitoring. Taken together, these approaches could result in improved availability with lower cost of maintenance. This study's main contributions are as follows:

1. A Markov model for the analysis of the reliability and operational availability of a typical FPSO LNG liquefaction system was developed. This model was also extended to consider redundancy options.

2. The estimation of the cost-benefit for partial and full redundancy of a typical FPSO liquefaction system.

3. An experimental compressor test-rig was built to simulate faults and generate data for condition monitoring studies.

4. An automated FFT-based segmentation, feature selection and fault identification algorithm for FFT-based CBM systems for typical high speed centrifugal equipment was proposed and demonstrated on experimental data. This algorithm proved to be robust, easy-to-implement, systematic, and highly responsive.

5. The algorithm was investigated for robustness using various signal time lengths and data window positions. The ability of the algorithm to identify machine degradation outside the datasets for which it had been trained was also investigated. A comparative study was performed to compare the performance of the proposed algorithm with other methods.

6. A "characterisation table" was developed to combine information from several fault detection techniques, namely AE, vibration, air pressure, Crest Factor, Energy

Factor, RMS, Amplitude, and spectral features. This approach was found to have great potential for the development of CBM systems for typical centrifugal equipment, and for the improvement of fault identification accuracy (when compared with a single technique).

Five indexed journal papers and one reviewed conference paper have arisen from this work. The list of publications is as follows:

Journal publications (5 papers)

S. Gowid, R. Dixon and S. Ghani, "Optimization of reliability and maintenance of liquefaction system on FLNG terminals using Markov modelling," International Journal of Quality & Reliability Management, Emerald, vol. 31, no. 3, pp. 293-310, 2014.

S. Gowid, R. Dixon and S. Ghani, "A Novel Robust Automated FFT-Based Segmentation and Feature selection Algorithm for Acoustic Condition monitoring Systems," Journal of Applied Acoustics, Elsevier , vol. 88, no. 1, pp. 66-74, 2015

S. Gowid, R. Dixon and S. Ghani, "Profitability, reliability and condition monitoring of LNG floating platforms: A review" Journal of Natural Gas Science & Engineering, Elsevier, vol 27, no. 3, pp. 1495-1511, 2015

S. Gowid, R. Dixon and S. Ghani, "Characterization of major fault detection features and techniques for the condition monitoring of high speed centrifugal equipment," Journal of Acoustics and Vibration, vol 27, no. 2, pp. 184-191, 2016.

S. Gowid, R. Dixon and S. Ghani, "Performance Comparison between FFT Based Segmentation Algorithm and Neural Network for the Condition monitoring of

Centrifugal Equipment" Journal of Dynamic Systems, Measurement and Control, ASME, vol139, no. 6, doi: 10.1115/1.4035458, 2017.

Conference publications (1 paper)

S. Gowid, R. Dixon and S. Ghani, "Assessment of liquefaction systems' process performance on LNG Floating export Terminals," International conference on Mechanical, Automotive and Aerospace Engineering (ICMAAE 2013), Paper # 30111, 2-4 July 2013, Kula Lumpur, Malaysia.

1.4 THESIS STRUCTURE

This thesis has eight chapters. Chapter 1 sets out the background, problem statement, objectives and contribution of the study, and presents the thesis structure. Chapter 2 is divided into six sections; the first five sections summarise the relevant existing research work, and the last section concludes with the outcomes of this literature review. Chapter 3 investigates the reliability of a typical C3MR LNG liquefaction system when several redundancy options are introduced; it then compares the new system with the standard system with a view to improving system availability and profitability. Chapter 4 explains the experimental setup which has been developed as part of this research project in order to evaluate the usefulness of the proposed maintenance methods. The acoustic emission transmission loss from the measurement system is quantified, and the calibration results and frequency responses are also presented. Chapter 5 proposes a novel, automated, fast, easy-to-implement and robust FFT Segmentation, Feature Selection and Fault Identification (FS2FI) algorithm which has its aim the improvement of the fault identification performance of CM systems. It also investigates and assesses the utilisation of AE spectral features for the diagnosis of faults in typical high speed centrifugal equipment. Chapter 6 studies the sensitivity of the proposed algorithm to changes in

signal time length and data window position. It also compares the performance of the proposed method with different fault identification methods. Chapter 7 presents the results of the experiments conducted in order to characterise the major fault detection techniques and signal features for the diagnosis of faults in typical high speed centrifugal equipment. Finally, Chapter 8 summarises the results of this research, identifies its limitations, and proposes possible future work.

Chapter 2. Literature review

2.1 INTRODUCTION

This chapter addresses the advantages of FPSOs, reviews the previous research related to LNG liquefaction processes, reliability and condition-based maintenance on FPSOs, and identifies a number of research directions with a view to improving the reliability and profitability of floating LNG terminals. The main purposes of this chapter are to ascertain the current state of the research, critically appraise the strengths and weaknesses of existing fault detection techniques, and identify research gaps and future work.

This chapter is divided into six sections. First, the objective and content of the chapter is introduced. Then in the second section, the economic performance of FPSO and LNG liquefaction processes are investigated and compared. Section three reviews the previous reliability studies of similar equipment. The most suitable reliability calculation approaches for LNG liquefaction systems are also investigated and prioritised according to their performance when solving reliability problems. Section four compares the benefits and cost-effectiveness of periodic preventive maintenance with CBM. In section five, the literature relating to the condition monitoring of rotating equipment is reviewed to identify the advantages and disadvantages of each fault identification and feature selection technique. Finally, the last section consists of a conclusion to the literature review, a presentation of the shortcomings of current research relating to reliability and maintenance system on FPSOs, and the research gaps that should be bridged to improve further the reliability and profitability of FPSOs.

2.2 FLOATING PRODUCTION, STORAGE AND OFFLOADING PLATFORM (FPSO)

The FPSO is a modern floating LNG production unit that can effectively and realistically exploit marginal and offshore gas fields in the event that (a) a pipeline network is not available, (b) there is a small number of wells, and/or (c) recoverable reserves are limited. FPSOs can produce and store LNG in the sea, and they have offloading facilities that enable the transfer of LNG product to LNG carriers and then on to world markets. The LNG-FPSO is not only compact and mobile, but can also be reused in other offshore fields [15]. Figure 5 shows a typical LNG-FPSO terminal.

The main disadvantages of current FPSOs are: (a) their low LNG production capacity (3.5 MTPA), (b) the poor reliability of their onboard centrifugal equipment, and (c) the inherent logistical difficulties of their operation and maintenance due to the remoteness of their offshore locations [16].
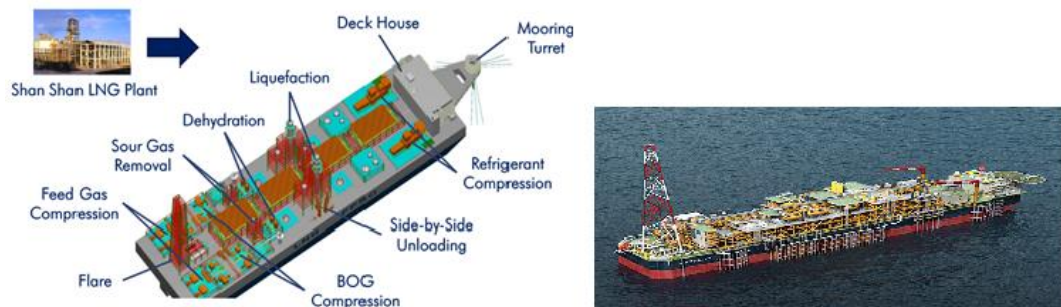
Figure 5: Sanha – World's first LNG FPSO Terminal [15].

2.2.1 BENEFITS OF LNG FPSOS

Yan and Yonglin, [15] and Haid [17] investigated and reported on the potential benefits of LNG FPSOs, and compared them to onshore LNG plants. They concluded that FPSOs are cost-effective and suitable for LNG production. Researchers have highlighted the extra benefits to the current LNG industry of using FPSO terminals in comparison with onshore LNG plants. In particular, cost savings and the operational effectiveness of

offshore terminals have been demonstrated. The following is a summary of the benefits of FPSOs [7] [17]:

- Cost savings of up to 40% compared with onshore LNG plants

- Shorter time to market

- High utilisation flexibility

- Compact in size

- Can be re-used in other offshore fields

- Early production

2.2.2 MAJOR LNG LIQUEFACTION SYSTEM PROCESSES

Liquefaction is a key process carried out on floating LNG platforms. The profitability of FPSOs is directly related to the availability of the liquefaction process. There are three major LNG liquefaction processes, namely the Propane Pre-Cooled Mixed Refrigerant cycle (C3MR), the Mixed Refrigerant cycle, (MR) and the Nitrogen Expander cycle (N2 expander). Despite the numerous advantages of the N2 expander system, the C3MR process remains the most utilised to date due to its highly economical process performance. However, further research should be undertaken to develop an LNG liquefaction process that combines the advantages of both C3MR and N2 liquefaction processes.

In a research study undertaken by Li and Ju [9], three major LNG liquefaction processes were described, analysed and systematically assessed. The study investigated the C3MR, MR and N2 expander processes used for the special gases associated with offshore production found in the South China Sea. These processes were analysed and compared by considering factors like performance parameters, economic performance, layout,

sensitivity to motion, suitability for different gas resources, safety, and operability. The study also considered the features of floating production, and the storage and offloading units for liquefied natural gas (LNG-FPSO) in the marine environment.

A typical pre-cooled mixed refrigerant process is shown in Figure 6. The process starts by lowering the raw natural gas pressure and temperature using a turboexpander (E). The low-pressure mixed refrigerants are compressed by a centrifugal compressor (P-1), and then precooled using a simple propane cooling system (P-2). The natural gas is then cooled by four huge series LNG heat exchangers (H-E1, H-E2, H-E3 and H-E4) to a very low temperature of approximately -165 degree Celsius. The liquefied form of natural gas is achieved when the output of the heat exchangers is throttled to the storage pressure. A booster/turboexpander set is employed in this system, and the work recovered from the turboexpander (E) is utilised to drive the booster (B). The pipelines are numbered according to the normal sequence of processes.

H-E1, H-E2, H-E3 and H-E4, LNG heat exchangers; S1, S2, S3 and S4, separators; M1 and M2, Mixers; P-1 and P-2, compressors; C-1 and C-2, after coolers; H, heater; B and E, booster/turbo expander; J1-J6, valves; P1, P2, Q1, Q2, E' and H', energy; numbers of 1-43, pipelines.
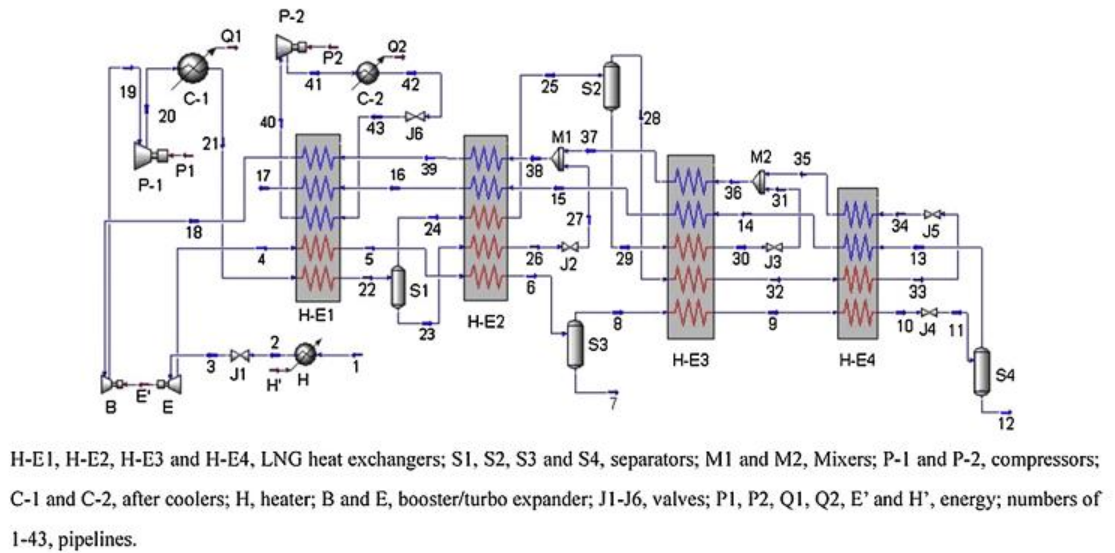
Figure 6: C3MR LNG liquefaction process [9].

The results indicate that the C3MR process has the highest economic performance, and that the Nitrogen expander process (N2 expander) has the highest energy consumption and poorest economic performance. Despite these two disadvantages of the N2 expander process, it remains more advantageous with respect to the C3MR and MR offshore application processes as it is easy to implement and more compact. The N2 expander requires less deck area, is less sensitive to LNG FPSO motion, is more suitable for other gas resources, is safer, and is easier to operate.

Figure 7 illustrates that, despite the various benefits of the N2 expander process, the C3MR and C3MR/split MR (precooled MR) were the most popular liquefaction process in 2013, accounting for 66% of the total capacity of in-service LNG trains [18].
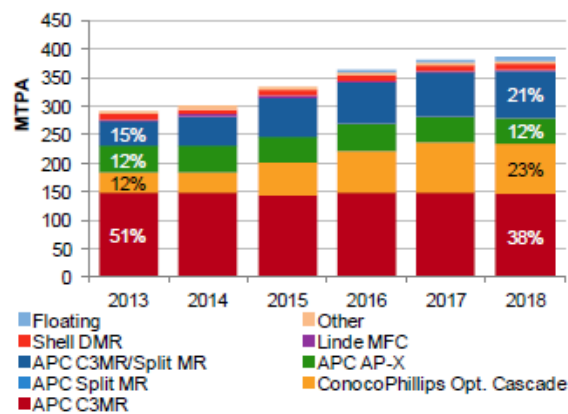
Figure 7: Liquefaction capacity by LNG process type [18].

According to the World LNG report, precooled MR processes such as C3MR and Dual Mixed Refrigerants (DMR) have the highest process efficiency, while the N2 process has the lowest. In terms of production capacity, the precooled MR process has the largest capacity, which normally ranges from about 1 to 5+ million tons per annum (MTPA). At the same time, the capacities of other liquefaction processes are limited to about 1 to 2 MTPA per train. The advantages of the N2 process is that it uses entirely non-flammable

refrigerant, and is insensitive to vessel motion; these features make it the safest LNG liquefaction process [19].

It can be concluded from the above research that various LNG liquefaction processes such as C3MR, MR, DMR, N2 expander and Cascade are currently utilised [9] [19]. The C3MR and DMR LNG liquefaction processes have the best economic performance, and most land-based LNG facilities use these processes in their LNG production lines.

2.2.3 SUMMARY

The research work reviewed in this section can be summarised as follows:

- The cost and operational effectiveness of FPSOs make them preferable to onshore plants; however, the poor reliability of their onboard centrifugal equipment represents their main disadvantage.
- The features that affect the selection of FPSO LNG liquefaction processes are: (a) process efficiency, (b) production capacity, (c) safety, and (d) impact of vessel motion.
- The N2 liquefaction process is the safest LNG liquefaction process as the N2 refrigerant is non-flammable, and the whole process is not sensitive to vessel motion.
- The C3MR liquefaction process has the greatest process efficiency, and most LNG onshore plants utilise this process in their LNG production lines.

2.3 RELIABILITY OF FPSO LIQUEFACTION SYSTEM

The reliability of LNG liquefaction plants is of paramount importance to the operation of FPSOs as it directly affects the profitability, availability and safety of FPSOs. Reliability can be defined as the probability of a component\system to perform their required

functions for a specific period of time without failure. On the other hand, availability can be defined as the probability that a system will not fail or undergo a repair action when it will be requested for use. Therefore, effective reliability models are utilized to calculate system availability which is a function of both system reliability and system maintainability [20] [21]. The failure rates for liquefaction systems on LNG floating platforms are high [22], and the repair times for failures are longer than for onshore repairs. All liquefaction system components are exposed to wear which results in increased failure rates over time if no maintenance is carried out [23]. The research papers summarised below (in Section 2.3.1) introduced reliability analysis as used in the oil and gas industry for a number of different types of terminals, and they discussed the different reliability analysis methods utilised to calculate the reliability of systems. This sets the context for the reliability work which is discussed later in this thesis, and also for the comparison of Markov-versus-Fault Tree methods undertaken in Section 2.3.2.

2.3.1 RELIABILITY IN OIL AND GAS INDUSTRY – STATE OF THE ART

Many authors have considered applying reliability analysis to the oil and gas industry [24] [25]. Researchers from academia and industry have also shown an interest in the area with papers addressing both the system level issues [26] [27] [28] [8] and the issues with particular components such as gas turbines [29], compressors [30], induction motors [31], pipelines [32], and bearings [33]. However, many of these studies are not directly relevant to this study.

Although there is a large amount of research which investigates reliability in different areas, a research gap can be observed when it comes to studying reliability and introducing redundancy to LNG liquefaction systems. Pil *et al.* [8] assessed the reliability

of Boil-off Gas (BoG) Systems on LNG carriers with a focus on maintenance strategies and redundancy optimisation. The objectives of their work were to: (i) consider and assess the feasibility, reliability and operational availability of the usual LNG re-liquefaction plant options for installation on a large LNG carrier; (ii) evaluate the financial benefits of both total and partial redundancies of the re-liquefaction plant, and suggest the preferred option for large LNG carriers; and (iii) offer a basic strategy for establishing a maintenance policy for LNG re-liquefaction plants.

Figure 8 is a diagram of a typical liquefaction and re-liquefaction plant. It shows the BoG (solid line) exiting the cargo tank, and entering the preparation system, where mist droplets are eliminated before compression. If necessary, the BoG is cooled, then compressed, and sent to the BoG liquefaction section, which in this case is a cold box (cryogenic heat exchanger) where the BoG is liquefied after indirect contact with cooled nitrogen gas. To minimise the investment cost, efforts were made to combine the LNG re-liquefaction plant with a Gas Combustion Unit (GCU) rather than introducing various redundancies into the LNG re-liquefaction plant. The results showed that, assuming no repair on board, the redundant system was more economical and efficient than the GCU.

This argument for choosing the redundant system was reinforced by the fact that the implementation of full onboard corrective maintenance is very likely to be impossible. Regarding the maintenance strategies, the results showed that both Preventive Maintenance (PM) and Corrective Maintenance (CM) costs were significantly higher than the cost of flaring. The reason for this was that the system availability was relatively high, even without introducing the redundancy into the N2 compressor system. This approach could be extended to study the reliability of existing LNG liquefaction systems.
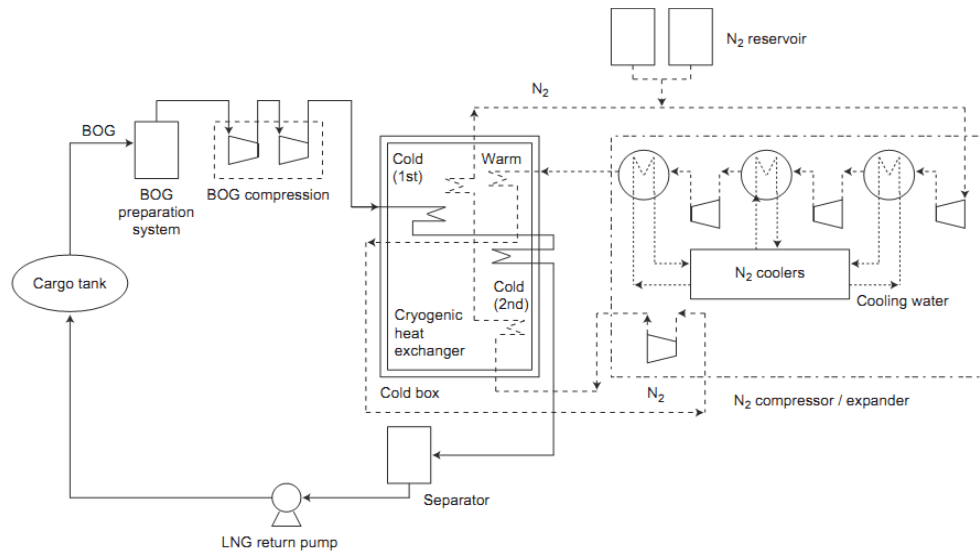
Figure 8: Diagram of a typical liquefaction and re-liquefaction plant [8].

Figure 9 shows a typical LNG FPSO. The length and weight of the FPSO shown in the figure are 310 m and 90,718 tones, respectively. A LM5000 LNG gas turbine driver has a length of 6.2 m and a weight of 12.5 tones [34]. While the length and weight of a typical 70M frame axial compressor are 5.8 m and 64.5 tones [35], respectively. Based on the dimensions of typical LNG drivers and compressors, the space required for a LNG liquefaction system will be about 25 m in length, 8 m in width and 3 m in height, excluding coldboxes. In comparison with the size and weight of FPSOs, the liquefaction system is considered to be small and the addition of a standby system will not present a problem in terms of space.

It can be observed from Figure 9 that there is a space on the ship deck for the installation of a new standby liquefaction system. However, in case of space constraint, the standby liquefaction system could be installed above the main system using a well-designed multi-story steel structure.

Figure 9: SENDJE-BERGE FPSO (348.75 m length, 51.87 m width and 274,333 tones deadweight) [36]

It can be concluded that the reliability and cost-effectiveness of the introduction of redundancy to LNG liquefaction systems have not yet been investigated and that there appears to be sufficient space to offer the potential of installing standby liquefaction systems/subsystems. Hence, a study should be carried out to identify the financial benefits of introducing redundancy to C3MR liquefaction systems on FPSOs.

2.3.2 RELIABILITY CALCULATION METHODS

Markov Analysis (MA) and Fault Tree Analysis (FTA) are two well-proven analytical techniques that are utilised in systems reliability calculation. Complexity of design and accuracy of results are the major parameters that should be taken into account before deciding on the most suitable analysis technique for the system [6].

MA is a technique developed by Andrei Markov to calculate the availability and reliability of systems with dependent components. The analysis is based on a state transition diagram that identifies all discrete system states as well as all possible transitions between these states. The MA model is time-dependent, and considers the transition rates which make this method a favoured option when calculating the reliability of time-dependent systems [37].

FTA is an analytical logic technique developed in Bell Telephone Laboratories to calculate the availability and reliability of systems. This analysis is based on a fault tree diagram which is built from the top down, and which takes discrete system state as the top event, and component states as the basic events. Logic gates (AND and OR) are then utilised to interconnect the events and conditions [38]. Although this technique does not consider the transition time from one state to another, its simplicity when modelling and calculating the reliability of a complex system makes it preferable to MA.

Andrews and Ericson [39] compared the accuracy of reliability calculation techniques. Table 1 shows that FTA and MA yielded the same results for series, parallel, and hot standby systems. The MA and FTA reliability equations for full monitoring, sequence parallel and cold/warm standby systems were different, making the results from FTA approximations. Although it was believed that MA produces more accurate results than FTA, this work asserted that FTA is accurate and that, for many design complexities, FTA produces a similar accuracy of results as MA. In addition, the authors observed that FTA is much easier for the modelling of large systems, for which it yielded acceptable results. Table 1 sets out a summary which compares the two approaches.

Table 1: Comparison between the accuracy of FTA and MA results.

Consideration	FTA	MA
Models Undesired Events	√	partially
Models Probability	√	√
Models Unavailability	√	√
Series System	√	√
Parallel system	√	√
Sequence Parallel System	Approx.	√
Full Monitor System	Approx.	√
Partial Monitor System	Approx.	√
Standby Redundancy System	Approx.	Difficult
Repair	√	√
Latency	√	√
Large models	Approx.	√
Dependencies	√	No
Coverage	√	√
Easy to follow model	√	No
Easy to document process	√	No

Norm [40] observed that fault trees and reliability block diagrams are widely utilised for predicting the safety of complex systems, and their maintainability and reliability; but they cannot accurately model the behaviour of dynamic systems. The author found that the MA technique is best for the analysis of dynamic systems, and attributed this to the ability of the MA technique to analyse complex, fault-tolerant, highly distributed and dynamic systems. The conclusions of Andrews and Ericson [39] and Norm [40] are consistent with each other, and showed that the effectiveness of both techniques depends on the size and type of system. The studies concluded that FTA is easier to use when modelling large systems, and that Markov gives a better accuracy when calculating the reliability of dynamic systems, such as sequential and cold/warm standby systems.

Ridley and Andrews [41] presented an improved model for the reliability calculation of standby dependencies and sequential systems using FTA and MA. In the case of MA, this achieved by introducing two new gates into the fault tree diagram. Both FTA and MA

were utilised to develop the new model. The authors extended FTA to produce accurate reliability calculations for both standby and sequential systems. The main disadvantage of FTA is that it only gives an approximation when used to calculate the reliability of dependent systems; while the main disadvantage of MA is the complexity of the construction of the Markov state transition diagram. The new model overcame the drawbacks of both FTA and MA by improving the accuracy of FTA, and by significantly reducing the complexity of MA. Pil *et al.* [8] modelled the BoG and N2 compressor systems using the Markov Approach. For each of these two systems, a redundant standby system of the same type was added. Because of the unique ability of MA to handle dynamic cases, it was used to calculate the reliability of the system. Cheng *et al.* [27] selected the FTA technique to calculate the reliability of the Emergency Shutdown System (ESD). The authors found that FTA is widely utilised for providing logical functional relationships between system components and subsystems, and for identifying the root causes of undesired system failures. Pil *et al.* and Cheng *et al.* utilised the MA and FTA approaches in modelling the BoG and ESD systems. Kwang *et al.* used MA when modelling the BoG system, taking into account the time-dependent transition rates (dynamic systems), while Cheng *et al.* utilised the FTA approach because of its proven effectiveness in modelling complex and big systems.

2.3.3 SUMMARY

The summary of the work reviewed in this section is as follows:-

- The introduction of a 100% standby system to the BoG liquefaction process significantly improved system reliability.

- The Markov Chain Approach is preferred over Fault Tree Analysis in calculating the reliability of time-dependent (dynamic) systems such as sequential and standby

redundancy systems; but the complexity of the construction of the Markov state transition diagram makes it difficult to implement on large systems. System availability is a function of system reliability and system maintainability.

- There are very few research papers addressing improvements in reliability of the LNG liquefaction systems, particularly for C3MR. Hence, the investigation of the system reliability of C3MR LNG liquefaction systems on FPSOs was identified as a research gap.

- The large size and weight of typical FPSOs mean that there is the potential to use redundant components in the liquefaction system.

2.4 PREVENTIVE MAINTENANCE VERSUS CONDITION-BASED MAINTENANCE

Maintenance cost is an important element in the overall cost of LNG production. Currently, four main maintenance strategies are being implemented in the oil and gas industry [42] [43], namely Corrective Maintenance (CM), Preventive Maintenance (PM), Risk-Based Maintenance (RBM) and Condition-Based Maintenance (CBM) or Predictive Maintenance. These maintenance strategies are described in [43] as follows:

Corrective Maintenance: "Maintenance is carried out following detection of an anomaly and aimed at restoring normal operating conditions. This approach is based on the firm belief that the costs sustained for downtime and repair in case of fault are lower than the investment required for a maintenance program" [43]. This strategy may be cost-effective except when frequent or catastrophic faults occur. The frequency and type of faults may significantly increase the shutdown cost of the plant and the repair cost for equipment.

Preventive Maintenance: "Maintenance carried out at predetermined intervals or according to prescribed criteria, aimed at reducing the failure risk or performance degradation of the equipment. The maintenance cycles are planned according to the need to take the device out of service and hence the incidence of operating faults is reduced" [43]. The implementation of this strategy has the potential for reducing the frequency of failures, and hence the overall shutdown cost. A cost optimisation should be carried out to minimise the total maintenance cost. Maintenance costs could significantly increase due to shutdown costs, the remaining life of spare parts, and labour charges in connection with preventive maintenance tasks undertaken during production stoppages.

Risk-Based Maintenance: "Maintenance carried out by integrating analysis, measurement and periodic test activities into standard preventive maintenance. The gathered information is viewed in the context of the environmental, operation and process condition of the equipment in the system. The aim is to perform the asset condition and risk assessment and define the appropriate maintenance program. All equipment displaying abnormal values is refurbished or replaced" [43]. This kind of maintenance programme has the potential to reduce risks and the number of catastrophic failures. On the other hand, it could increase the total cost of maintenance by requiring equipment to be overhauled or replaced on the basis of a risk assessment which does not explicitly consider the profitability of the plant.

Condition-Based Maintenance (CBM): "Maintenance based on the equipment performance monitoring and the control of the corrective actions taken as a result. The actual equipment condition is continuously assessed by the on-line detection of significant working device parameters and their automatic comparison with average values and performance. Maintenance is carried out when certain indicators give the

signals that the equipment is deteriorating and the failure probability is increasing. This strategy, in the long term, allows reducing drastically the costs associated with maintenance, thereby minimizing the occurrence of serious faults and optimizing the available economic resources management" [43]. Hence, effective CBM can offer the best maintenance strategies in terms of cost and risk management.

In summary, Preventive (also known as scheduled) Maintenance is a strategy that aims to reduce failures and equipment degradation through a number of planned maintenance tasks over the lifetime of the equipment at fixed time intervals. In contrast, Condition-Based Maintenance is a strategy that aims to predict and avoid failures through a system that monitors the current dynamic condition of the equipment as it changes over time, and then decides what maintenance action is required. Maintenance should only be performed when performance degradation is detected or a future failure is predicted.

Reduction in failure rates and their associated maintenance costs is the main objective of any maintenance strategy. Therefore, several models which optimise preventive maintenance frequency have been developed. These optimisation models are based on probabilistic techniques which develop solutions based on static probabilistic information. Probabilistic systems cannot provide an optimised maintenance schedule which performs better than stochastic and dynamic maintenance systems such as CBM. This can be attributed to the fact that CBM systems simultaneously optimise maintenance in accordance with dynamic changes in the status of equipment [44] [45] [46]. Hence, CBM has the potential to dramatically reduce maintenance costs in comparison with a normal preventive maintenance strategy [46]. In recent years, a sustained effort to shift from preventive maintenance to CBM maintenance has been observed [44] [45] [46]. Hence, the latter approach will be examined further in the next section.

2.4.1 TYPICAL CBM SYSTEMS, VALUE POTENTIAL AND CHALLENGES

As identified above, and as distinct from the preventive maintenance strategy, condition-based maintenance helps to reduce or eliminate unnecessary repairs, and increase profit by reducing maintenance costs and prevent disastrous machine failures [47] [48]. CBM relies on the regular monitoring of the mechanical condition of different trains of actual equipment. With CBM, maintenance is carried out when failure is imminent (ideally before it occurs), and significant reductions in unexpected machine failures, maintenance costs, and repair downtime are possible. The overall effects of CBM are to reduce maintenance costs, and improve the operational availability of systems. Thus, efforts to shift from PM to CBM are continuing [49].

Effective fault detection and condition monitoring systems are key enablers of successful and effective condition-based maintenance. Their effectiveness depends on their accuracy when identifying the correlation between fault situations and signal features using different fault detection techniques and machine process information such as vibration, sound and acoustic emission. Many fault detection techniques have been proposed in this field. Each technique has its own merits and demerits. Fault diagnosis is still a challenging problem as numerous fault situations can possibly affect the accuracy of detection due to the improper selection of signal feature sets, or due to the existence of fault interference and noise.

Figure 10 shows part of a typical CBM system which utilises a condition monitoring system that continuously monitors the status of system components in order to identify faulty components and their locations, and acts only when a component is about to fail. The process starts by reading the signals coming from sensors through a data acquisition system. The signals received are then passed onto a feature extraction algorithm which

extracts the best features of the signals, and then passes them onto a fault diagnostic/decision-making algorithm. The decision-making algorithm is trained using a set of machine fault signatures which are collected through either signal-based or model-based techniques. After the decision-making algorithm has been trained, the algorithm compares the signal features of the current machine condition with the signal features of the fault signature. Then, a real-time decision (faulty or healthy) is made, based on the prevailing condition of system components. Effective fault diagnostic algorithms will also identify the faulty components and the fault type.
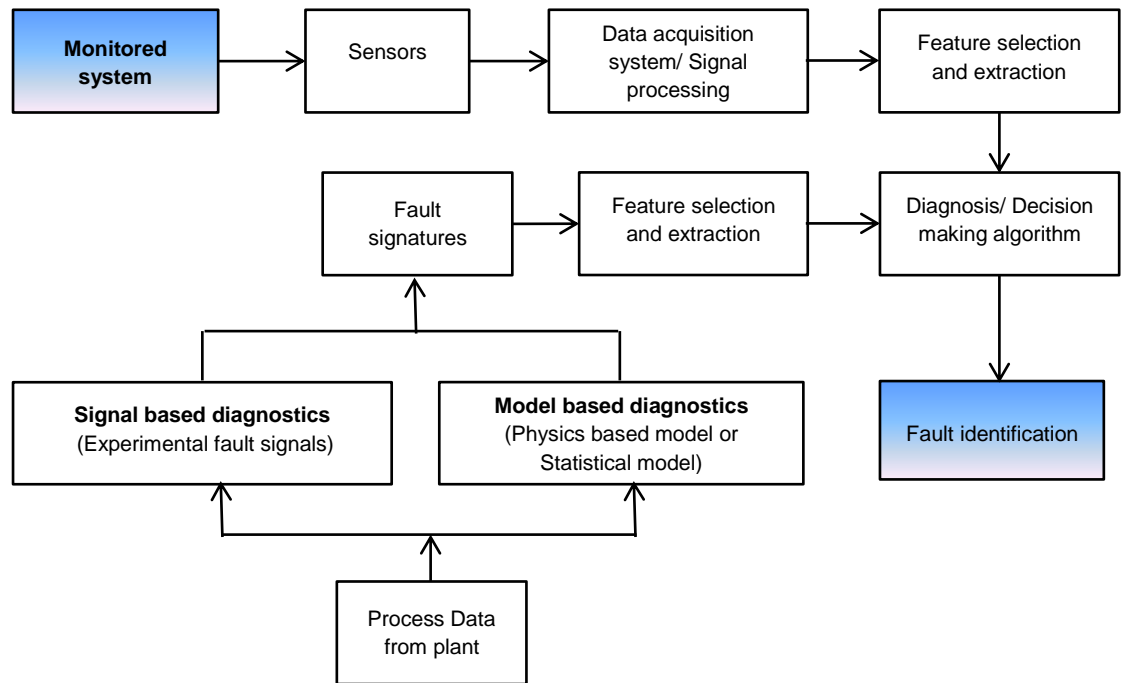
Figure 10: A typical CM and fault identification system.

Figure 11 illustrates the feature selection process using the signal-based diagnostic technique. The experimental determination of the best feature sets of different machine fault signatures consists of three major processes. The first process determines the most suitable fault detection technique (or combination of techniques) for the diagnosis of

machine faults; in the second process, the feature type that gives the greatest differences between fault patterns (time domain, frequency domain or time frequency domain) is selected; and the last phase is the processing and analysis of the signal in order to select the best set of features for the automated detection of machine condition.
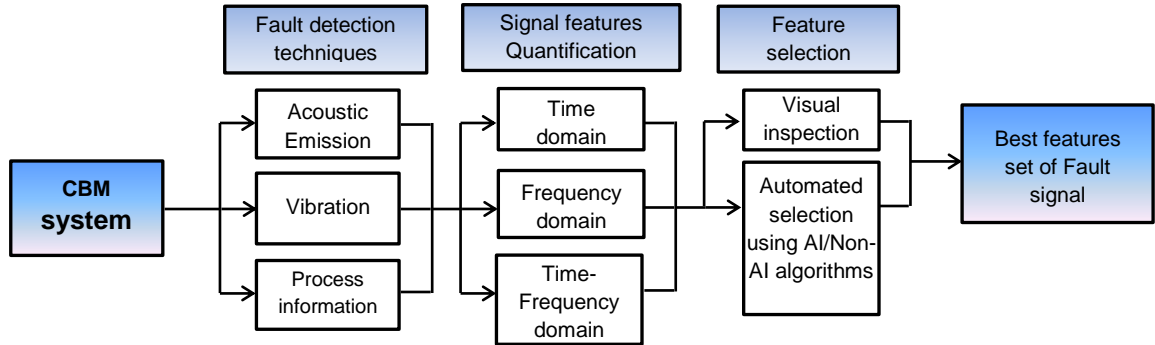
Figure 11: Selection process of best features set of a machine fault signal using the signal-based diagnostic technique.

2.4.2 SUMMARY

From the discussion above, it can be seen that condition-based maintenance offers a number of benefits over other approaches. Condition monitoring is crucial to the success of CBM. Hence the effectiveness of CBM systems depends on the proper selection of condition monitoring fault diagnosis and feature selection techniques. This will be considered next.

2.5 CONDITION MONITORING

According to [50], condition monitoring is a process that utilises the most appropriate fault detection technology and sensors to record a number of characteristics or equipment parameters. These characteristics and equipment parameters are then analysed to identify significant changes which are indicative of developing faults. Condition-Based Maintenance (CBM), as identified above, is a maintenance strategy that utilises the Condition Monitoring (CM) process to decide on the time and type of maintenance that

should be carried out. The decision is usually based on an analysis of various measurable data in relation to operating machines. CBM detects early changes in characteristics, parameters and trends, in order to identify the state of the deteriorating component prior to failure. This helps to enable rectification without experiencing the disadvantages which may result from loss of production due to unplanned shutdowns. In contrast, fault diagnosis processes identify faulty components along with the cause of the fault after the occurrence [50]. Of course, both elements are important.

Many authors have considered developing condition monitoring systems for rotating machinery. A number of research papers have assessed the suitability of implementing various condition-based maintenance techniques for fault detection in rotating equipment; the techniques in question are sound, vibration, process information and AE. The advantages of utilising multi-fault detection techniques have been investigated and discussed in other research articles, which are summarised in section 2.5.3. However, a lack of research has been observed in respect of the development of a "characterisation table" that prioritises the best fault detection and feature selection techniques for fault detection in rotating equipment. Automated feature selection process has a key role in CM systems, and comes next in sequence and importance to the selection of the most appropriate fault detection technique. Therefore, as summarised in section 2.5.4, many researchers have developed different approaches to the selection of the best sets of signal features in order to improve the detection of faults, and to decrease development costs and the duration of the feature selection process. However, the existing feature selection approaches are still in need of development in terms of accuracy and development time.

This section contains a survey of the recent techniques and results of CM systems. The section is organised into typical fault diagnosis techniques (section 2.5.1), the model-

based diagnostic techniques (section 2.5.2), the signal-based diagnostic techniques (section 2.5.3), and the feature selection techniques (section 2.5.4).

2.5.1 FAULT DIAGNOSIS TECHNIQUES

The diagnosis of faults is divided into three stages: detection, diagnosis, and prognosis. Faults are detected when a change in condition parameters is observed. Fault diagnosis consists of two main processes: fault identification and fault isolation. These processes determine the location of faulty components [50] [51] [52]. A fault can be diagnosed through a quantitative comparison between different machine condition patterns. The fault identification process provides information about the size of the fault and the time of onset; while the prognosis provides a long-term prediction for industrial applications. As the probability of a future failure event occurring is arguably stochastic in nature, the formulation of a prognosis is more complex than for diagnosis [53].

Diagnostic technique can be effectively assessed by considering the following factors: (a) detectability of fault, (b) the effect of noise on fault detection, and (c) the ability to easily distinguish a specific fault from other known and unknown faults. The major CM fault diagnosis techniques are divided into model-based and signal-based techniques. Section 2.5.2 reviews the previous research related to model-based fault detection techniques, while section 2.5.3 reviews the signal-based techniques.

2.5.2 MODEL BASED FAULT DIAGNOSIS TECHNIQUES

Models utilised for fault diagnosis can be categorised into physics-based models and statistical models.

2.5.2.1 PHYSICS-BASED MODELS

Physics-based models usually utilise mathematical models that are directly related to physics parameters that have direct or indirect effects on the health of system

components. The diagnostic process is based on the values of residuals which are commonly generated using Kalman Filters, Interacting Multiple Models, and Parity Relations [54].

Figure 12 shows a typical physics-based model for CM systems. The physics-based model approach aids the understanding of the physics of the system, and hence helps to develop an advanced model that considers system deterioration. However, the difficulties associated with this technique are that the model developed must be validated using an adequate number of actual datasets, and the model development process requires special knowledge of mathematics and theories relevant to the system monitored.
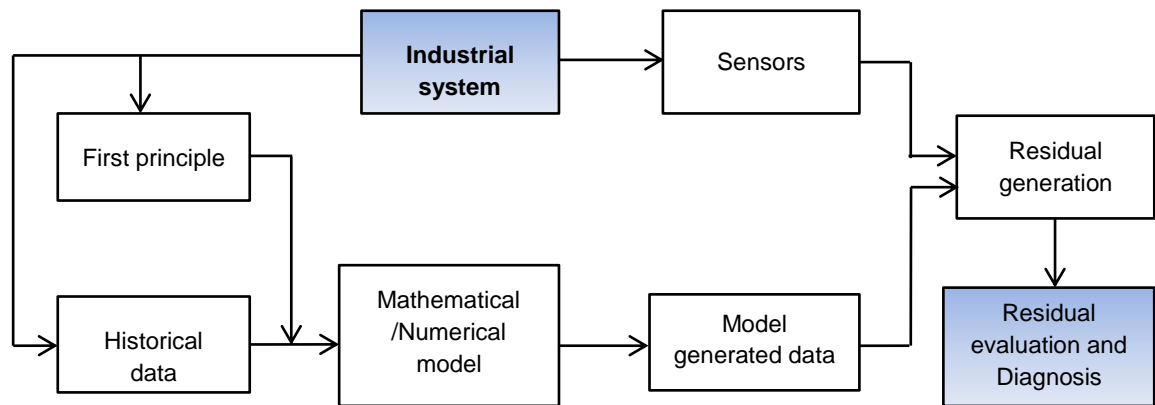
Figure 12: Flowchart of a typical physics-based model for CM systems [54].

Ginzinger et al [55] presented a model-based condition monitoring system for an auxiliary bearing. A multi-body simulation environment was utilised in the modelling of the rotor system, as shown in Figure 13. A number of fault simulation parameters were optimised in order to align the simulation results with the measurements. Two different faults were successfully identified as a result of the model developed.
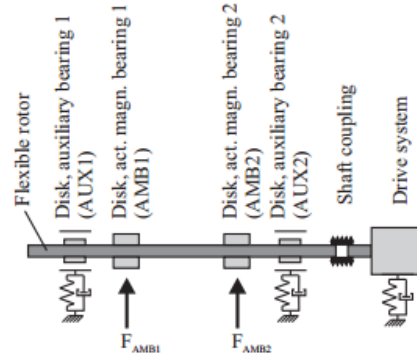
Figure 13: Modelling of auxiliary bearing dynamic response [55].

Charles *et al.* [56] developed two models to simulate wheel-rail profile and low adhesion contact. The results of the simulation were fed into a CM system which monitored the condition of the wheel-rail interface. Ugechi et al [57] proposed a model-based condition monitoring system for the diagnosis of faults in centrifugal pumps. The model was validated using vibration data from a centrifugal pump. Guo and Parker [58] developed a model-based condition monitoring system for planetary gear tooth wedging in a wind turbine system. The model shown in Figure 14 was utilised to predict vibration signals resulting from non-linear tooth wedging behaviour. The model also considered back-side contact, tooth separation, and bearing clearance. It can be observed that Guo and Parker verified the analytical model using a non-verified finite element solution. Further investigations should be carried out in order to apply the results of this study to an actual wind turbine system.
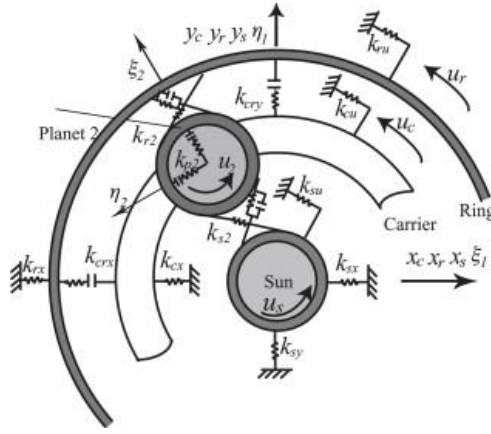
Figure 14: 2D dynamic model of planetary gear lumped parameter [58].

In practice, for all model-based approaches, modelling inaccuracies can and do occur due to simplifications and assumptions that are made when defining the models, and due to model parameter drift over time. Whilst there are some non-linear approaches, the majority of the models developed for monitoring are linear, and cannot handle non-linear systems. Linearisation of non-linear systems results in a significant reduction in effectiveness of the technique. Furthermore, the modelling of a full-scale process (such as in the case of FPSO compressors) can be difficult as it includes a number of non-linear relationships, which in turn increase the computational complexity and hence the likelihood of model error [59].

2.5.2.2 STATISTICAL MODELS

Statistical models are based on statistical time series measurements. In faulty conditions, statistical parameters, such as mean and standard deviation values, deviate from their benchmark values. Multivariate statistical techniques, such as Principal Component Analysis (PCA) and Partial Least Square (PLS), have proved their effectiveness in compressing data, and in handling correlation and noise in order to effectively extract

true information. The main function of these techniques is the transformation of a large number of process-related variables into a smaller set of uncorrelated variables [59].

Hundle et al [47] and Lampis et al [48] developed two CM approaches based on the available historical data. The first approach was based on FTA, while the second approach was based on Bayesian Belief Networks (BBNs). Both approaches utilised historical training data, and then performed the detection by comparing the current system performance to the historical data from sensors. The BBN approach proved more advantageous than FTA as it ranked all possible faulty components according to their failure probabilities, while FTA identified a single component only.

Although Hundle *et al.* and Lampis *et al.* demonstrated the effectiveness of FTA and BBNs, the fault diagnosis performance was not properly determined. Shang *et al.* [60] investigated the difficulties in implementing intelligent diagnostics in reciprocating compressors due to a lack of actual fault samples. Thus, the authors proposed the Support Vector Machine (SVM) technique, which relies on statistical learning theory in order to overcome the deficiency identified, and to provide a new approach to diagnostic technology. This approach was implemented in an intelligent diagnostic process which could accurately and rapidly recognise faults.

The main disadvantage of the work of Shang *et al.* is that the model was not verified due to a lack of actual fault signatures. The percentage difference between the features of machine condition signals in question was not determined. The quantification of these differences can be utilised when measuring the certainty/confidence level of CM approaches. Galka and Tabaswewski [53] utilised lifetime historical data from machine diagnostics. Symptom value fluctuations were utilised for machine learning. An energy

processor model was developed to verify the results of this study. It can be observed that Galka and Tabaswewski did not verify the model developed. The main advantage of having a model of the system is that it can allow numerous faults to be simulated, and this can contribute to the investigation of the issue of fault interference.

In general, statistical models are easier to use than explicit system models. Although statistical models are powerful in that they reveal the presence of abnormalities, the fact that they do not provide a fault signature for each fault makes the fault isolation process difficult. However, enhanced PCA showed some improvement to the differentiation of process conditions [61].

2.5.3 SIGNAL-BASED FAULT DIAGNOSIS TECHNIQUES

Sensory fault diagnosis systems provide real-time or continuous condition monitoring of rotating plant equipment for CM systems. Sensory inspections are considered the cornerstone of any dynamic maintenance system. These fault diagnosis systems utilise either statistical or experimental datasets for training and learning. They vary from simple devices such as vibration switches that can produce alarm signals and machine shutdown information, to highly sophisticated multi-channel monitoring devices that include time-to-frequency domain conversion, expert data analysis, and feature extraction algorithms. In order to detect various machine faults, and also to evaluate the severity of each fault, diagnostic systems detect and isolate faults, and this in turn ensures smooth and safe machine operation. The different signal-based fault detection techniques are investigated in section 2.5.3.1 and section 2.5.3.2.

2.5.3.1 SOUND, VIBRATION AND PROCESS INFORMATION-BASED TECHNIQUES

The previous research on CM using sound levels, mechanical vibration and process information is reviewed in this section. Toprak and Iftar [62] utilised sound pressure level

to diagnose machine faults. The authors studied five of the most common faults of compressors by means of a Multi-Layer Perceptron Network (MLPN). The MLPN was trained using the backpropagation algorithm. Training and testing data were obtained from the Sound Pressure Level (SPL) measurements of ten selected compressors. Recordings and measurements were carried out in a semi-anechoic sound test room with 12 microphones. Two different techniques were utilised for data analysis. In the first approach, the weighted average of the measurements of all 12 microphones was used. The second technique was based on the separation of data produced by the individual microphones; the results showed that the MLPN training required larger data files and more computational time when compared to the first approach. As shown in Table 2, each fault was precisely identified using the second approach, when sufficient training was performed.

Table 2: Summary of the experiments results [62].

	Analysis	# of Training Data Set	# of Test Data Set	# of Success	# of Failures
1st Approach	1	30	30	30	0
	2	60	6	6	0
	3	30	30	14	16
	4	54	6	3	3
	5	54	6	3	3
2nd Approach	6	30	30	20	10
	7	30	30	19	11
	8	54	6	6	0
	9	54	6	6	0
	10	54	6	6	0
	11	54	6	6	0

Toprak and Iftar's research stated that using the data of each microphone was more accurate than using the weighted average SPL of the data of the 12 microphones as long

as sufficient training was performed. The authors managed to demonstrate their proposed approach, but the high ratio between the training and testing datasets, the long computing time, and the high computing cost were the main disadvantages of their research.

Condition monitoring using vibration signatures is one of the major fault detection techniques. The main vibration signal analysis methods extract spectral and time domain features. Vibrations from machines usually result from the dynamic forces present in moving structures and parts. Different machine conditions can be detected by identifying their corresponding fault symptoms, for example mechanical vibration, and changes in process parameters such as temperature, efficiency and airborne noise [63].

Detections using vibration analysis show repetitive motion in the surfaces of rotating or oscillating machines. This repetitive motion may be caused by imbalances, misalignments, resonances, electrical effects, rolling element bearing faults, or any number of other causes.

To determine the current and future operating condition of the machine, it is vitally important to know its previous degradation pattern and history. The major vibration characteristics of rotating equipment are displacement, velocity acceleration, frequency, and phase angle [49]. In vibration spectra, "low" and "high" frequency ranges can be observed. The various types of vibration frequencies in a rotating machine are directly related to its geometry and operating speed. By knowing the relationship between the frequencies and the type of defect, vibration analysts can define the cause and severity of faults or problem conditions. The low vibration range contains component frequencies produced by rotational motion (harmonics), while the high vibration range contains component frequencies resulting from the interaction between the fluid-flow system and

the flow medium. In a power steam turbine, blade frequency range typically ranges from a few hundred hertz to about 10 to 20 kHz, depending on the turbine design [64] [65]. The amplitude of FFT spectra is the most common FFT feature used. A large number of recent papers have demonstrated the effectiveness of this feature in the detection of rotating equipment faults [66] [67] [68].

Shang e*t al.* [60] introduced a SVM-based intelligent diagnostic system for reciprocating compressors. This approach was utilised due to the lack of actual fault signatures of the different fault situations for compressors. The main disadvantage of this approach is that the differences in values between the features of the machine condition signals addressed were unknown. The quantification of the said differences can be utilised when determining the certainty/confidence level of CM approaches.

Wang and Hu [69] utilised the vibration technique to investigate the ambiguities and uncertainties that exist among pump failures and fault symptoms. A new approach for solving the existing problems of pump fault diagnosis was presented. Fuzzy logic was used to model the ambiguity and uncertainty relationship between different pump faults, analyse the fuzzy at different phases of fault diagnosis, and determine the frequency spectrum relevant to the pump faults in question. Analysis of the vibration signals of each pump was undertaken in order to extract the diagnostic features from the spectra. A fuzzy membership function, which was necessary for the pump fault diagnosis, was then created using condition variables based on dynamic signal processing.

Figure 15 shows two vibration spectra for the faulty device. The authors concluded that the faults in question were detected according to de-fuzzy diagnostic criteria and through a fuzzy comprehensive discrimination.

It can be concluded that it is difficult to adopt the traditional spectral vibration signature technique for the effective diagnosis of pump faults. The difficulty comes from the fact that the differences between various fault symptoms and events are uncertain. Thus, the authors established a new fuzzy membership function in order to address the interference problem. However, the accuracy of the proposed technique was not quantified, and it is not clear whether the proposed approach completely addressed the problem. Furthermore, the work did not investigate the problem of fault interference that strongly affects vibration signals.
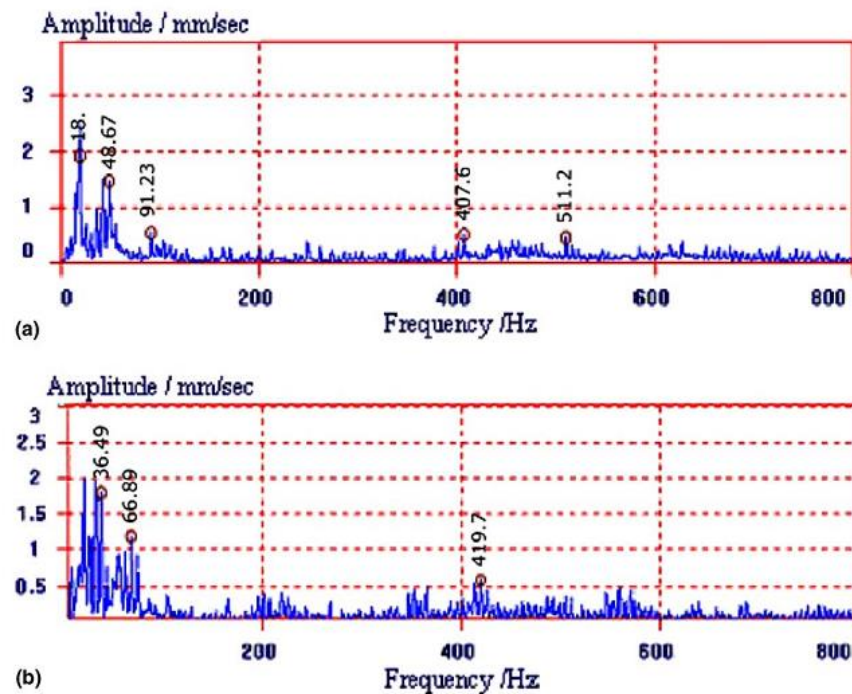
Figure 15: Two frequency spectra represent (a) sample fault, (b) second fault with the same sample fault on the second inlet valve [69].

Liao and Huang [70] observed that windowing the signals in Fourier Transform causes misrepresentation of vibration signals, and that frequency distribution spectra were not clear enough for shock vibration. Furthermore, it was difficult to extract good spectral

features because of the distribution of frequency sidebands in frequency modulation. Auto-Regressive analysis (AR) was utilised to analyse the signals in time domain. The changes in the AR coefficients were calculated by deducting the AR coefficients of faulty centrifugal compressor signals from those of healthy ones. A neural network was trained to model the relationship between the faults and the differences between the AR coefficients. The diagnosis results obtained from this neural network together with the differences between AR coefficients were better than the results of the neural network in conjunction with AR coefficients and distance approaches.

It can be concluded from this paper that time series analysis has some advantages over frequency domain analysis, and that the accuracy of a neural network, in conjunction with the differences between AR coefficients, is better than that of a neural network in conjunction with AR coefficients. The detection accuracy was not reported numerically, and no concurrent faults were emulated to investigate the issue of fault interference.

Condition maintenance using process information has demonstrated its effectiveness in diagnosing a number of machine faults. Fault detection performance is directly affected by the selection of proper information, as well as by the utilisation of an effective decision-making algorithm such as a decision table and fuzzy logic-based algorithms [71] [72]. Zanoli *et al.* [73] proposed a fault detection method for a compression process that was built into the Integrated Gasification and Combined Cycle (IGCC) part of a refinement plant. Single and multiple faults, which may have been capable of causing sensor reading errors in the process actuators, were considered. Principal Component Analysis (PCA) was used in a multivariable data-driven approach to monitor chemical process performance. The Analysis of Variance (ANOVA) procedure was proposed in

order to determine the number of principal components. Fault detection and isolation systems were tested and validated on the plant.

Elhaj *et al.* [72] studied the practical usage of two different CM techniques, namely, Dynamic Cylinder Pressure (DCP) and crankshaft Instantaneous Angular Speed (IAS). Conventional transducers for machine monitoring were utilised in their experiments. The authors proposed a monitoring technique for valve fault detection in reciprocating compressors. DCP and IAS were used to build two truth tables that showed the cases in which each method could be applied. The two truth tables were merged into one decision table. This combination provided a unique and reliable method for the detection and diagnosis of each individual fault in the compressor.

Zanoli *et al.* [73], Hafaifa *et al.* [71] and Elhaj *et al.* [72] utilised the process information technique to detect faults in machines. The authors addressed the selection of signal features and decision-making algorithm through different approaches, namely PCA, fuzzy logic and a decision table approach based on two truth tables.

The disadvantages of the research of Hafaifa *et al.* and Elhaj *et al.* are that the authors did not consider multiple faults, and hence they did not investigate the issue of fault interference. Elhaj *et al.* [72] did not determine the fault diagnosis performance of the proposed diagnostic approach. In addition, this approach was validated for use with compressor valves, but it is not known whether it will accurately diagnose main compressor faults such as bearing faults.

Condition maintenance using two or more combined techniques demonstrated a better performance in detecting machine faults [74] [75]. Schultheis *et al.* [75] studied different techniques used in machine health condition monitoring. The authors also compared

online against periodic monitoring, and proven against effective techniques. The following techniques were judged to be effective: ultrasound vibration, mechanical vibration, temperature, rod runout, and pressure velocity measurements. The measurements in respect of the crankshaft case and the crosshead piece of each cylinder were proven to be effective. For gas leaks, ultrasonic vibration measurement was preferred over mechanical vibration. Online monitoring was effective in decreasing the chances of catastrophic failures and consequential maintenance and shutdown costs.

It can be concluded that the combination of two or more monitoring techniques maximises the efficiency and accuracy of fault diagnosis in reciprocating compressors. Moreover, the appropriate fault detection technique must be selected according to the type of fault. Despite the promising results of combining fault detection techniques and signal features, only few researchers considered the characterisation of the major techniques and signal features for the CM of centrifugal equipment.

To summarise the work reviewed in this section:

- Utilisation of sound pressure, ultrasound vibration, mechanical vibration, and process information techniques such as temperature, rod runout, and pressure velocity techniques are effective in detecting rotating machinery faults.

- Vibration spectral (frequency domain) features are well proven for detecting faults in rotating equipment. FFT Amplitude is the most common spectral feature.

- The characterisation and combination of fault detection techniques and signal features have the potential to improve the accuracy of CM systems.

2.5.3.2 ACOUSTIC EMISSION-BASED TECHNIQUES

The majority of current CM systems utilise the vibration technique to detect faulty components, although most fault vibration signatures change with load and speed, and are affected by strong harmonics and interference. Thus, CM systems usually result in false indications of faults or inaccurate severity assessments for existing faults. The main concerns of most CM studies are the accuracy of the evaluation of machine condition, the identification of fault severity, and the prediction of the remaining life of machines under a broad range of operating states. The relationships that exist between machine conditions and machine process information, vibration and Acoustic Emission (AE) can be integrated into machine fault models during fault identification to assess fault severity.

AE is defined as the science that deals with the generation, transmission, reception, and effects of sound. It deals with demonstrable physical or airborne sounds that can manifest themselves as: signals relating to mechanical objects, pressure waves associated with leaking vapour or gases, or the humming of electrical equipment. Acoustic technology includes frequencies that can be as low as 2 Hz, or as high as the mega-Hertz range. Acoustic testing, which includes sensor selection, signal filtration and amplification, and low and high pass filters can be used to diagnose machine condition [49].

AE provides indicative data on levels of friction, rubbing, random impacting, and energy produced by the machine at the location of sensor. Ultrasonic monitoring is useful as a first line defence instrument because it collects information relatively fast and inexpensively. Ultrasonic monitoring can be used to detect the early onset of faults [76] [77]. Hence, ultrasound is utilised to perform preliminary diagnoses, and to alert operators to changes in machine condition. It should be noted that surface defects such as

cracks and scratches attenuate Rayleigh waves. Moreover, the surface finish of metals can also influence attenuation [78].

AE sensors cover a wide frequency range from 40 kHz to 1MHz. The time domain waveforms associated with AE are of two types: burst and continuous. A continuous AE signal refers to a waveform whose transient bursts are not differentiable. Both waveform types are associated with rotating machinery. For instance, a continuous emission may be the result of turbulent fluid flow within a pipe, while a burst emission could be associated with the transient rolling action of meshing gears [79].

With rotating machinery, the typical background operational noise is of the continuous type. The most commonly measured time domain AE parameters for diagnostics are amplitude, Root Mean Square (RMS), energy, kurtosis, Crest Factor (CF), counts and events. The FFT features provide useful information for rotating components since well-defined frequency components are associated with machine conditions [80]. The FFT-based feature selection process is key in CM systems as it directly affects the efficiency of the diagnostic process. Unlike the mechanical vibration technique, the AE technique is less affected by noise, and detects faults such as friction in bearings in their early stages.

High frequency AE signals are produced by rotating machinery due to frictional forces, and are often masked by low frequency vibrations and ambient plant noise [81]. AE RMS, maximum amplitude, and kurtosis values increase in line with defect size. However, observations of corresponding parameters from vibration measurements were disappointing [82].

Tandon and Nakra [83] investigated the suitability of AE peak amplitudes and the count method for the detection of outer race defects in bearings using a resonant type

transducer. The AE counts increased with rotational speed and load. However, although AE peak amplitudes provided an indication of defects irrespective of the defect size, AE counts did not provide any such indication when the defect was less than 250µm in diameter. The authors disagreed with Al-Ghamd and Mba [82] regarding the effect of defect size on the AE maximum amplitude.

The research of Al-Ghamd and Mba was more detailed: the authors emulated different crack sizes, and observed increases in AE amplitudes. Rogers [84] utilised the AE technique for monitoring the condition of slow rotating anti-friction slew bearings in cranes. The AE CM technique was found to give better results than the vibration technique. Grinding of the metal fragments in the bearing, rubbing of crack faces, and impacts between the damaged parts and the rolling elements in the loaded zone were all identified as sources of detectable AE signatures.

Schoess [85] presented the results of an assessment of six different relevant technologies for the onboard monitoring of railcar bearings. It was concluded that the AE technique offered the greatest potential advantage. Rogers and Schoess demonstrated the potential advantage of the AE-based condition-based maintenance technique for the detection of faults in bearings. However, Rogers focused only on the kurtosis parameter, and did not investigate the other time domain and frequency domain signal analysis methods.

Neil *et al.* [86] described how AE techniques could be implemented as a condition-based maintenance strategy to monitor the inlet and outlet valves of reciprocating compressors. AE sensors required very little space and were non-intrusive, which was a major benefit in hostile conditions. The results indicated that AE sensors could be practically deployed for condition monitoring applications.

Alfayez and Mba [87] presented a case study on the application of high-frequency acoustic emissions as a means of detecting the early stages of loss of mechanical integrity in low-speed Rotating Biological Contactors (RBCs). An RBC was used for sewage treatment in small communities, and rotated at approximately 1rpm. The stub shaft of the RBC was fractured. The potential of AE to diagnose serious mechanical defects was demonstrated, however the vibration technique was found to be ineffective.

Gill et al, Alfayez and Mba highlighted the effectiveness of AE-based fault detection techniques for the detection of both reciprocating compressor and mechanical integrity faults. The AE technique was found to be more informative than the vibration technique. The study did not investigate nor compare other time domain parameters.

Dane [88] discovered that ultrasonic flow measurement offered significant advantages over widely utilised turbine instruments. These instruments were known to be positively biased by at least 5% due to their inherently non-linear aerodynamics. The response of the AE signal to velocity fluctuations was not affected by the rotor inertia of turbine due to the fact that non-linear aerodynamics do not affect AE signals.

Puttmer [89] proved that AE sensors could play an important role in the condition monitoring of machinery. The author developed a CM system for reciprocating positive displacement pumps, and stated that the AE technique was found to have been better than the vibration technique, in particular in noisy environments.

Schulthesis *et al.* [75] showed that the preferred approach in valve condition analysis is ultrasonic. Ultrasonic energy is often associated with gas leaks, and so a leaking valve is a strong generator of ultrasonic energy. Ultrasonic measurements are usually taken in conjunction with compressor pressure-volume analysis.

Goodman [90] attributed the effectiveness of ultrasonic to the fact that most leakage problems and all operating equipment produce a broad range of sound. The high-frequency ultrasonic components of these sounds are are waves of an extremely short nature. These short wave signals are directional, and it is relatively easy to determine their exact location. This can be achieved by separating these signals from operating equipment and background plant noises. Moreover, the directional nature of ultrasound allows these potential warning signals to be detected at the onset of faults, and before they are detected by infrared or vibration techniques. Dane, Puttmer, Schulthesis *et al.* and Goodman demonstrated that AE-based CM systems are effective in detecting a number of machine faults. The authors demonstrated that the AE technique responds well to flow velocity fluctuations and gas leaks.

In practice, the generalisation capacity outside the training fault signature is considered to be a limitation of the signal-based technique [59]. Fault interference is one of the major disadvantages of the signal-based technique. When faults interfere, their signatures change, and sometimes faults are masked due to the interactions of different fault signals. This issue can be solved by avoiding taking decisions if there are no similar fault patterns in that region, or by increasing the number of samples/fault patterns in order to specify all possible faults. This process should explicitly include the combination and degradation of all fault conditions, though this is considered to be very difficult. However, in the event of dissimilar fault patterns and multiple faults, this technique will effectively detect abnormalities in operation, though with a limited ability to classify the fault.

It can be observed that there is a lack of research into the combination and characterisation of the major fault detection techniques and signal features for high speed centrifugal equipment.

To summarise the work reviewed in this section:

- Fault interference and noise represent the main obstacles when implementing CM systems.

- AE is more effective than the mechanical vibration technique in detecting faults such as friction in bearings at their early stages. Also, it is less affected by noise and by the non-linear aerodynamics of rotors.

- AE proved its effectiveness over the mechanical vibration technique in detecting the size of cracks.

- The most commonly measured AE diagnostic parameters are amplitude, RMS, Energy, kurtosis, crest factor, counts and events.

- The collective utilisation of several fault detection techniques and signal features has the potential to improve the accuracy of fault diagnosis [91] [92] [93]. However, there is a research gap when it comes to combining and characterising the major fault detection techniques for high speed centrifugal equipment.

2.5.4 FEATURE ANALYSIS AND FAULT CLASSIFICATION TECHNIQUES

Currently, the majority of fault diagnosis systems are based on two feature analysis and fault/pattern classification techniques. The first technique analyses time and frequency domains using traditional methods, while the second technique analyses the time and frequency domains using Artificial Intelligence (AI), which takes the neural network method as representative. The traditional methods have a reasonable performance when detecting faults, but prior knowledge and numerous fault samples are needed.

AI techniques also have a reasonable performance but need a long computational time and are costly [50]. The accuracy of the results is always dependent on various design parameters which should be set based on the training, validation and testing sets.

The global structure of the monitoring system which is generally used can be divided into the following three main sequential processes: data collection, data acquisition (which includes the calculation of statistical functions and values in both time and frequency domains), and finally, automated fault diagnosis. Automated fault diagnosis is the most difficult phase, and is still under development.

Fault diagnosis approaches should undergo continuous development to adapt to the necessities of industrial applications, and to avoid dependency on operators [94]. The frequency domain signal analysis technique is essential when using vibration or AE-based monitoring. For vibration-based monitoring, the magnitude of the vibration signal is utilised essentially for establishing the severity of the vibration, while the frequency content is utilised for identifying the cause or origin. The AE-based method is widely used for monitoring the condition of rotating machinery. Compared to traditional vibration-based methods, the high frequency approach of AE has the advantage of significantly improving the signal-to-noise ratio [95].

Since the Artificial Neural Network (ANN) architecture and connection weights (initial conditions) significantly affect the performance of the ANN network, it is desirable to identify the best possible set of ANN design parameters. As shown in Figure 16, the performance of the ANN technique is directly affected by the following major design parameters: (a) the type of training algorithm, in addition to the initial values of connection weights, (b) the number of training cycles (Epochs), and (c) the number of

hidden NN layers, as well as the number of neurons in each layer. Although one hidden layer is always sufficient to approximate any continuous function, the use of two hidden layers can improve the generalisation in complex problems [96].
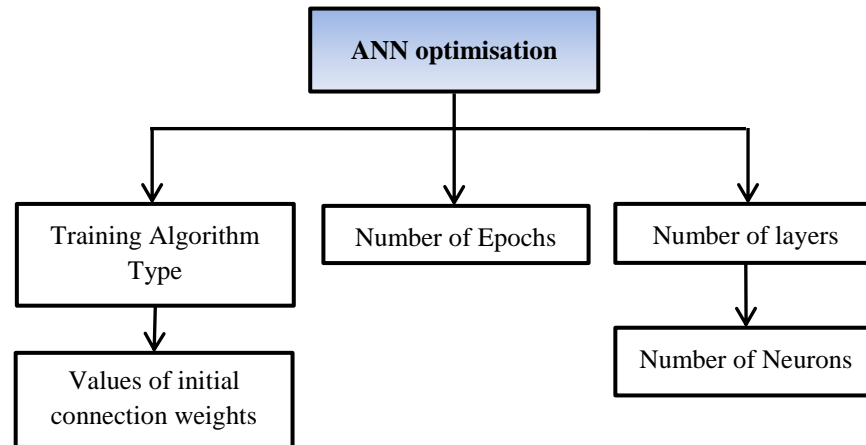
Figure 16: ANN optimisation design parameters.

Chan and Gu [97] investigated the accuracy of the Adaptive Neuro-Fuzzy Inference System (ANFIS) and ANN AI approaches, and observed that accuracy increased when the design parameters were optimised in terms of number of training epochs, number of Membership Functions (MF) of ANFIS per input, number of ANN neurons, and type of transfer function for ANN. The values of these design parameters were obtained from over 50 runs and with final fuzzy if-then rules of 81 for the ANFIS based turbine cycle model.

Saxena and Saad [80] proposed the utilisation of Genetic Algorithm (GA) with ANN for identifying a near-optimal feature set for ANN fault diagnostic systems. Health conditions were simulated using nine bearings, eight of which had different crack sizes and one of which was healthy. The cracks were constructed using an Electric Discharge

Machine (EDM). Three accelerometers and one AE sensor were utilised. Five feature options were set as inputs for the GA, namely statistical features, statistical parameter for the sum, the differences between signals, spectral features, and all the features together.

The FFT analysis was based on 32 values for each signal. The results showed that the technique of using GAs for selecting an optimal feature set for a classification application of ANNs was powerful, and that the collective use of all features was the best. The GA optimised the best combination based on the performance obtained directly from the success of the classifier; the mean classification success was 99.94%.

Chan and Gu's investigation and the Fuzzy Interference System (FIS) structure and parameter adjustment theory demonstrate that the ANFIS is complex as the number of membership functions, number of training cycles (epochs), and number of rules must be set. The accuracy of prediction mainly depends on the design parameters, which in turn usually depend on the training and testing datasets. Hence, every time the training datasets change, the design parameters should be adjusted to ensure the maximum accuracy. This consequently increases the complexity, computational time, and cost of AI-based approaches. The algorithm developed by Saxena and Saad included the FFT in the analysis, but it did not change the number of segments to better identify the fault. The number of segments should be automatically changed according to both the number of faults and the difference between the values, in order to optimise the detection of different faults. Moreover, the FIS algorithm was not tested for the diagnosis of simultaneous faults. The authors did not investigate the effectiveness of the proposed algorithm when selecting the best feature set for multiple-fault classifiers, and when investigating the issue of fault interference.

The Support Vector Machine (SVM) is an artificial intelligence method based on the principle of statistical learning theory. The SVM method was utilised for both feature selection and classification [98]. Meng and Feng [99] presented a new condition monitoring and analysis method for small sample studies on a reactor coolant pump using SVM. The data were passed through a multi-band FIR filter to eliminate noise and useless frequencies. Kernel principal component analysis was utilised to decrease the dimension of the vector, processing time and accuracy. This method was used as a multiple classifier, and was able to separate the different machine conditions successfully.

Gryllias et al. [91] developed an SVM-based feature selector for the selection of optimal features in the absence of actual experimental data. The input features were divided into two groups: (a) time domain statistical features such as RMS, SK, VAR and kurtosis, and (b) spectral features such as energy values calculated at the specific frequency bands of the demodulated and measured signals. The main contribution of this work resulted from basing the SVM training on a model that considered the dynamic behaviour of defective rolling element bearings. This enabled the SVM to select a set of good features without the need for experimental data in relation to defective bearings.

The approach developed by Meng and Feng was not validated for the detection of simultaneous faults (unbalance and friction faults). Neither did these authors investigate the effectiveness of the proposed technique in distinguishing multiple machine faults. Furthermore, Gryllias et al. did not consider the fault interference problem, and only studied the occurrence of a single fault. Finding the dynamic equation of each component was difficult and time-consuming. The performance of the proposed CM approach was

not verified experimentally. Moreover, this approach will not help to minimise the development cost and time needed for CM systems.

Samhouri et al. [92] proposed a new approach based on a combination of the axial vibration time signal features of the carnallite surge tank pump, namely RMS, variance, skewness, kurtosis, and normalised sixth central moment. These features were utilised as inputs for both *Adaptive* ANFIS and ANN. Three different faults with three different fault codes were simulated. A total of 92 runs were conducted, of which 73 were for training and 19 for testing. The comparison showed that the adoption of the time root mean square and variance features achieved the minimum fault prediction errors for both ANFIS and ANN. The trapezoidal membership function in ANFIS achieved a fault prediction accuracy of 95%, while the cascade forward back-propagation ANN achieved a better fault prediction accuracy of 99%.

Gupta and Wadhwani [93] proposed a robust Genetic Programming-based (GP) feature selector for the selection of the best features from a large feature dataset for bearing fault classification. The ANN classifier was utilised for the recognition of fault patterns. Vibration time domain features were extracted from the statistical measures of Median, RMS, crest factor, histogram Lower Bound (LB), histogram Upper Bound (UB), Entropy (ENT), Skewness (SK), Kurtosis (KT), Variance (VAR), Shape Factor (SHF), Impulse Factor (IMF), and Clearance Factor (CLF). Experimental data were collected for four bearing conditions namely health, defective outer race, defective inner race, and defective ball fault condition. The algorithm was utilised to effectively select a smaller subset of features. All eight features were selected by the GP and yielded a detection accuracy of 99.99%.

Zhao et al. [100] addressed the limitations of existing spectral feature selection algorithms when handling redundant features. Since redundant features can have significant adverse effects on learning performance, the authors proposed a novel spectral feature selection algorithm for an embedded model. The proposed algorithm evaluated the utility of a set of features after the efficient removal of redundant features. The algorithm was based on sparse multi-output regression. The algorithm yielded an average feature selection redundancy rate of 0.24. This rate is much lower than the redundancy rates of the existing spectral features algorithms.

Samhouri et al. observed the effectiveness of the ANN technique against the ANFIS technique. The authors neither addressed the fault types nor utilised the spectral analysis technique as one of the major vibration analysis techniques. As no multiple-fault simulations were carried out, the effectiveness of the proposed approach in distinguishing simultaneous faults was not demonstrated.

Zaho et al. demonstrated that the existing spectral feature selection algorithms such as Laplacian Score, Fisher score and trace ratio failed to handle the problem of redundant feature identification. It can be observed that the existing algorithms evaluate features individually, and cannot identify redundant features.

In recent years, several studies of bearing fault diagnosis using wavelet analysis have been conducted. Lin and Qu [101] used wavelet analysis, and varied the shape factor of the Morlet wavelet to achieve minimum wavelet entropy for bearing fault feature selection. Qiu et al. [102] used the Shannon entropy and singular value decomposition to optimise the wavelet entropy and kurtosis parameters. Bozchalooi and Liang [103]

introduced a smoothness index to guide the parameter selection of the complex Morlet wavelet for de-noising bearing fault signals.

Wang et al. [104] proposed a novel Adaptive Wavelet Stripping Algorithm (AWSA) to extract simulated transients from bearing fault signals. A comparison between the periodic multi-transient model and the AWSA was carried out to show that the proposed approach was better in selecting the random characteristics of real transients. An enhanced AWSA was also developed to reduce the computing time.

Shen et al. [105] proposed an automated sensory feature selection method to reduce the development time and cost of condition monitoring systems for machining operations. Force, acceleration, sound and acoustic emission sensors were utilised for the detection of high-speed milling operations. Time domain, frequency domain and wavelet analysis techniques were employed to analyse the signals measured. Gradual tool wear was used for evaluating the proposed self-learning automated sensory feature selection method. The results showed that the proposed method could be utilized in automated and self-learning monitoring system for the selection of the most suitable sensors. Lin, Qu, Wang et al. and Shen et al. proved the effectiveness of wavelet analysis technique for fault detection in bearings. It can be concluded that the majority of research has focused on the improvement of feature selection algorithms as well as on the minimisation of computing time and cost.

The complexity of the ANN technique comes from the fact that its performance is significantly affected by a large number of design parameters [96]. The development time and cost of the ANN-based CM techniques are significant due to the need for the optimisation of all design parameters, which in turn should be customised according to

the training datasets; in any event, 100% accuracy cannot be guaranteed. Many approaches such as SVM, AWSA and GA have been utilised to improve the accuracy of the ANN-based feature selection algorithm and to reduce its computing and development times.

2.5.5 SUMMARY

To summarise the work reviewed in this section:

- FFT analysis is a key approach in vibration and acoustic-based monitoring. It is judged to offer more potential for correct fault diagnosis than the model-based monitoring approaches summarised above.

- AE demonstrated a superior efficiency in detecting faults at their early stages, and in identifying crack sizes when compared with the mechanical vibration technique. Moreover, it was less affected by rotors, non-linear aerodynamics, and noise.

- Neural Networks or AI approaches are judged to have some potential for monitoring and feature selection. However, they are complex and can be computationally demanding.

- The majority of the existing automated feature selection and fault classification techniques utilise artificial intelligence methods such as ANN, ANFIS, GA and SVM, while other tools utilise standard fault classification algorithms.

- Combination of different types of signal features has some potential to improve the fault identification accuracy of CM systems.

- A research gap was identified in respect of the development of effective, automated, non-artificial intelligence, fast and non-complex FFT-based fault identification algorithms for CM systems.

2.6 CONCLUSION AND RESEARCH DIRECTIONS

Having reviewed a broad range of relevant literature, a number of conclusions can be drawn: It is clear that the importance of FPSOs has stimulated a number of researchers to research them, and to recognise the advantages of floating production over on-shore. It was also found that the C3MR is the most popular liquefaction system currently available (mainly due to economics). However, the relatively poor reliability of the rotating equipment on FPSOs is one of its weaknesses.

This weakness was highlighted in sSectionection 2.3 where it was found that there was a research gap in the area of introducing redundant systems or components to the C3MR liquefaction system in order to achieve full operational reliability of FPSOs. This section also suggested that Markov modelling should be the preferred approach for analysing the reliability and availability of the liquefaction system.

Section 2.4 identified dynamic maintenance, namely condition-based maintenance (CBM), as a useful approach for improving the operational reliability (and hence the availability) and reducing the maintenance costs of FPSOs. CBM is therefore preferred over preventive maintenance-based systems. The section went on to discuss the importance of effective condition monitoring methods which do not generate too many false alarms (thereby allowing CBM to take place).

Finally, Section 2.5 discussed in more detail some methods and algorithms for machine condition monitoring which included model-based, time signal-based, and frequency domain approaches. Approaches for the feature selection of particular faults were also considered. The major conclusions were that, for rotating equipment, sound, vibration

and process information such as pressure, temperature and other operating information are judged to be the most appropriate fault detection techniques.

Whilst the acoustic emission technique seems to have a better performance in respect of detecting faults in machines, it was concluded that a combination of two or more fault detection techniques should improve diagnostic accuracy, and reduce the potential for false alarms (Section 2.5.5).

With regard to fault diagnosis, artificial intelligence-based feature selection approaches were most common. However, these were considered complex, time-consuming and sensitive to a number of design parameters. Hence, it was felt there was some scope for further research in this area.

Overall, the outcomes of this literature review have set the context for this research. Therefore, with the top level aim of improving the profitability of FPSOs through reducing unavailability resulting from failures and preventive maintenance shutdowns, this study will investigate the introduction of redundancy to the C3MR liquefaction system. It will also propose and demonstrate a new CM system that (if it can be applied in practice) has the potential to reduce the maintenance costs of FPSOs. The required robustness and effectiveness of the proposed system can be achieved by developing a spectral feature selection and fault identification algorithm which is automated, fast and effective, and by characterising and suggesting combinations of the major fault detection techniques for the condition-based maintenance of high speed centrifugal equipment. The effectiveness of the proposed algorithm will be analysed by means of identifying changes in data, firstly, in terms of robustness and secondly, by comparison with a standard FFT-based scheme and a Neural Network-based scheme.

Chapter 3. Reliability of C3MR liquefaction system

3.1 INTRODUCTION

The liquefaction of gas on LNG-FPSOs is an essential process, yet one which experiences a high failure rate [22]. The profitability of an entire LNG production plant is strongly affected by the availability of the liquefaction system; the factors which are relevant to its availability are redundancy and maintenance strategy. The majority of FPSO manufacturers tend to utilise the C3MR cycle on account of its strong economic performance. As the liquefaction system has the highest system failure rate on FPSOs, the latter's availability is severely affected by the former's performance. Hence, this chapter will investigate the introduction of redundancy to C3MR liquefaction systems on FPSOs with a view to making potential improvements to the availability of FPSO units.

The C3MR liquefaction process and three other proposed options for introducing redundancy to its key operational components are described herein. A set of Markov models have been developed to facilitate the calculation of the reliability of the four options. This chapter seeks to contribute to the overall research by investigating the introduction of redundancy as a possible means of improving both reliability and profitability of LNG liquefaction plants.

3.2 DESCRIPTION OF THE C3MR LIQUEFACTION PROCESS

Figure 17 depicts a typical propane precooled mixed refrigerant cycle (C3MR). This cycle is currently the most common in the LNG industry, and it is used in over 80% of the world's completed production trains. The C3MR system utilises a multi-component refrigerant, usually nitrogen, propane, butane, methane, ethane, or pentane in order to evaporate and condense natural gas in a cycle over a wide range of temperatures. The propane refrigerant is first utilised to precool dry and treated gas to around $-30^{o}C$. The precooling step condenses any remaining heavy hydrocarbons and some LNG in a scrub column (separators). These liquids are separated into their components in a fractionation train, then used as refrigerant makeup, and are finally re-injected into the liquefaction feed gas (up to the gas specification limit), or sold as separate Natural Gas Liquid (NGL). The precooled feed gas is then sent to the Main Cryogenic Heat Exchanger (MCHE), where it is condensed and then subcooled at elevated pressures using mixed refrigerants. The subcooled Liquid Natural Gas (LNG) leaving the MCHE is then flashed to near storage tank pressure, cooling the LNG to $-163 \ ^{o}C$, and causing the ejection of nitrogen-rich steam [106].

Propane-precooled mixed-refrigerant process.
LP= Low Pressure; MP= Medium Pressure; HP= High Pressure
(Source: Air Products)
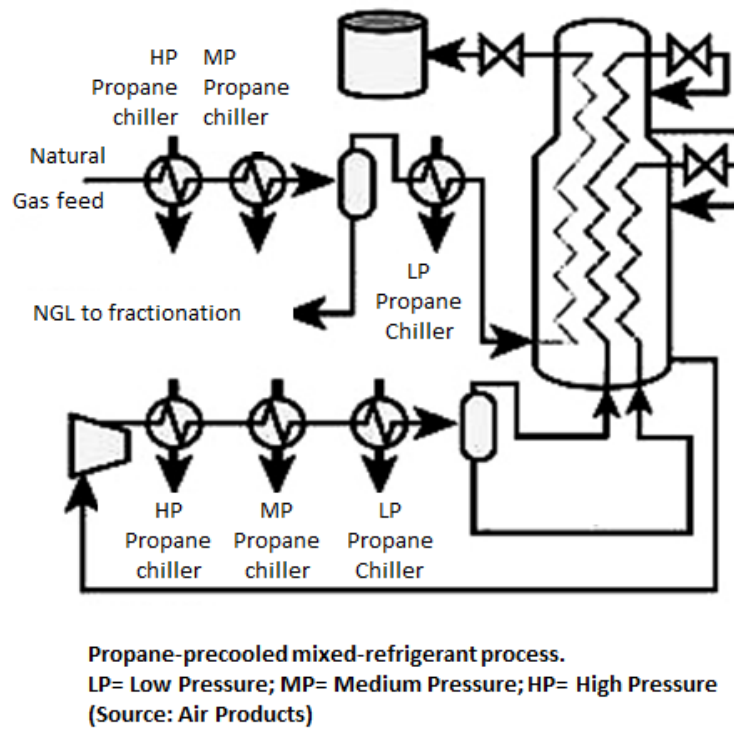
Figure 17: Propane precooled mixed refrigerant process [106].

3.3 MODEL DEVELOPMENT

Figure 18 shows a typical Air Products C3MR liquefaction process (APCI-C3MR). Air Products is a world-leading industrial gases company that provides gas-related equipment to manufacturing markets. The high-pressure natural gas (pipeline 1) is first expanded, thereby reaching a lower pressure, with concomitant temperature drop, via a turbo-expander (optional), and then cooled to -163oC by a series of LNG heat exchangers (H-E1, H-E2, H-E3 and H-E4). These heat exchangers cool down the natural gas to a sufficiently low temperature, and then liquefy it when throttled to the storage pressure.

The low-pressure Mixed Refrigerants (MR) are first compressed and then pre-cooled by a simple propane cycle. The mixed refrigerants provide cooling for the natural gas in the LNG heat exchangers. The booster/turboexpander is employed in this process, and the

work recovered from the turboexpander is employed to drive the MR booster. The pipeline numbers describe the sequence of process steps. The booster/turbo-expander is not widely used in FPSOs [9].
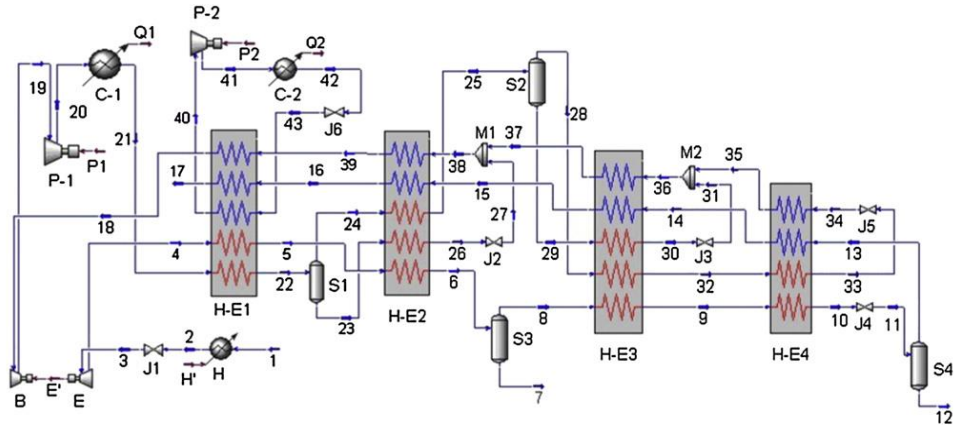
H-E1, H-E2, H-E3 and H-E4, LNG heat exchangers; S1, S2, S3 and S4, separators; M1 and M2, Mixers; P-1 and P-2, compressors; C-1 and C-2, after coolers; H, heater; B and E, booster/turbo expander; J1-J6, valves; P1, P2, Q1, Q2, E' and H', energy; numbers of 1-43, pipelines.

Figure 18: C3MR liquefaction process [9].

The low-pressure Mixed Refrigerants (MR) are compressed through a three-stage compression system (Low Pressure MR, Medium Pressure MR and High Pressure MR), while the propane is compressed through a single compressor. The liquefaction process consists of two liquefaction trains, and has two Gas Turbine (GT) drivers, and four heat exchangers (coldboxes) and separating units, as shown in Figure 18 and

Figure 19 [107] [9] [108]. The first GT drives the propane and high pressure MR compressors, while the other drives the low and medium pressure MR compressors. This configuration allows the power split between propane and MR refrigeration to be optimised while fully utilising the power available from the two drivers [108].

The reliability model introduced in this chapter is based on the widely used APCI-C3MR liquefaction process, and is illustrated in Figure 20.

Figure 19: 5 MTPA APCI-C3MR process schema [108].

Figure 20: LNG Liquefaction system. Two gas turbines, three MR Split refrigerant compressors, one propane compressor and four coldbox and separator units (CB & SEP).

3.4 SYSTEM RELIABILITY MODELLING

As identified in Section 2.3.2 of the literature review, Markov analysis offers better utility when analysing failure rates in systems where redundant units are introduced sequentially

following failure of the primary unit. Although the complexity of MA makes it a very difficult option when evaluating large systems, it gives a better accuracy than FTA in predicting fault rates in standby and sequential systems, as shown in Table 1 [39].

This section is divided into three subsections. The first subsection presents reliability and cost data, and gives an estimate of the cost of a standby liquefaction system. The second subsection shows the proposed system redundancy schemes which will be analysed later in this chapter. Finally, the last subsection explains the Markov state transition diagrams developed and utilised in the calculation of system reliability, and presents the results of the reliability and cost analysis.

3.4.1 RELIABILITY AND COST DATA

The basic input data needed for analysis of cost is presented in Table 3 [109] [8] [110]. Based on this information, it is possible to calculate the approximate total cost of a standby liquefaction system, including installation and commissioning. The approximate cost of a full standby system is US$66 million, while the cost of a standby system for MR and Propane compressors only is approximately US$5.5 million. The most expensive component in any liquefaction plant is the gas turbine, at approximately US$40 million; while the component with the lowest price is the coldbox which costs approximately US$120,000.

Table 3: Basic input data.

Item	Estimated cost (US$)
MR/C3 LNG centrifugal compressor	850,000
LNG gas turbine driver (Frame 6/7 – 40 MW)	40,000,000
Coldbox (Heat exchanger)	120, 000
LNG separator	150,000
Piping of a 100% redundant system	600,000
Standby system for compressors	5,500,000
Standby system for compressors & gas turbines	94,000,000
100% standby system	98,580,000
One LNG Ton	500 *(average LNG prices in 2013 and 2014)*

Table 4 gives the failure and repair rates of the liquefaction subsystems considered in this study. The reliability data presented in Table 4 are taken from OREDA Handbook [22]. OREDA is a project organisation sponsored by eight oil and gas companies with worldwide operations. The main task of OREDA is to collect and exchange reliability data among its participating companies, and to act as the forum for the co-ordination and management of reliability data collection within the oil and gas industry. The failure rates presented in Table 4 are given per train and not per unit. A train consists of one of more compressors installed in series.

Table 4: Transition rates [22].

Transition rate (per hour)	Description
$\lambda_{Ct} = 9.15 \times 10^{-4}$	Failure rate of centrifugal compressors train (MR, C3). $\lambda_C = 0.5\lambda_{Ct}$, where λ_C is the Failure rate of a single compressor.
$\lambda_D = 2.68 \times 10^{-3}$	Failure rate of Driver (Gas turbine)
$\lambda_H = 0.2 \times 10^{-4}$	Failure rate of Coldbox (Heat exchanger)
$\lambda_S = 0.51 \times 10^{-4}$	Failure rate of LNG separator
$\lambda_E = 0.55 \times 10^{-4}$	Failure rate of LNG expander
$\mu_C = 5.3 \times 10^{-2}$	Repair rate of centrifugal compressor (MR/C3)
$\mu_{CA} = 1.3 \times 10^{-2}$	Repair rate of all centrifugal compressors in the system
$\mu_D = 4.5 \times 10^{-2}$	Repair rate of GT Driver
$\mu_{DA} = 1.1 \times 10^{-2}$	Repair rate of all GT Drivers in system
$\mu_H = 2.5 \times 10^{-1}$	Repair rate of Coldbox (Heat exchanger)
$\mu_S = 1.8 \times 10^{-1}$	Repair rate of LNG separator
$\mu_E = 3 \times 10^{-2}$	Repair rate of LNG expander
$\gamma_G = 4$	Estimated rate of starting up standby compressors and drivers (from 15 to 20 minutes) [109]

3.4.2 PROPOSED REDUNDANCY SCHEMES

As shown in Figure 21, a typical C3MR liquefaction system consists of two gas turbine drivers, four centrifugal compressors, four coldboxes (heat exchangers), and four separators (see

Figure 20). Of all the system components, the centrifugal drivers and compressors are mostly likely to fail, as shown in Table 4. Several redundancy options have been introduced (see Figures 12-15) including a complete cold standby system with a configuration similar to the primary liquefaction unit, as outlined in

Figure 24. In this study, the reliability performance of the following four system options is investigated.

The proposed system options are as follows:

1. A basic configuration with a 1 x 100% C3MR liquefaction system. This system may be considered as a series configuration with no redundancy, as presented in Figure 21.

2. A basic system with the addition of 100% cold standby MR and Propane compressors, as presented in

   Figure 22. A cold standby system is a method of redundancy in which the secondary system is only called upon when the primary system fails.

3. A basic system with the addition of 100% cold standby MR and Propane compressors, and a cold standby GT driver, as presented in Figure 23. The standby units will switch on when the primary units fail.

4. A basic system with the addition of 100% cold standby MR and Propane compressors, a cold standby GT drive, and 100% cold standby LNG coldboxes (heat exchangers) and separators, as presented in

   Figure 24. Since the prices of coldboxes and separators are low in comparison to drivers and compressors, option 4 introduces a standby redundancy system for both coldboxes and separators in addition to compressors and drivers. This system will activate whenever any primary unit fails. Therefore, Option 4 may be considered as a combination of options 1, 2 and 3, as shown in

   Figure 24.
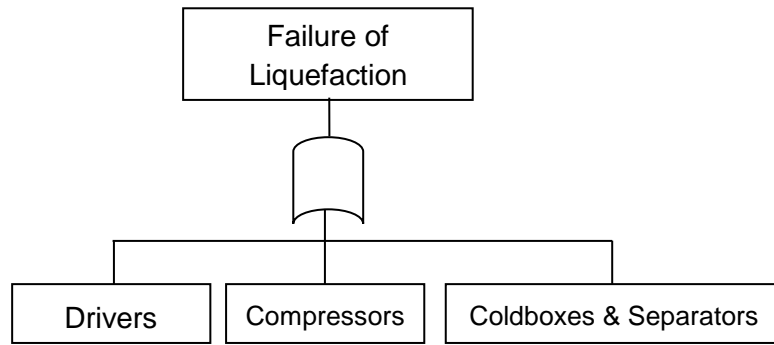
Figure 21: Fault tree of a liquefaction system without redundancy (Option 1).
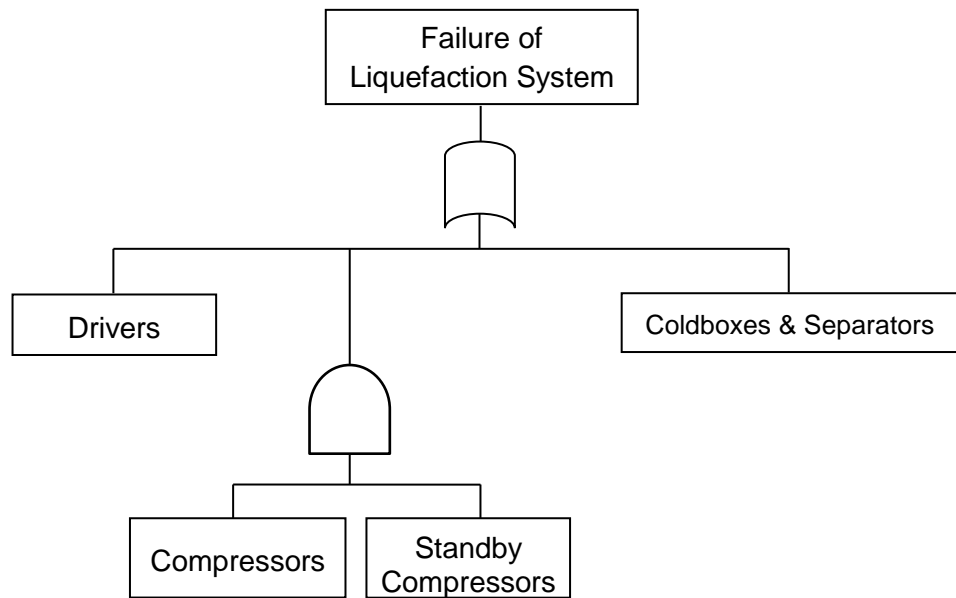
Figure 22: Fault tree of a liquefaction system with 100% cold standby compressors

(Option 2).
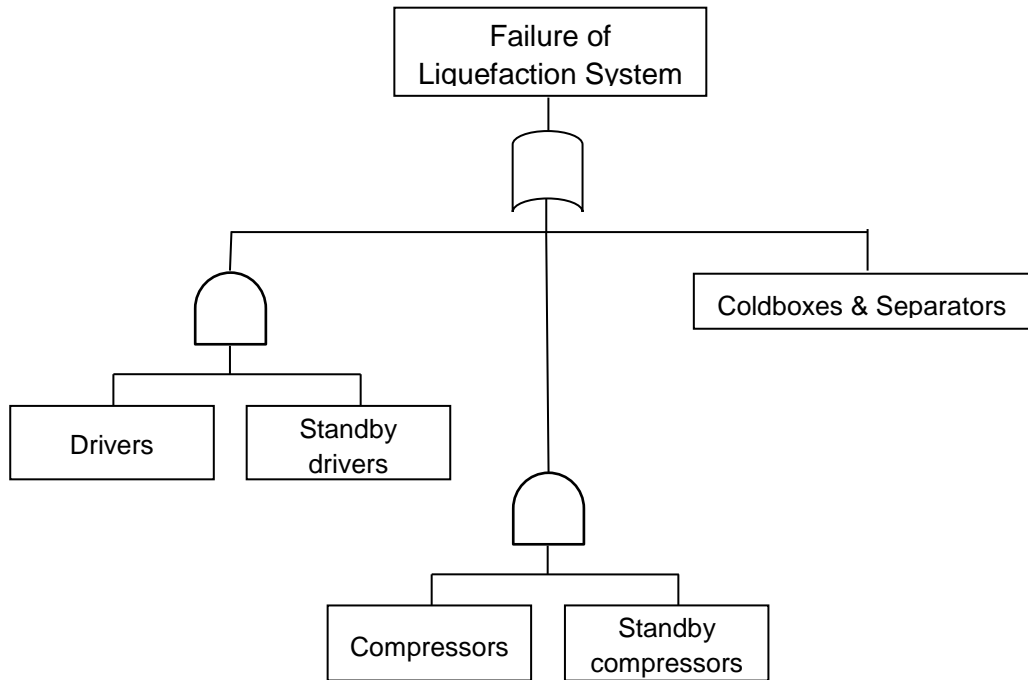
Figure 23: Fault tree of a liquefaction system with 100% cold standby compressors and drivers (Option 3).
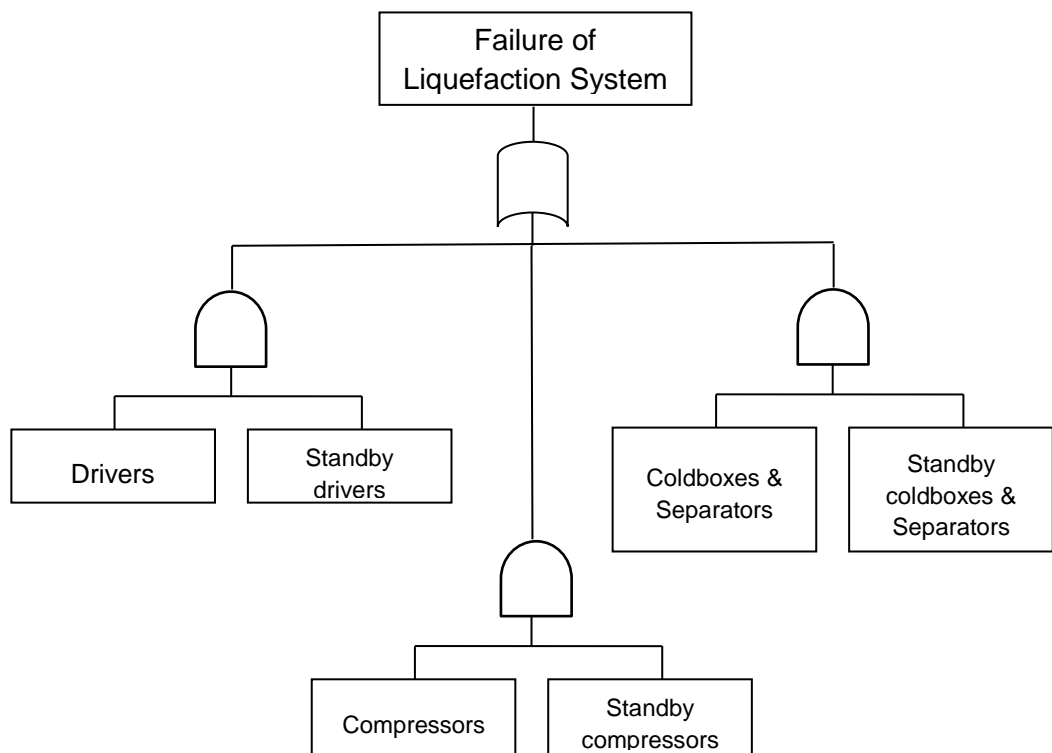
Figure 24: Fault tree of a liquefaction system with full redundancy (Option 4).

3.4.3 SYSTEM RELIABILITY MODEL AND RESULTS

During the development of Markov state transition diagrams, the following assumptions were made:

- The model considers the useful working life of the machine. According to the Bathtub curve, failure and repair rates are constant over time and statistically independent of each other.

- At any given time, the system is either in the operating state or in the failed state.

- Sufficient repair facilities are available.

- Standby units are of the same nature and capacity as that of active systems.

- Service schedule includes repair and/or replacement.

- The performance of the repaired unit is as good as that of a new unit.

- The standby system is perfectly reliable (cold standby) and the production will immediately be switched to the cold standby system if any of the main system components have failed.

- System failure/repair follows the exponential distribution [111] as it describes the time between events in a Poisson process. An example of such a process is when events occur continuously and independently at a constant average rate, and the Mean Time To Repair (MTTR) is equal to the inverse of the mean downtime after failure.

Markov Analysis and system reliability calculations are carried out using Isograph reliability software. This software has an integrated environment for performing Reliability Prediction, Maintainability Prediction, Failure Mode Effect and Criticality Analysis (FMECA), Reliability Block Diagram (RBD) analysis, Fault Tree Analysis, Event Tree Analysis, and Markov Analysis [112].

Table 5 shows all possible states of the system proposed. The system has 21 states and includes two gas turbines (GT and Gt2), four LNG compressors (C1 to C4) and four coldboxes and separators (CB1 to CB4).

Table 5: Markov State transitions of a main and a standby C3MR LNG liquefaction system. GT, C and CB stand for Gas Turbine, Compressor and ColdBox, respectively.

State	Description	GT1	GT2	C1	C2	C3	C4	CB1	CB2	CB3	CB4
0	The main system is working and the Standby system is functioning	W/S	W/S	W/S	W/S	W/S	W/S	W/S	W/S	W/S	W/S
1	The main system has failed (GT1 failed) and the standby system is functioning	**F**/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S
2	The main system has failed (GT2 failed) and the standby system is functioning	S/S	**F**/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S
3	The main system has failed (C1 failed) and the standby system is functioning	S/S	S/S	**F**/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S
4	The main system has failed (C2 failed) and the standby system is functioning	S/S	S/S	S/S	**F**/S	S/S	S/S	S/S	S/S	S/S	S/S
5	The main system has failed (C3 failed) and the standby system is functioning	S/S	S/S	S/S	S/S	**F**/S	S/S	S/S	S/S	S/S	S/S

Table 5 (Continued)

State	Description	GT1	GT2	C1	C2	C3	C4	CB1	CB2	CB3	CB4
6	The main system has failed (C4 failed) and the standby system is functioning	S/S	S/S	S/S	S/S	S/S	**F**/S	S/S	S/S	S/S	S/S
7	The main system has failed due to a failure in CB1 and the standby system is functioning	S/S	S/S	S/S	S/S	S/S	S/S	**F**/S	S/S	S/S	S/S
8	The main system has failed due to a failure in CB2 and the standby system is functioning	S/S	S/S	S/S	S/S	S/S	S/S	S/S	**F**/S	S/S	S/S
9	The main system has failed due to a failure in CB3 and the standby system is functioning	S/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S	**F**/S	S/S
10	The main system has failed due to a failure in CB4 and the standby system is functioning	S/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S	S/S	**F**/S
11	The main system has failed due to a failure in GT1 and the standby system is working	**F**/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W
12	The main system has failed due to a failure in GT2 and the standby system is working	S/W	**F**/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W

Table 5 (Continued)

State	Description	GT1	GT2	C1	C2	C3	C4	CB1	CB2	CB3	CB4
13	The main system has failed due to a failure in C1 and the standby system is working	S/W	S/W	**F**/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W
14	The main system has failed due to a failure in C2 and the standby system is working	S/W	S/W	S/W	**F**/W	S/W	S/W	S/W	S/W	S/W	S/W
15	The main system has failed due to a failure in C3 and the standby system is working	S/W	S/W	S/W	S/W	**F**/W	S/W	S/W	S/W	S/W	S/W
16	The main system has failed due to a failure in C4 and the standby system is working	S/W	S/W	S/W	S/W	S/W	**F**/W	S/W	S/W	S/W	S/W
17	The main system has failed due to a failure in CB1 and the standby system is working	S/W	S/W	S/W	S/W	S/W	S/W	**F**/W	S/W	S/W	S/W
18	The main system has failed due to a failure in CB2 and the standby system is working	S/W	S/W	S/W	S/W	S/W	S/W	S/W	**F**/W	S/W	S/W
19	The main system has failed due to a failure in CB3 and the standby system is working	S/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W	**F**/W	S/W

Table 5 (Continued)

State	Description	GT1	GT2	C1	C2	C3	C4	CB1	CB2	CB3	CB4
20	The main system has failed due to a failure in CB4 and the standby system is working	S/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W	S/W	**F**/W

W: Working / S: Standby/ F: Failed/ -/-: Main system/standby system (for example, S/W means that the main system is in "standby mode" and the standby system is working)

Figure 25 presents the main Markov state transition diagram developed to investigate the introduction of partial and full redundancy options to the main C3MR LNG liquefaction system. Cold standby systems are considered in all partial and full redundancy system options. The rates of failure (λ) and repair (μ) for all system components are indicated in Table 4. The standby system will not be considered if $\gamma_G = 0$. States (1) to (10) present the unavailability of the system. System unavailability percentages were calculated using Isograph software at four different system configurations.
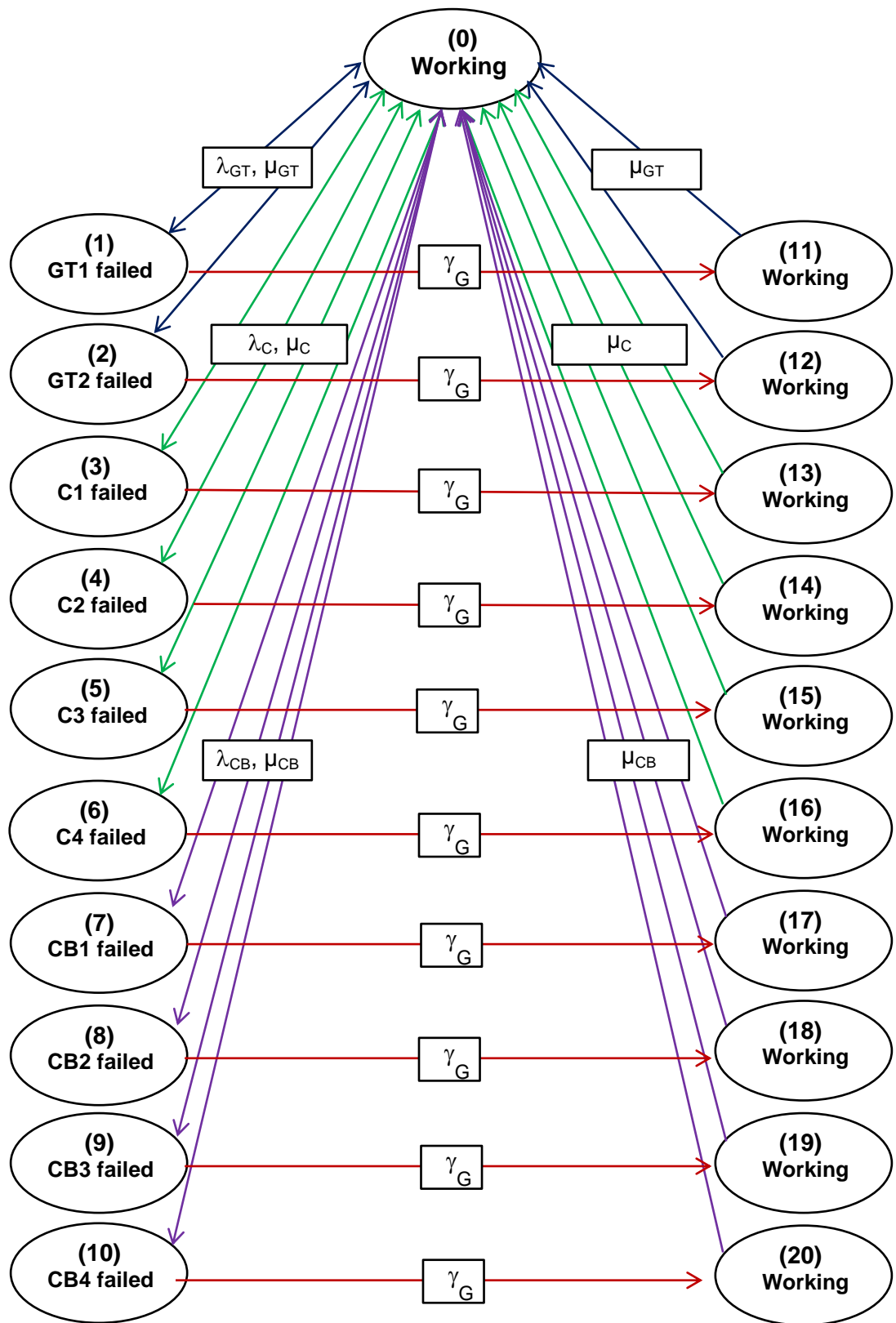
Figure 25: Markov state transition diagram for the C3MR LNG liquefaction system.

Figure 26 presents the availability percentage of a basic liquefaction system with no standby redundancy (option 1). The system unavailability percentage at this system configuration is 15.9%.
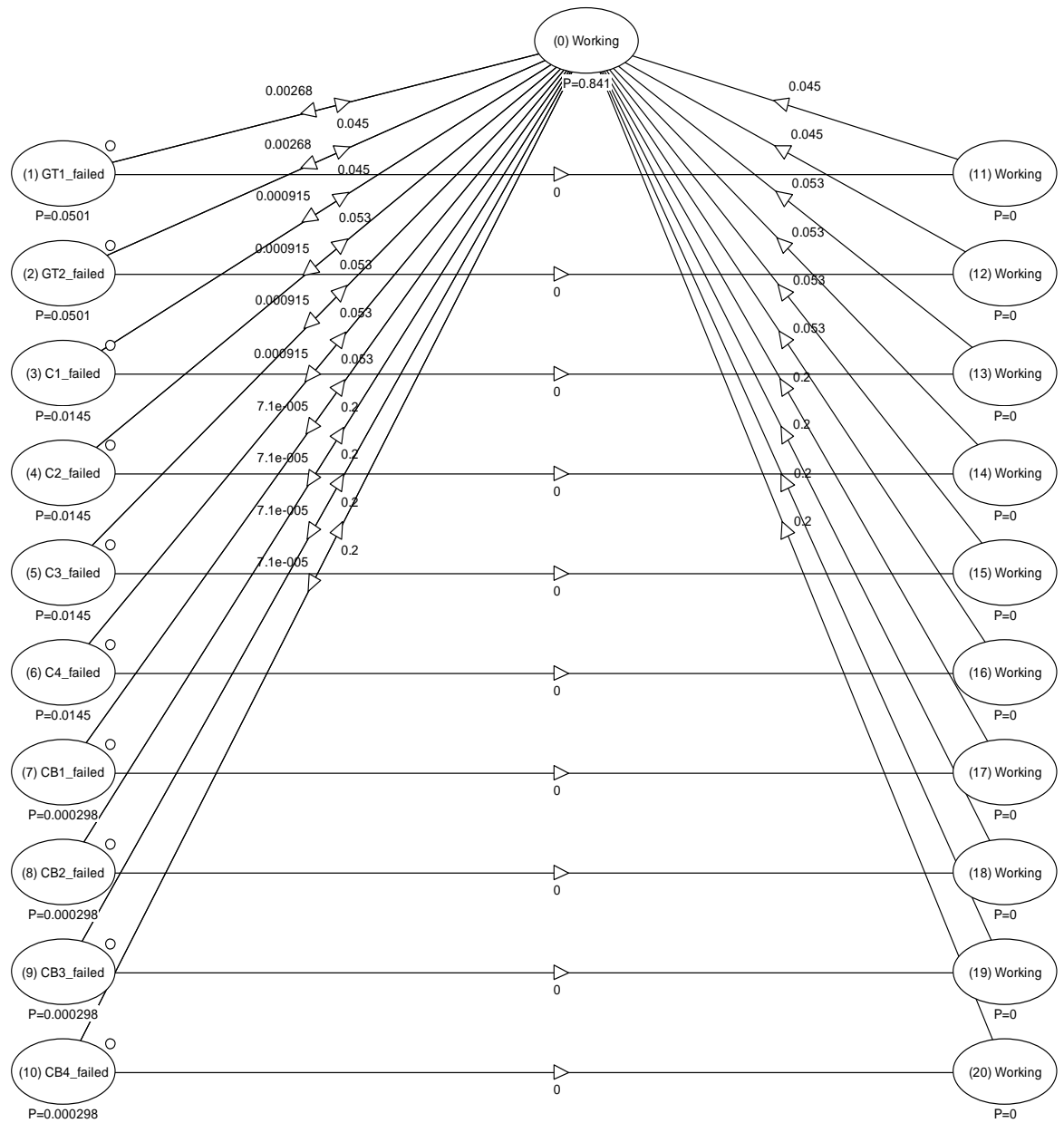
Figure 26: Availability results of a basic C3MR liquefaction system with no standby redundancy using Isograph software.

Figure 27 shows the availability percentage of a basic liquefaction system with 100% standby gas turbine drives (option 2). The system unavailability percentage at this system configuration is 6.04%.
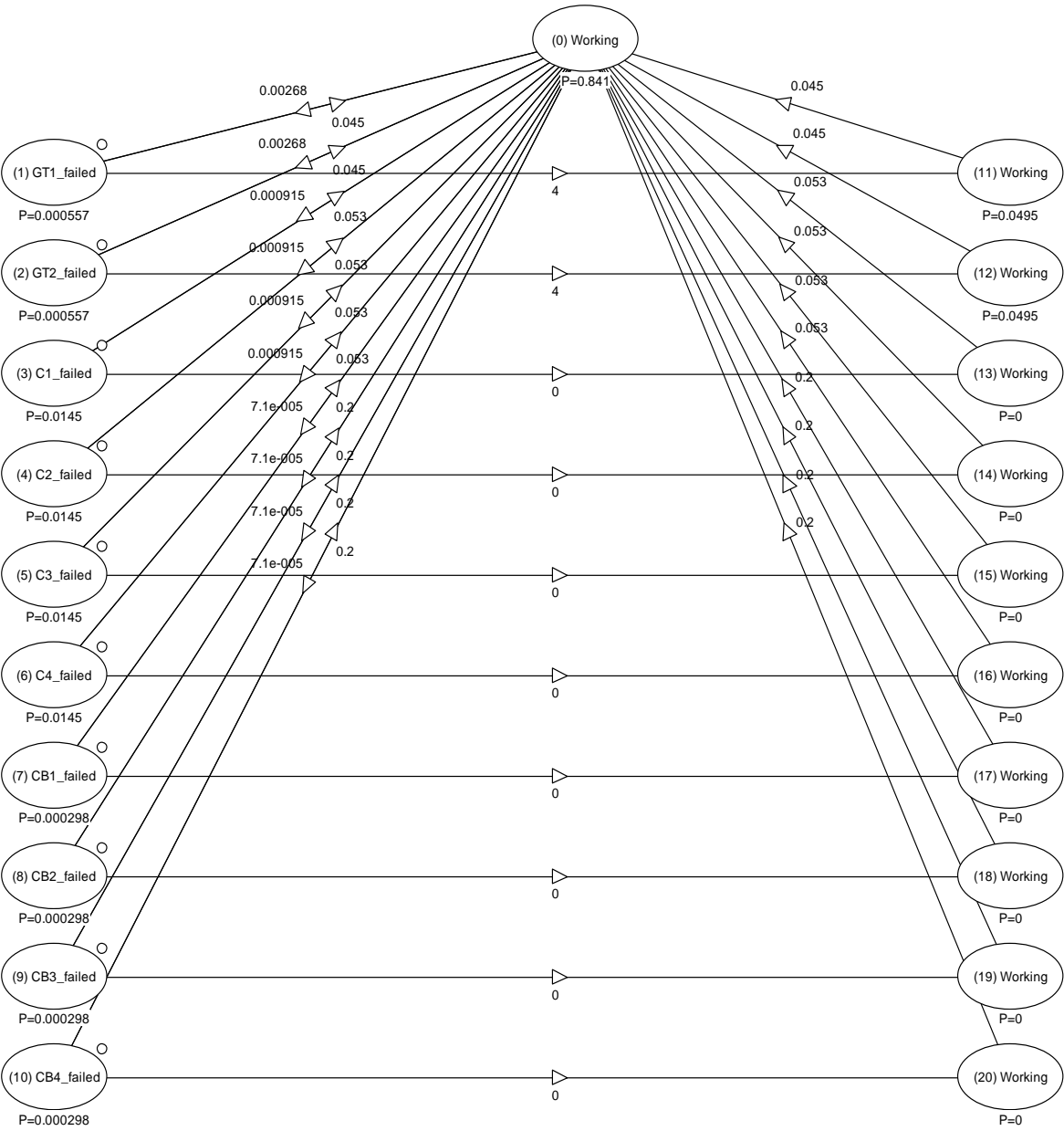
Figure 27: Availability results of a basic C3MR liquefaction system with the addition of 100% cold standby GT drivers using Isograph software.

Figure **28** depicts the availability percentage of a basic liquefaction system with 100% gas turbine drivers and compressors (option 3). The system unavailability percentage at this system configuration is 0.3%.
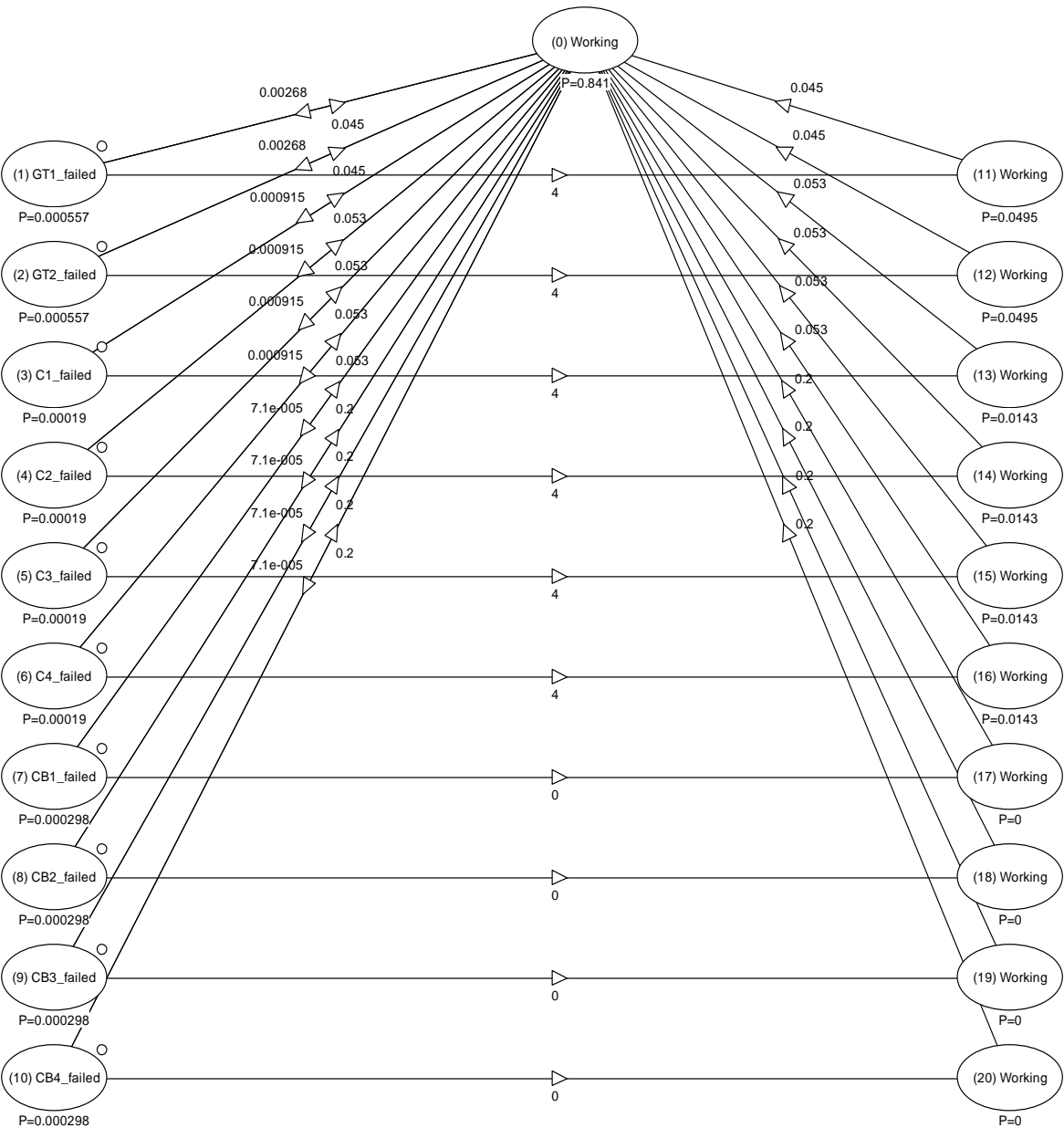
Figure 28: Availability results of a basic C3MR liquefaction system with the addition of 100% cold standby GT drivers and 100% MR and Propane compressors using Isograph software.

While Figure 29 presents the availability percentage of a basic liquefaction system with 100% standby system (option 4). The unavailability percentage at this system configuration is 0.19%. The failure and repair rates for coldboxes and separators were calculated using equations (1) and (2).

$$\lambda_{CB} = \lambda_{coldbox} + \lambda_{separator} \tag{1}$$

$$\mu_{CB} = \mu_{coldbox} + \mu_{separator} \tag{2}$$
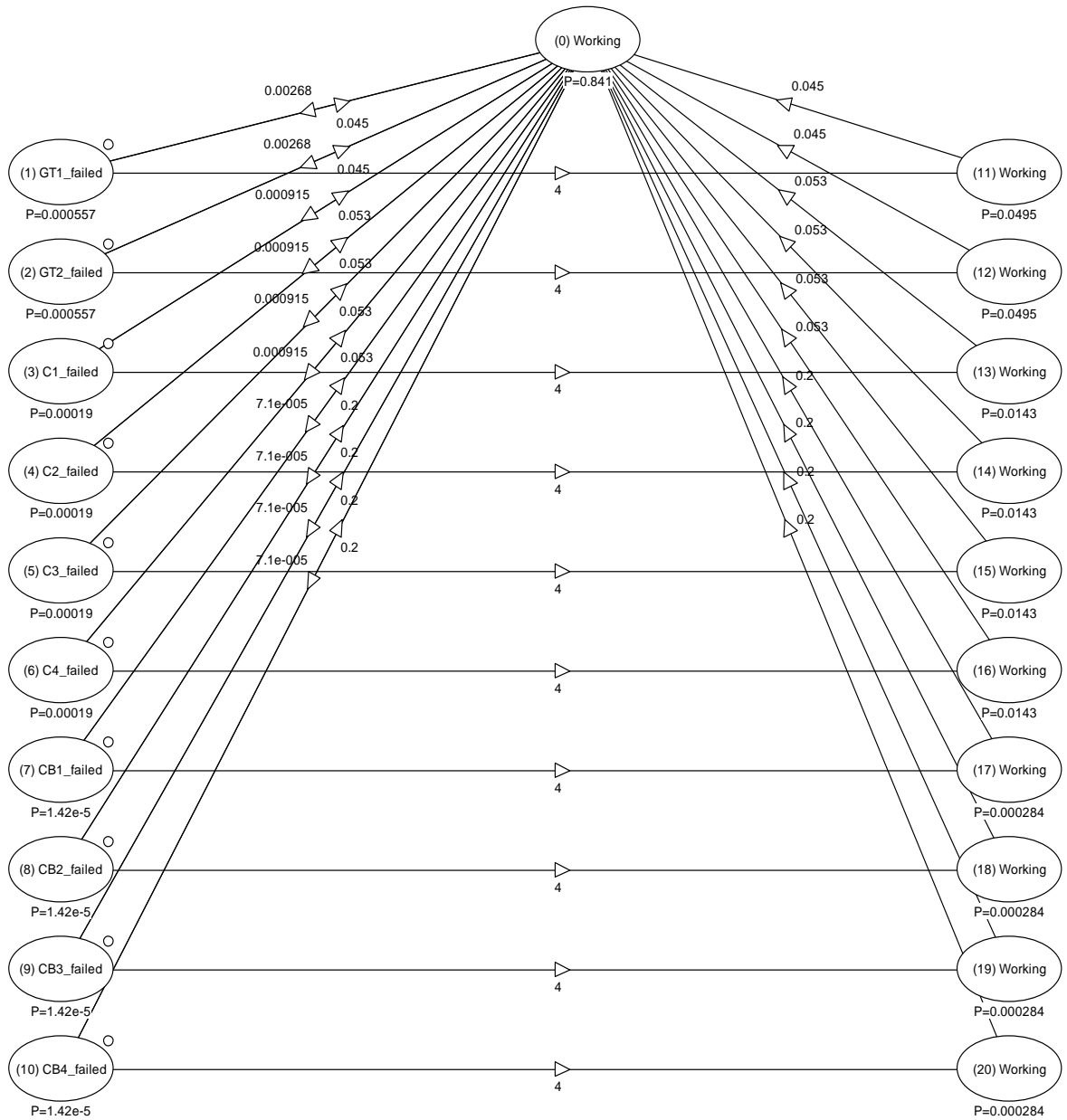
Figure 29: Availability results of a basic C3MR liquefaction system with the addition of 100% cold standby GT drivers, 100% cold standby MR and Propane compressors and 100% standby coldboxes and separators using Isograph software.

The results of the reliability analysis are presented in Table 9. The cost calculations were based on a typical 3 MTPA LNG liquefaction plant. It was assumed that all repairs are carried out on board, and that the majority of spare parts are available on site. The time-dependent Markov Analysis technique was used to determine the system unavailability

percentage. Typical production capacity of LNG FPSOs ranges from 3 to 3.5 MTPA. For the smallest capacity (3 MTPA), the annual FPSO income will be around US$1500 million, if no failures occur. The annual instalment calculations are based on a 15-year loan period with a 3% interest rate.

Three factors were to be considered in the evaluation of the cost-effectiveness of different system redundancy options: the system shutdown cost, which can be calculated on the basis of the unavailability profile of the system; the annual investment cost, which indicates the annual instalments that have to be paid for the installation and commissioning of the corresponding redundant systems; and the annual maintenance cost for the different options. It was assumed that there will be one scheduled shutdown per year for preventive maintenance costing a total of US$70 million. The estimated cost was based on the minimum shutdown time length (ranging from two weeks to two months [113]), and on the time required to restore full production capacity after the scheduled shutdown. A minimum of 15 emergency/unscheduled shutdowns a year for corrective maintenance was also assumed, at a total cost of US$1.5 million. The cost of emergency shutdowns included labour and spare parts only, as the shutdown cost was already included in the reliability calculation for the system.

When redundancy is partially or fully applied, the corrective maintenance for the failed redundant component will be carried out while the liquefaction plant is fully in service. Although the frequency of failure is subject to increase due to the exclusion of PM, the cost of major corrective maintenance is not expected to increase, as the cost of repair depends on the urgency of the maintenance request. All of the previously mentioned factors affect the profitability of LNG production plants, and hence the best liquefaction system is the one that has the lowest total shutdown, annual investment and maintenance costs.

Figure 29 and Table 9 show that the introduction of a standby setup dramatically reduced the unavailability time from 15.9 %, with no standby units, to 0.19 %, with 100% standby units. This study has not considered the detailed practical considerations of applying partial redundancy. It is possible that the standby system may need to be installed as a separate production line that includes all system components; to avoid complexity in installation and operation of a standby unit parallel to each of the main system components. However, this needs further investigation.

Table 6: Summary of the reliability study results.

Measure	Option 1	Option 2	Option 3	Option 4
System Unavailability	15.9%	6.04%	0.3%	0.19%
Shutdown cost (USD)	238,500,000	90,600,000	4,500,000	2,850,000
Annual cost estimate of major preventive maintenance tasks (USD) including shutdown cost– Once a year.	70,000,000	42,000,000	7,000,000	0
Annual cost estimate of major corrective maintenance tasks (USD)	1,500,000	2,100,000	2,,850,000	3,000,000
Total shutdown and maintenance cost (USD)	310,000,000	134,700,000	14,350,000	5,850,000
Annual investment (USD) [to be paid over 15 years with a 3% annual interest rate]	0	455,784	7,789,761	8,169,305
Annual profit (USD)	1,190,000,000	1,364,844,216	1,477,860,239	1,485,980,695

Table 9 shows that option four is the best in terms of cost-effectiveness as it dramatically reduces the unavailability time from 15.9 % to 0.19 %. This option introduces a 100% redundant LNG liquefaction system, which is estimated to increase profitability by

approximately US$296 million per year (from US$1,190 million to US$1,485.98 million per year).

Option 2 introduces a 100% redundant system for gas compressors only and this redundancy option would increase annual profitability by approximately US$174.8 million. Option 3 introduces a 100% redundant system for gas compressors and GT drivers only, and this option would increase profitability by approximately US$287.9 million per year. Finally, the last option would increase profitability by around US$296 million per year. The cost analysis results are based on the average LNG prices (US$500/ton) in 2013 and 2104.

3.5 SUMMARY

The redundancy optimisation of the C3MR liquefaction system on FPSO terminals was investigated. Maintenance intervals were also optimised in order to reduce the total associated maintenance cost. The main reliability results were obtained using the Markov approach. Isograph software and MATLAB were utilised to perform the reliability and maintenance optimisation calculations.

The implementation of a 100% standby system drastically reduced unavailability from 14.3% to 2.6% of the total operational hours. Based on the proposed reliability model, the annual system profit would increase by approximately US$236 million if redundancy option 4 were implemented on FPSOs. Hence, the cost-effectiveness of introducing a 100% standby system for the primary liquefaction system on FPSOs (option 4) has been demonstrated.

If redundancy option 4 is applied, the cost of maintenance will significantly reduce due to the introduction of a 100% standby system and the elimination of PM costs; in turn, the overall cost of shutdown and PM maintenance will decrease by US$236 million. The PM

cost will be almost zero since only a corrective maintenance strategy will be applied [8] [49].

In order to swiftly switch between main and standby systems, a dynamic CM system will be developed and presented in the following chapters with a view to providing a more cost-effective solution for maintenance problems and eliminating production stoppage during switching over between main and standby systems.

The study was only applied to the C3MR LNG liquefaction plant. Hence, the proposed reliability model shall be utilised only to investigate the reliability of similar plants. The estimates of the costs of different LNG liquefaction components, and of CM and PM maintenance are approximations.

To ensure that the results were as realistic as possible, the reliability calculations were based on the only available source of information, which was the OREDA Handbook 5th edition. The results of this study have the potential to improve the commercial impact of FPSOs by reducing both shutdown and maintenance costs, and so increase overall profitability.

Chapter 4. Experimental setup for fault detection

4.1 INTRODUCTION

Centrifugal equipment has the highest failure rate of all LNG production line equipment [22]. Bearing problems account for 40% of the failures which occur in all machines, including centrifugal equipment [114]. Hence, the experimental work in this thesis is focused on bearing faults in centrifugal equipment.

The main contribution of this chapter is to describe the experimental hardware which was commissioned as part of the research project in order to facilitate experiments relating to the development and testing of monitoring algorithms for centrifugal equipment.

In order to monitor acoustic signals that can be indicative of faults, two different types of AE sensors were installed and evaluated and, based on experiments using these sensors, the most suitable of the two was selected. This decision was informed by a combination of experience and experiments that quantify the AE transmission loss. The preferred locations of the vibration and pressure sensors were also determined.

Pressure and triaxial vibration sensors were installed to monitor pressure and vibration information that could also be indicative of faults; they were then utilised in combination with the AE technique to characterise the major fault detection techniques for the CM of high-speed centrifugal compressors (see Chapter 7). The measurement system including the antialiasing filter is also described in this chapter.

4.2 EXPERIMENTAL SETUP

This study aims to develop and test an efficient CM system for LNG centrifugal equipment, in particular LNG refrigerant compressors. Ideally, the experiments presented in this thesis should have been carried out using a full-scale industrial compressor. Due to lack of access to such a compressor, a functionally similar lab-scale unit was utilised.

The experimental setup consisted of a 20 HP Paxton industrial compressor, four AE sensors with different measurement ranges, a triaxial accelerometer, a pressure sensor, preamplifiers with programmable filters, a high-speed NI data acquisition system, and an AC inverter for motor speed control. The models of bearings (A) and (B) are DKT-7203BMP and FAG-2203TV, respectively. The data was collected using an MSeries-PCI 6250 National Instruments data acquisition board with 16 channels, a 16-bit resolution and a 1.25 MS/s sampling rate. A schema of the overall system and the bearing casing with sensor locations can be seen in Figure 30 and
Figure 31.

The AE sensors were positioned as close as possible to the bearings, and attached to signal conditioners and programmable low-pass filters with isolated grounds, in order to combat the problem of aliasing in sampling signals. A cut-off frequency of 200 kHz was set to attenuate high-frequency AE signals. The vibration sensor was positioned midway between the bearings, and the pressure sensor was installed inside the air hose 50 cm away from the compressor outlet.
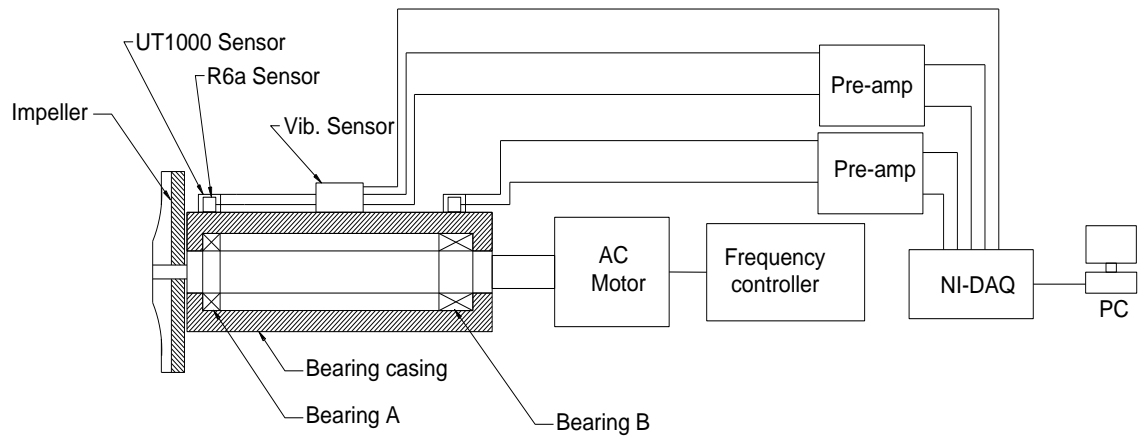
Figure 30: Schema of the experimental setup
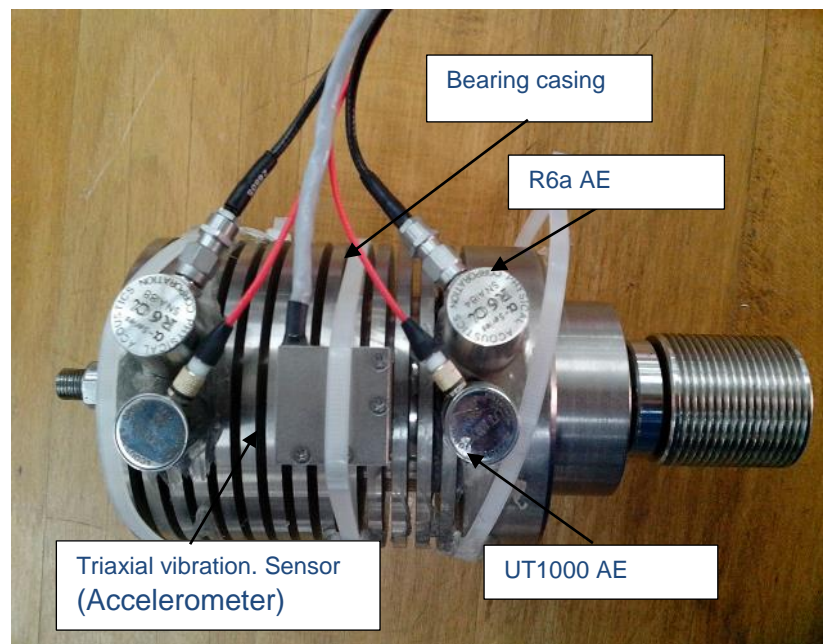
Figure 31: Locations of the sensors on the bearing casing.

4.2.1 COMPRESSOR

A Paxton AT1200 industrial single-stage centrifugal air compressor system was selected as the subject for the experimental parts of this project. The blower has a maximum flow rate of 1954 m^3/hr (1150 CFM) and a maximum pressure of 1.254 BarA (100 inch water) as shown in appendix B. Figure 32 shows the disassembled compressor.
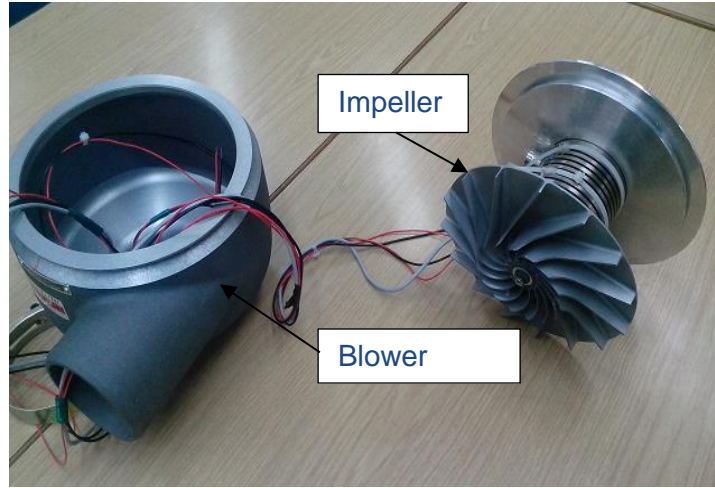
Figure 32: The single stage centrifugal compressor.

The similarities between the lab scale compressor and the full scale LNG compressor are based on the fact that the vibration component frequencies of bearing faults (bearing fundamental frequencies) in rotating equipment are a function of rotational speed and bearing geometry [115] [116]. A quantitative example is given in Table 10 in order to explain the change in bearing fundamental vibration frequency due to bearing geometry and machine operational speed changes. The comparison results were calculated using equation (3). The results showed that the bearing utilized in the prototype has an outer ring ball pass frequency of 1072.5 Hz in comparison to 343.6 Hz for the NSKHPS 7956A5 high precision industrial bearing.

$$\mathbf{BPF} = \frac{\mathbf{Nb}}{\mathbf{2}}\mathbf{S}\left(\mathbf{1} - \frac{\mathbf{Bd}}{\mathbf{Pd}}\cos\emptyset\right) \tag{3}$$

where

BFP: Ball Pass Frequency

Nb: Number of ballsS: Rotational speed in RPS

Bd: Ball diameter

Pd: Pitch circle diameter

∅: Contact angle

Table 7: Ball pass vibration frequency of different types of bearings based on bearing geometry and rotational speed.

Parameter	Prototype bearing	Industrial compressor bearing
Bearing model and max. static load capacity	DKT 7203BMP, 6.6 KN	NSKHPS 7956A5, 390 KN
Ball diameter (Bd)	7 mm	60 mm
Number of balls (Nb)	11	14
Pitch circle diameter (Pd)	28 mm	330 mm
Rotational speed (S)	260 RPS	60 RPS
Contact angle (∅)	0	0
Ball Pass Frequency of outer ring (BPF)	1072.5 Hz	343.6 Hz

With ultrasonic acoustic signals, the situation is similar as the component frequencies depend on the high-frequency natural resonances of bearings. High natural frequencies are generated by impacts betweenthe internal parts of bearings. These impacts are the results of bearing imperfection, degradation or variation in load [117].

It is common for a technique called High-Frequency Natural Bearing Resonance Indicator (HFNBI) to be utilised to monitor the friction between bearing rollers and race [116] [118] [119]. When the motion of the bearing rollers degrades, the rollers slide momentarily, and this usually results in a friction force change between the rollers and the race. Again, the amplitude of the AE signal is mainly a function of rotational speed, geometry and of course the level of degradation. HFNBI normally ranges from 3 to 50 kHz, and can be expressed as a percentage rise in the spike energy when compared to its normal good condition.

The Shock-Pulse technique is one of the most commonly utilised methods for monitoring the condition of bearings. This technique is based on the fact that any bearing

imperfection generates a white noise which contains all frequencies. As the natural frequency of the AE sensor is one of these frequencies, the sensor will have a resonance in the 30-90 kHz range depending on its natural frequency [120].

In summary, imperfection and sudden shock makes the AE sensor vibrate at its resonant frequency. Friction, wear and shock normally excite the resonant frequency of the AE sensor [116] [118] [119]. Thus, the acoustic signal component frequencies are a function of the natural frequency of bearings and AE sensors, and normally range from 3 to 90 kHz.

4.2.2 SENSOR SELECTION

Four factory calibrated AE sensors from Physical Acoustics were initially selected to measure a broad range of acoustic signals; they consisted of two R6a low frequency range sensors with an operating range of 35 to 100 kHz and with peak sensitivity of -65 dB, and two UT1000 high frequency range sensors with an operating range of 100 to 1000 kHz and a peak sensitivity of -70 dB. However, AE sensors can measure signals at frequencies outside their operating bandwidth albeit with less sensitivity and a growing phase lag (Note: the phase-lag is not an issue in this application where only amplitude is considered).

The final selection of the most useful sensor was based on the normal range used in HFNBR (normally from 3 to 50 kHz), backed up by experimental results, as presented in Figure 33. The detail of these experiments will be discussed more fully below. However, the graphs of frequency response from the two types of sensor confirm that the most significant frequencies (peak amplitudes) are in the 2-122 kHz range. This favours the use of the lower frequency R6a device. It should be noted that the device does still pick

up some peaks outside of this range. This is normal as AE sensors can measure outside their normal working bandwidth, though with reduced sensitivity.

Having selected the sensor type to be used, the final decision related to the placement of the sensor. Since there are two R6a sensors, one near bearing (A) and another near bearing (B), an AE transmission loss analysis was carried out to select the sensor with the highest output signal (section 4.3).
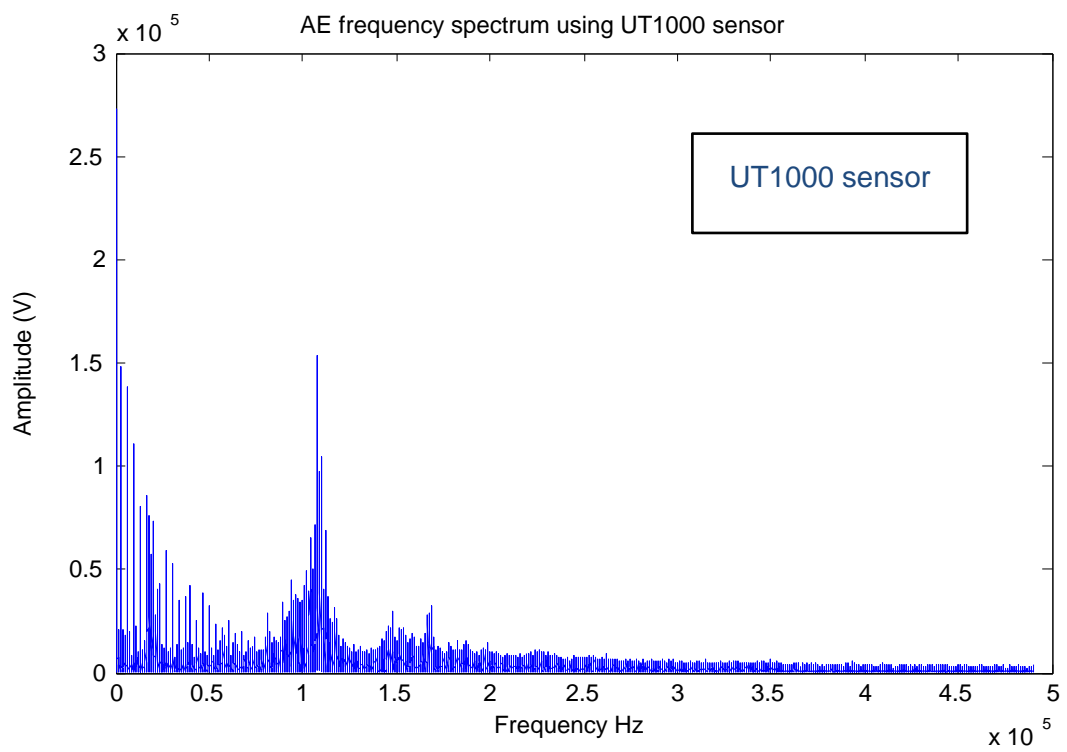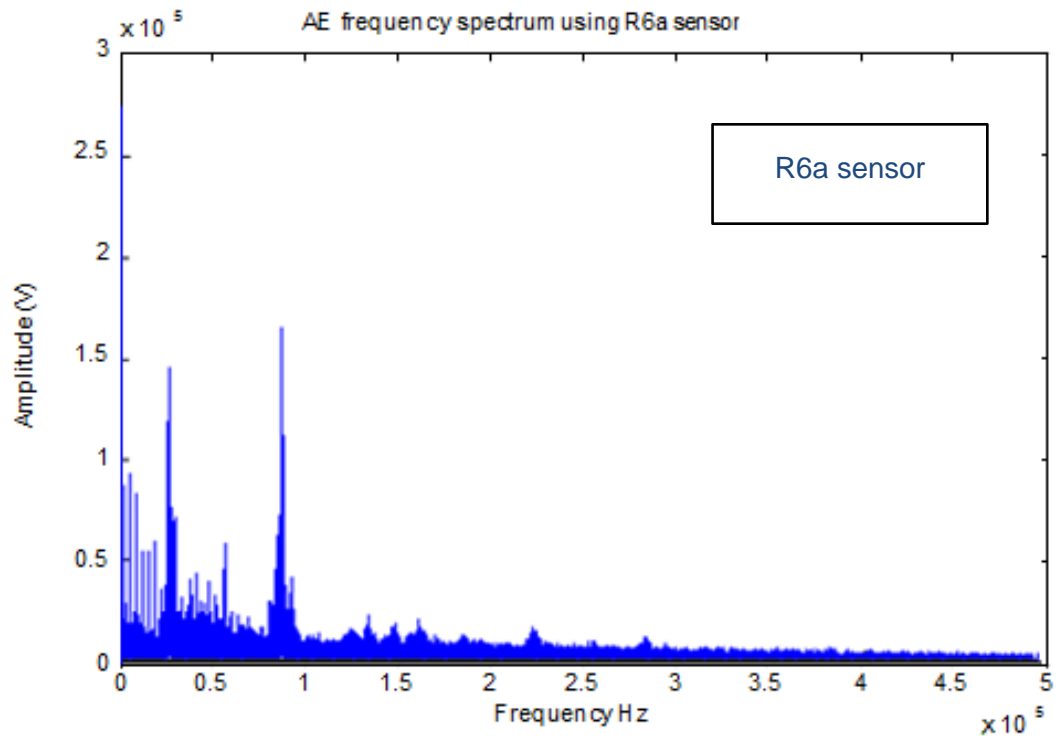
Figure 33: AE frequency spectra using R6a and UT1000 sensors with a frequency range of 2 to 499 kHz.

A +/-70 g triaxial vibration sensor, model SUMMIT 34207A with a sensitivity of 31.55 mV/g, was positioned midway between the shaft bearings. The selection of the accelerometer range was based on a number of experiments that were conducted in the lab. The RMS values of vibration signals related to different machine conditions presented in Table 40 were considered during the selection process of accelerometer range.

Modal analysis was carried out using SOLIDWORKS software to determine the natural frequencies of the bearing casing. SOLIDWORKS is a solid modelling Computer-Aided Design (CAD) and Computer Aided-Engineering (CAE) piece of software which includes finite element analysis and a multi-body dynamics program. A model for the bearings case was drawn to reflect the actual shape, dimensions and material. The modal problem was solved using the FEA modal superposition method with a fixed left-end-flange constraint.

The modal analysis results presented in Table 11 show that the first natural frequency value was 1,668.9 Hz, which was much higher than the 260.8 Hz operating frequency of the compressor. Therefore, due to the high stiffness level of the bearing casing, the accelerometer could be positioned at any location on the surface. Figure 34 illustrates that, with the second mode shape, the maximum dynamic magnification factor took place at the end of the casing at a natural frequency of 1669.2 Hz. However, the sensor can still be installed at any location even if the operating frequency is close to the first natural frequency of the system; this is due to the fact that pattern recognition algorithms are based on machine condition signatures, which are captured by placing the same sensor at the same location. On the other hand, all vibration signals will be magnified by the same factor and hence, the frequency spectrum shape (or pattern) will remain unchanged.

A pressure sensor was first selected on the basis of the maximum outlet pressure of the blower, installed in the outlet pipe, and then positioned 50 cm away from the blower outlet. The data sheets for the components utilised are provided in appendix D.

Table 8: Modal frequencies of the bearing casing.

Mode No.	Frequency (Hertz)
1	1668.9
2	1669.2
3	3028.8
4	5488.7
5	5489.1
6	6100.9
7	9100.5
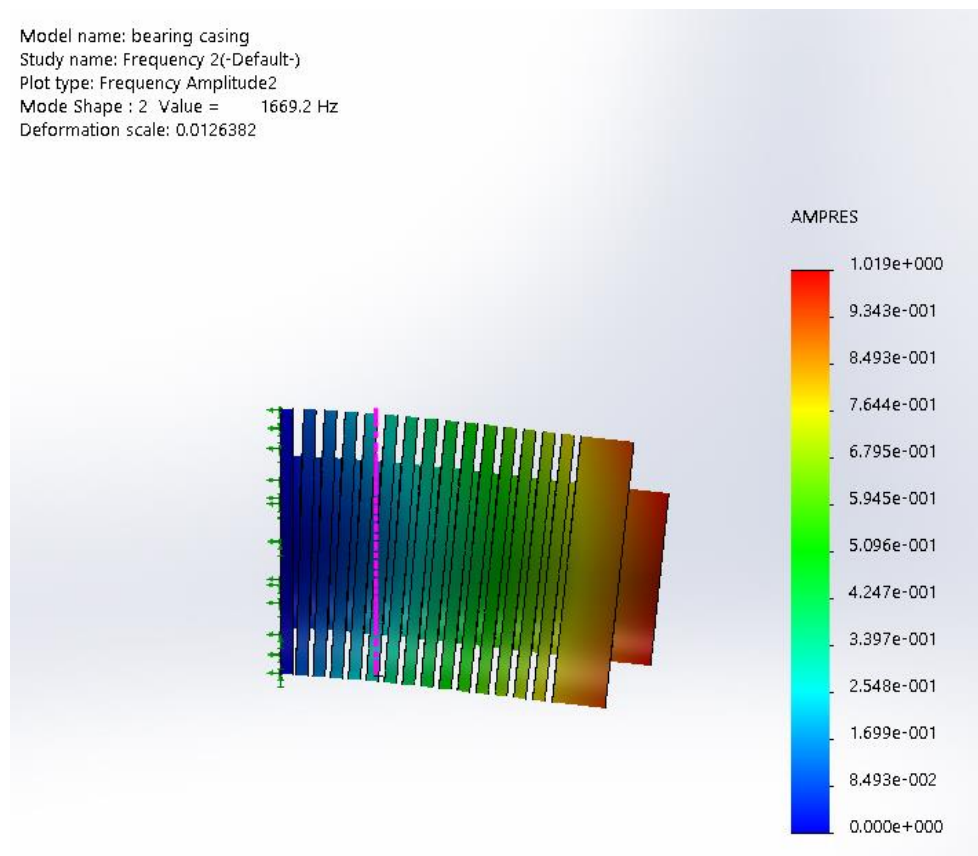
Figure 34: Simulation of the bearing casing second mode shape using SOLIDWORKS software.

4.3 QUANTIFICATION OF AE TRANSMISSION LOSS

AE transmission loss is the power lost when transmitting a signal from one point to another through a medium. Hence, the transmission loss percentage has to be quantified in order to calculate the actual power of the AE signals generated, and to select the single sensor with the highest level of fault readings in both bearings (A) and (B).

A test was carried out using a hollow aluminium tube, pin and steel ball as shown in Figure 35. The ball was set at a height of 25 mm from the bottom of the aluminium tube. Then, the pin was pulled out to allow the ball to hit the surface. The AE signal generated was measured by the R6a AE sensor, which was connected to the NI DAQ system.
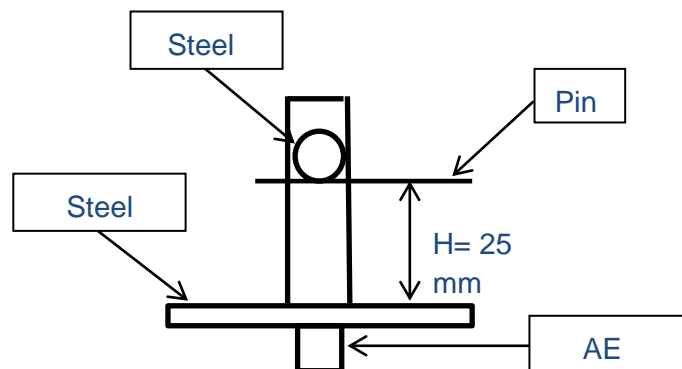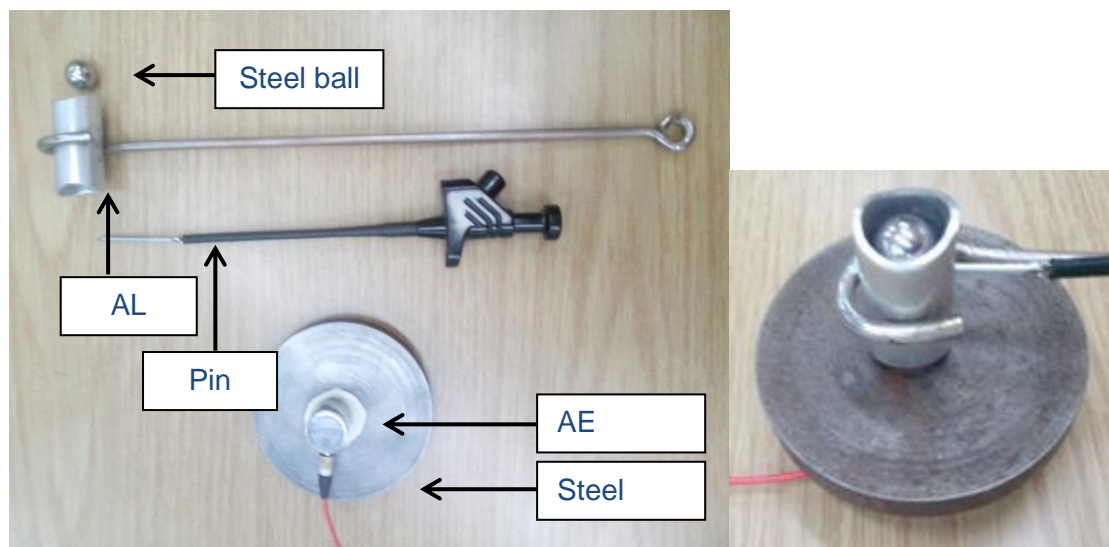
Figure 35: AE transmission loss test using a ball setup.

Five different tests were carried out and compared in order to calculate the AE transmission loss percentages, as shown in Table 12. The ball setup was placed in three different positions: the first was above the steel plate shown in Figure 35 (Benchmark), the second was above bearing (A), and the third was above bearing (B).

Figure 36 shows the results of tests numbered four and two where the loss percentages in AE signals were 10.68% and 4.48%, respectively. The transmission loss factors of test 3 and 4 were high for the following reasons:

Since bearing (B) had an upper flange and bearing (A) located very close to the impeller, AE sensors were not positioned exactly above the bearing locations.

The AE sensors were positioned above a grooved bearing case. Although these grooves were filled with silicon, air gaps may have been present. These air gaps strongly attenuated the strength of AE signals.

AE signal at Bearing (A) when the ball setup was placed at bearing (B)	AE signal at Bearing (A) when the ball setup was placed at bearing (A)
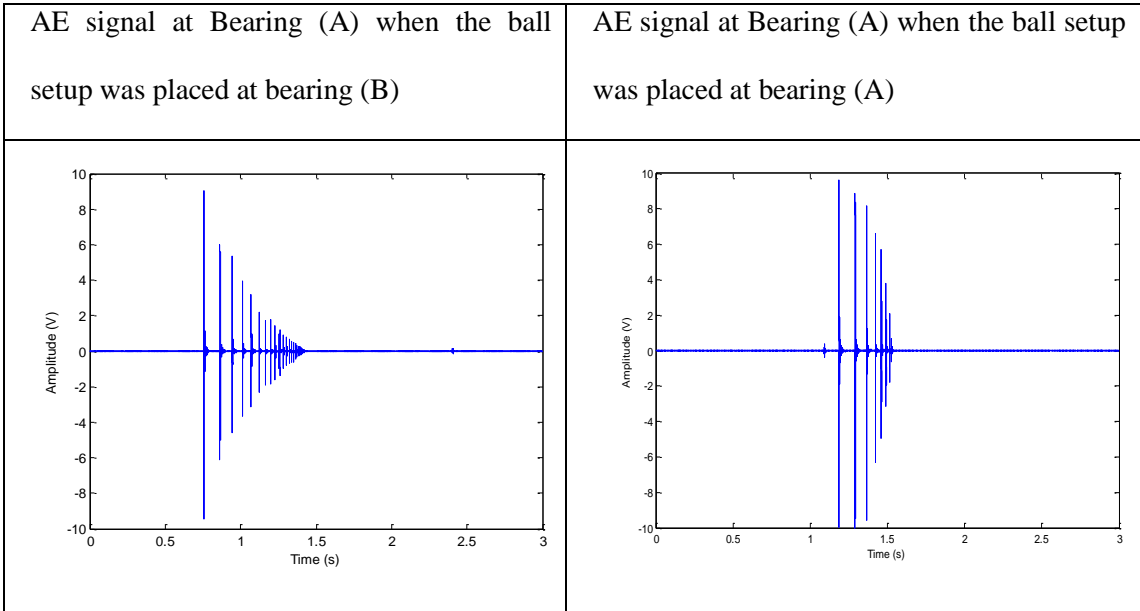
Figure 36: The AE signals acquired during the impact test.

Table 12 illustrates the results of the transmission loss experiment and shows the difference in AE amplitude values when the ball impacted the surface first at bearing (B), and then at bearing (A). The signal at bearing (A) was attenuated by 10.68% when the ball impacted the surface at bearing (B). Hence, as 10.68% is the lowest peak percentage loss, the AE R6a sensor which was positioned above bearing (A) was selected for its accuracy and sensitivity.

Table 9: Summary of the AE transmission loss results.

Test No.:	R6a Sensor/Signal location	Max. amplitude (V)	Difference (V)	Loss (%)
1	AE sensor/ Steel plate (Benchmark)	20.5	--	0%
2	Bearing (A)/ Bearing (A)	19.58	0.92	4.48%
3	Bearing (B)/ Bearing (A)	16.55	3.95	19.26%
4	Bearing (A)/ Bearing (B)	18.31	2.19	10.68%
5	Bearing (B)/ Bearing (B)	20.02	0.48	2.34%

It was also observed that the amplifier has a zero error value ranging from -0.04 to 0.04 V. The zero error value and transmission loss percentage must be considered if the original AE signal strength is required.

4.4 FREQUENCY RESPONSE AND CALIBRATION OF THE MEASUREMENT SYSTEM

All the components of the system employed were factory calibrated. In addition, a frequency response test was carried out to ensure that the antialiasing filter works properly as well as to quantify the attenuation values at different frequencies. A function generator was utilised to generate sinusoidal signals with fixed amplitude of 5 V and a frequency ranging from 10 kHz to 600 kHz. These signals were utilised as inputs to the measurement system (low pass filters and DAQ system). The difference between the raw FFT of the input and output signals were compared and the attenuation was calculated.

As the frequency range of the AE signals was determined as between 2 and 122 kHz, the cut-off frequency of the low pass filter was set to 200 kHz. Although the amplitudes at frequencies higher than 200 kHz were small, low-pass filters were integrated into the system to attenuate any possible or sudden increase in the amplitudes at any frequencies beyond 200 kHz during experimentation.

Figure 37 shows the frequency response of the measurement system, and illustrates the effect of the low-pass filter which was utilised to combat the problem of aliasing by attenuating frequencies above 200 kHz. As a result of the value of the cutoff frequency, the attenuation percentage went up from approximately 50% to 98% in the frequency range of 150 kHz to 400 kHz, and consequently the proper functionality of the measurement system was proved. The results shown in Figure 37 confirms that the filter does significantly attenuate the signal at frequencies beyond 200 kHz and hence should be no issues of aliasing with the measurement system at sampling rate of 1 MS/s. The attenuation percentages shown should be considered by researchers if the original AE signal strength is required.

Frequency response curve
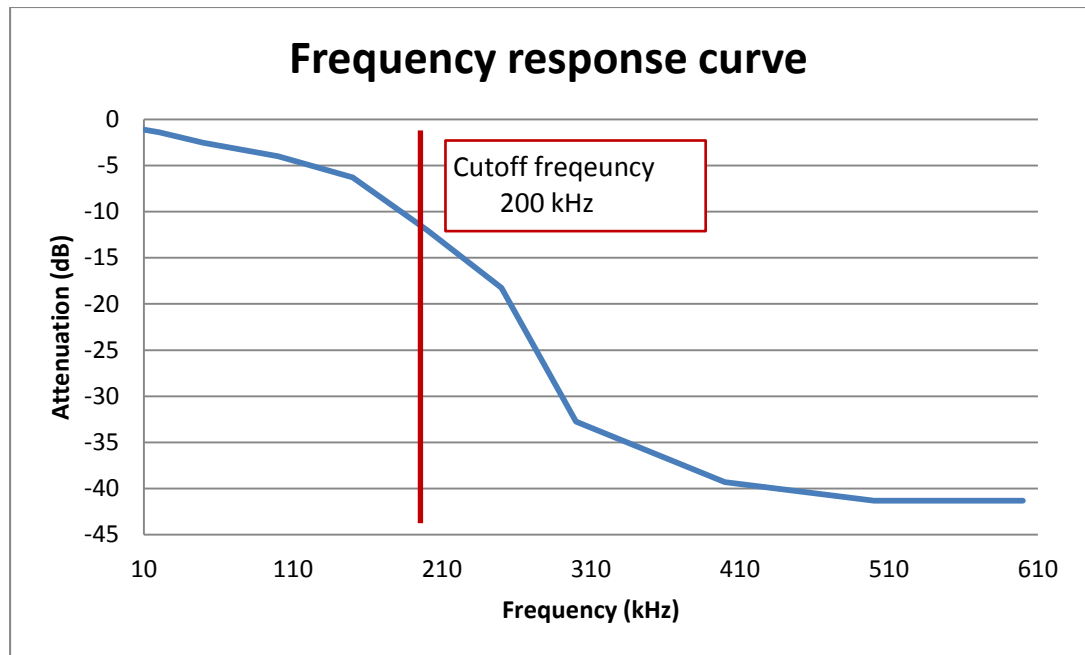
Figure 37: Magnitude frequency response curve at a filter cut-off frequency of 200 kHz.

4.5 SUMMARY

In this chapter, an experimental setup was designed in order to facilitate experiments relating to the development and testing of monitoring algorithms for high-speed centrifugal equipment. A number of experiments were conducted to decide where best to place the AE (at bearing (A) or bearing (B)), and the vibration and pressure sensors. It was found that the AE sensor received a stronger signal if installed at bearing (A), that the accelerometer could be positioned at any location, and that the pressure sensor worked properly when installed 50 cm away from the compressor outlet.

Experiments were conducted to calculate the frequency response of both the data acquisition system and antialiasing filters. The resulting "loss" and "attenuation" factors were determined to allow correction of the measured signal amplitudes, if required.

Chapter 5. Development of acoustics-based condition monitoring system

5.1 INTRODUCTION

LNG production trains availability can be improved by implementing a standby system (see chapter 3), or by using an effective CBM system in order to reduce the number of planned shutdowns [110] [49]. Effective CBM systems can reduce the overall cost of maintenance and allow the early detection of faults which usually lead to catastrophic consequences and are extremely expensive to repair. Unlike preventive and corrective maintenance strategies, the implementation of CBM sharply lowers the overall maintenance cost by preventing major failures, and by delaying scheduled maintenance until a more convenient time or until it becomes necessary.

The performance of CBM systems is affected by the performance of the CM system in use. The performance of CM systems is affected by the fault detection technique and feature selection algorithm employed. The effectiveness of acoustic emission-based fault detection techniques has been demonstrated, especially with respect to the early detection of machine faults and the identification of crack size (section 2.5.3.2). Feature selection algorithms are based on various feature types such as time domain features and FFT features. FFT features provide useful information on rotating components since well-defined frequency components are associated with machine condition [80].

In a large number of recent studies, the amplitude feature of FFT spectra was considered as the most common and well-proven feature for the identification of faults in rotating

equipment [66] [121] [122] [67] [68]. The existing non-AI and AI feature selection algorithms have a number of drawbacks which adversely affect the cost of implementation and generalisation of these algorithms to all systems (section 2.5.4). Existing non-AI feature selection techniques cannot handle redundant feature identification, and this has an adverse effect on the learning process and accuracy of detection [100]. Existing AI feature selection techniques and classifiers are case-dependent and complex, particularly when compared with systematic non-AI approaches; this can be inferred from the research work related to feature selection. AI approaches require significant research to provide near-optimum accurate results [80] as the accuracy of AI approaches is highly dependent on the architecture of the network and changes in the various design parameters such as number of training cycles, number of neurons, initial conditions, number of membership functions, and number of rules. The development costs include the cost of computing and of hiring skilled CM system developers experienced in AI approaches and rule-based knowledge. The development time includes the time needed to develop, train and validate custom FFT feature selection and classification systems based on artificial intelligence approaches [80] [99] [91] [92].

Since it is difficult to visually select the most informative FFT features for a very large number of datasets, and in order to overcome some of the disadvantages of existing feature selection techniques, a robust, self-learning and automated feature selection and classification algorithm is proposed for the development of effective CM systems. An FFT feature selection and classification tool which is flexible, systematic, automated, and non-AI-based could significantly decrease the cost and time needed for the development of efficient CM systems in comparison to the development cost and time of existing FFT-based feature selection classifiers.

LNG production lines use several items of centrifugal equipment such as LNG compressors, gas turbines and blowers. Bearings are essential components of all centrifugal equipment. Bearing faults occur due to fatigue (even under normal and balanced operations), improper lubrication, contamination or installation errors. Major bearing fault identification systems are usually based on vibration or acoustic signatures, both of which tend to increase in line with bearing deterioration. Major vibration and acoustic signal features consist of RMS, crest factor, energy, counts and peaks, amplitude, and frequency response via Fast Fourier Transform (FFT) [7].

In the light of the literature review (chapter 2), the AE technique was judged to be very effective in detecting rotating machine faults. Unlike the vibration technique, the AE technique is less affected by noise emanating from other nearby machines or by structural vibration. For example, it is difficult to use vibration spectral features of faulty bearings for the identification of their faults as the resonance frequencies of the structures between the bearings and the transducers will be excited, and consequently might change the vibration signatures of the faults. Thus, most informative bearing fault signatures occur at high-frequency resonance bands [123]. Hence, AE spectral features will be utilised here to verify and validate the proposed algorithm. Other CM techniques and signal features will also be utilised in chapter 7 for the assessment, comparison and characterisation of different fault identification techniques used in the CM of typical centrifugal equipment.

In summary, significant research has been undertaken with a view to developing and implementing efficient automated machinery fault identification and diagnostic tools. The feature selection algorithms of the majority of existing techniques are based on non-automated scheme or on AI approaches (see section 2.5.4). The existing AI-based feature selection and fault identification techniques are complex, computationally intensive and require long development times. Therefore, the main purpose of this chapter is to develop

an automated, robust, fast, easy-to-implement, and non-AI AE-based fault identification algorithm that overcomes some of these disadvantages.

5.2 DESIGN OF EXPERIMENT

Bearing problems account for over 40% of machine breakdowns. Thus, this experimental work focused on bearing faults in centrifugal equipment. Typical causes of bearing faults in all rotating equipment are: excessive load, overheating, false brinelling, true brinelling, normal fatigue failure, reverse loading, contamination, lubricant failure, corrosion, misalignment, loose fits, and tight fits [124]. Two common bearing failures were selected to evaluate the proposed fault detection algorithm; true brinelling and normal fatigue failures. Brinelling occurs when loads exceed the elastic limit of the ring material. Brinell marks show as indentations in the raceways, and these indentations increase bearing vibration and noise. Severe brinell marks can cause premature fatigue failure. Fatigue failure, usually referred to as spalling, is a fracture of the running surfaces and subsequent removal of small discrete particles of material. Spalling can occur on the inner ring, outer ring, or balls [124] [125] . Hence, indentations (small holes) and notches were created in bearings in order to emulate the most common bearing failures in rotating equipment.

Using the test rig detailed in Chapter 4, nine machine conditions were experimentally emulated in the laboratory at rotational speeds of 3600, 6960 and 15650 RPM.

Figure 39 depicts the machine conditions emulated. 81 experimental datasets were utilised to train, verify and test the proposed CM system at the three rotational speeds; there were three datasets for each machine condition at each rotational speed (3x9x3=81 datasets). Machine Condition 1 (MC 1) refers to the healthy condition, while others refer to the faulty conditions. These three rotational speeds were selected in order to emulate normal low, normal high, and high-speed centrifugal equipment. The changes in

signatures resulting from to speed changes will help better understand the behaviour of faults at these different rotational speeds.

Figure 38 illustrates the machine faults emulated at bearings (A) and (B). Bearing (A) has a 3 mm hole throughout the outer race, while bearing (B) has four notches, two on each side with a maximum groove width of 2 mm. This hole was drilled using a 3 mm carbide drill bit, and the notches were engraved using a Dremal shank diamond taper point bit. The machine condition faults shown in Figure 39 are a combination of partial or full bearing lubrication removal and outer race defects.
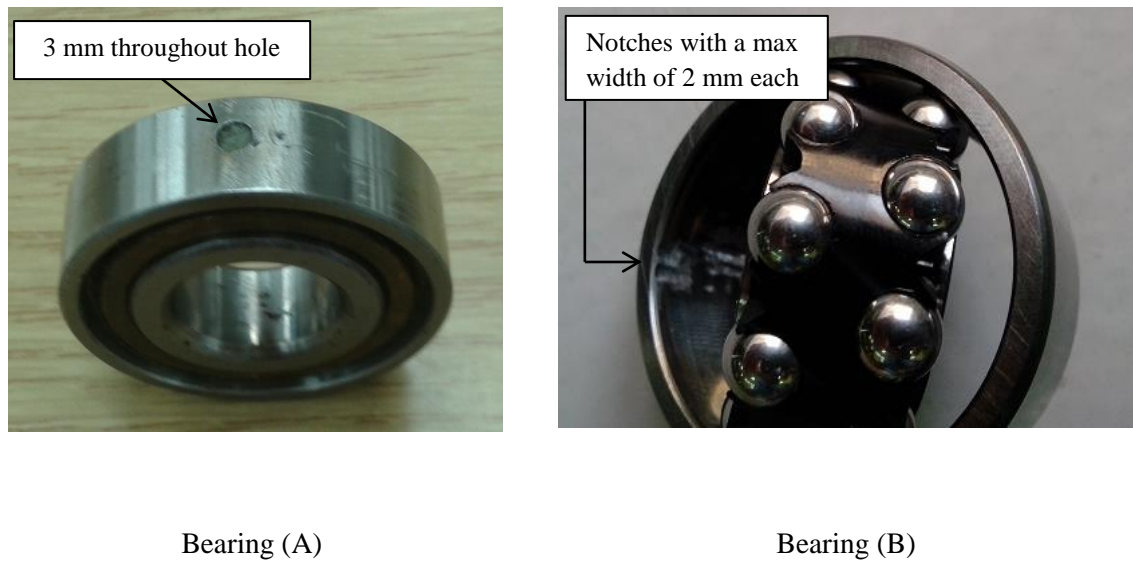
Bearing (A)   Bearing (B)

Figure 38: Notches in the outer races of bearings (A) & (B).
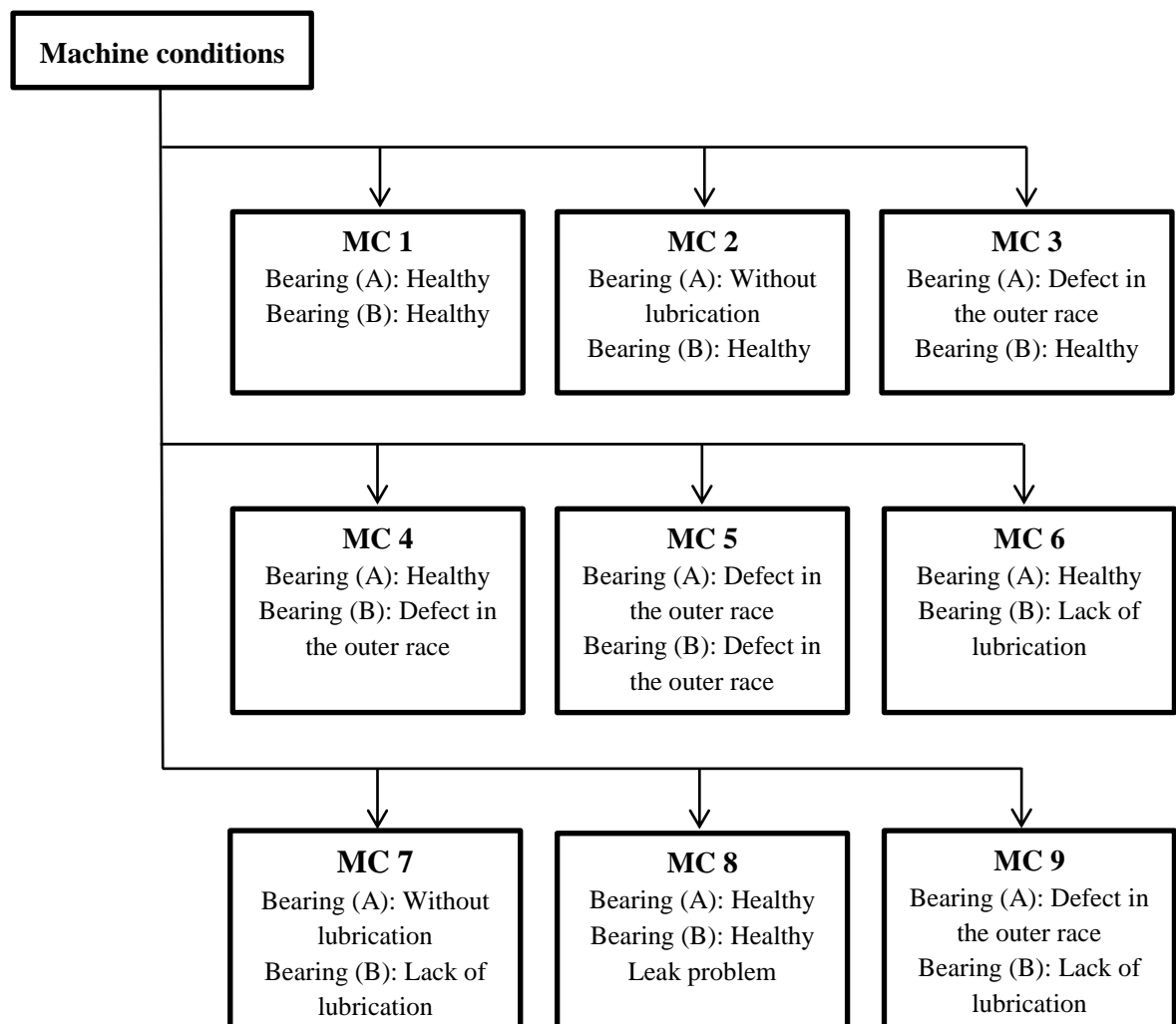
Figure 39: Emulated machine conditions.

The data were collected over a sampling duration of one second using a high speed NI DAQ board at a sampling rate of 1 MS/s with a total number of 81 AE datasets. Three datasets were collected at a fixed time interval of 5 seconds for each machine condition, one set every 5 seconds. The first dataset for each experiment was recorded 30 seconds after the blower reached its full rotational speed. For each machine condition, one dataset was utilised for training, while the remaining two were used for verification and testing.

According to the Nyquist Theorem, the sampling rate must be twice as fast as the highest frequency of the measured signal. However, sampling at exactly two times the highest frequency is often unacceptable, especially in applications where the shape (time-domain representation) of the signal is important. Thus, the data were sampled at a rate of 1MHz to allow accurate measurement of signals up to 200 kHz AE. This sampling rate is five times higher than the target frequency. Based on the sampling rate and period, the size of each dataset is 1×10^6 samples.

5.3 DEVELOPMENT OF FFT-BASED SEGMENTATION, FEATURE SELECTION AND FAULT IDENTIFICATION (FS2FI) ALGORITHM

The spectral-feature selection process is a key process for FFT-based AE and vibration condition monitoring and maintenance systems. Due to the large number of machine fault patterns, and in order to overcome the disadvantages of the existing CM techniques summarised in the introduction of this chapter, a segmentation, automated feature-selection and fault identification FFT based algorithm is proposed. The total number of machine conditions (NoC) utilised to test the performance of the algorithm is nine. Based on the literature survey (chapter 2), the proposed algorithm utilises the well-proven amplitude feature of the FFT spectrum, and AE as a well-proven fault detection technique. The main objective of the FS2FI algorithm is to segment the FFT spectra of all benchmark signals and to identify the segment size at which all machine fault patterns are

identifiable. The FS2FI algorithm starts by computing the Fast Fourier Transform (FFT) of the discrete benchmark machine condition signals using equation (4) [126]. Then, it breaks down the frequency spectrum of each signal into smaller groups of frequencies (segments). The segment size depends on the number of feature differences required to clearly differentiate all spectra of all machine conditions. The FS2FI algorithm shown in Figure 41 managed to select the spectral features required to differentiate all the machine fault patterns addressed.

$$X_k = \sum_{n=0}^{N-1} X_n e^{-i2\Pi K \frac{n}{N}} \qquad (4)$$

where

X_k = Transform values [Amplitude and phase, a complex number]
K   = Current frequency [from 0 to (N-1) Hz]
X_n = Sample values
n = Number of sample

Figure 40 illustrates a simplified flowchart that describes the main steps for the processing and analysis of the data. The proposed algorithm starts by processing training machine fault signatures (benchmark signatures). All time domain signals are converted into FFT spectra, and then segmented into a number of equal segments (Nd). The maximum amplitude value for each segment of each signal is selected as an FFT feature, and compared with the corresponding maximum amplitude values for other signals. If there is a minimum difference of 10% between the maximum amplitude values for signals at all machine conditions, the algorithm produces Nd as a possible solution. Otherwise, the algorithm iteratively increases the number of segments (Nd) using a for-loop until all fault patterns are successfully differentiated.
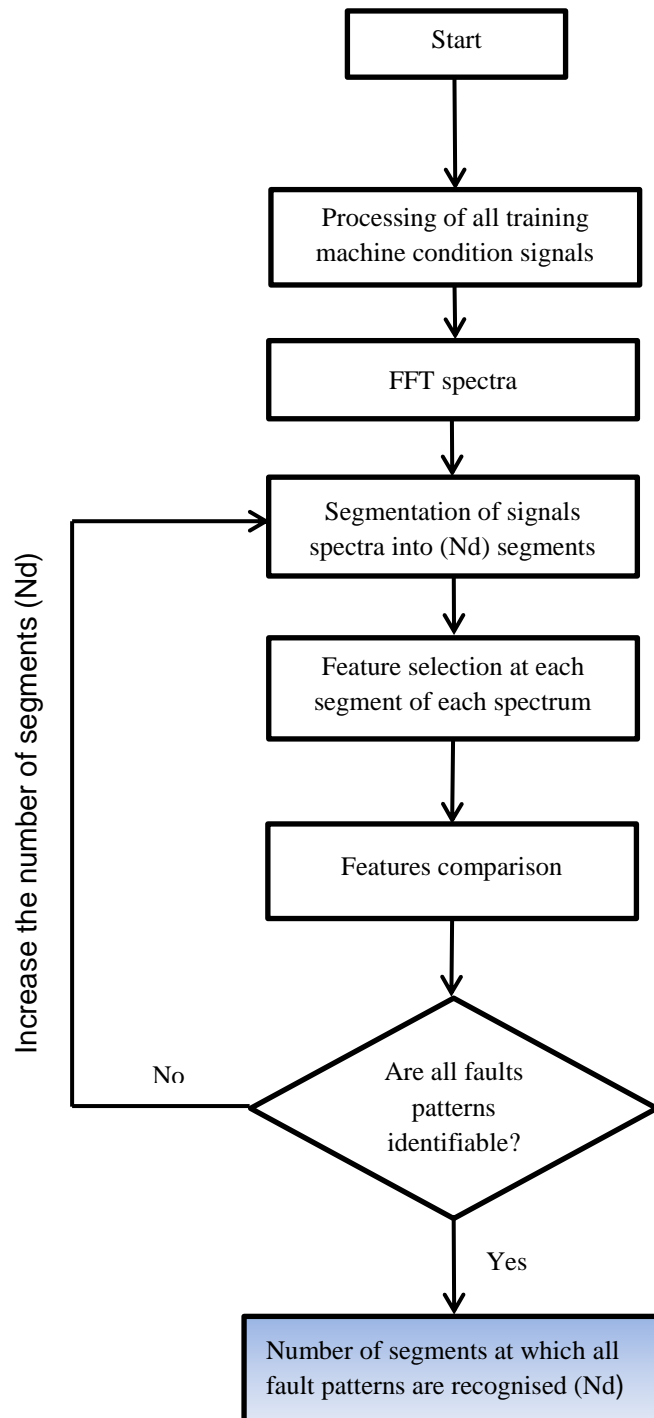
Figure 40: Illustrative flowchart of the proposed FS2FI algorithm.

Figure 41 shows the process of the FS2FI algorithm in detail. The algorithm consists of four phases: in the first phase, the program loads all the recorded datasets, transforms the datasets from time to frequency domain using the MATLAB FFT function, and then divides the frequency range into two or more equal segments (Nd). A sample of

amplitude matrix (A) is shown in Table 13. The bin size was set to 1 Hz in order to maximise the resolution of FFT spectra (considering all component frequencies). This was carried out by using the fft() MATLAB function to set the FFT length as equal to the length of the signals (each of which contains 1 million samples).

In the second phase, the program calculates the peak FFT amplitude value at each segment, and then calculates what is known as the "difference matrix" of size $(NoC_T * NoC_T, Nd)$. NoC_T is the total number of machine conditions. The terms in this matrix are calculated by comparing the peak values of each fault spectrum with all the corresponding peak values of the other fault spectra (process A). If there are no intersections in a +/- 10% range of the peak values of the amplitudes of a spectrum, and the corresponding peak amplitudes of the other spectra, a value of 0 is inserted in the difference matrix at the corresponding row and column. This 10% range was selected to consider the median change in the measured values if the R6a sensor is removed and remounted. This percentage should not be less than the median remounting uncertainty of the R6a sensor, which is 5.4% [127].

In the third phase, the algorithm analyses the difference matrix to produce a Boolean Decision matrix of size (NoC_T, NoC_T) that can output the recommended segment size for successful pattern classification. Then, it sums the elements of each difference matrix row, converts the summed values into "zero" or "1", and then inserts the results at the corresponding element into what is a "decision matrix" (process B). A zero value in the decision matrix indicates that there is an intersection between peak amplitudes, and that the pattern is not recognised. Hence, to recognise all patterns, all decision matrix elements have to be equal to 1, except for the diagonal elements which should be equal to zero. A sample of a decision matrix is shown in Table 14.

In the last phase, the algorithm analyses the decision matrix to check whether or not all patterns have been successfully recognised. If not, the Nd variable will be increased by 1, and thus the number of features will increase in an effort to differentiate all fault patterns across a larger number of segments. The last phase ensures that the CPU's computational time is reduced to a minimum as the algorithm stops the iteration process once all patterns have been recognised (process C).

The set of utilised variables and conditions are as follows:

- Nd: Total number of frequency divisions (or segments)
- NoC: Number of machine fault conditions (from 1 to 9)

Process A: this compares all peak values of the benchmark AE frequency spectra together, and stores the results in a matrix called a "difference matrix" (MDifference) using the MATLAB equation (5). If two peak values are equal, or the difference between them is less than the median uncertainty percentage (+/- 10%), a value of 1 is inserted in the difference matrix at the corresponding row and column; this in turn gives an output that this segment cannot be utilised to differentiate between both fault patterns. Otherwise, if they are different, a value of zero is inserted.

$$\mathbf{M}_{Difference}(\mathbf{I} + \mathbf{II}, \mathbf{K}) = \begin{array}{l} \mathbf{1} \; if \; A(II,K) - A(II,K) < (+/-)\mathbf{0.1} * A(II,K) \\ \mathbf{0} \; if \; A(II,K) - A(II,K) >= (+/-) \; \mathbf{0.1} * A(II,K) \end{array} \qquad (5)$$

where

$NoC_T = 9$ (total number of machine conditions), $I = 0:NoC_T:NoC_T*(NoC_T-1)$, $II = 1:1:NoC_T$ and $K = 1:1:Nd$.

Process B: this sums the elements of each row of the difference matrix, converts the summed values into "zero" or "1", and then inserts these values into a matrix called a "decision matrix" using the MATLAB equation (6).

The decision matrix is then rearranged and put into the form shown in Table 14. If the summed value is equal to "zero", this means that the fault pattern is detectable, and a value of "1" is then inserted into the decision matrix. Otherwise, a value of "zero" is inserted.

$$M_{decision}(I + II, NoC) = \begin{array}{l} 1 \ if \ \sum_{K=1}^{Nd} M_{Difference}(I+II,K)=0 \\ 0 \ if \ \sum_{K=1}^{Nd} M_{Difference}(I+II,K)>0 \end{array} \tag{6}$$

Process C: this iteratively increases the number of divisions (Nd), and calculates new difference and decision matrices if any decision matrix element except diagonal elements is equal to "0" , and until all elements are equal to "1". The maximum value of Nd was set to 1000 divisions.
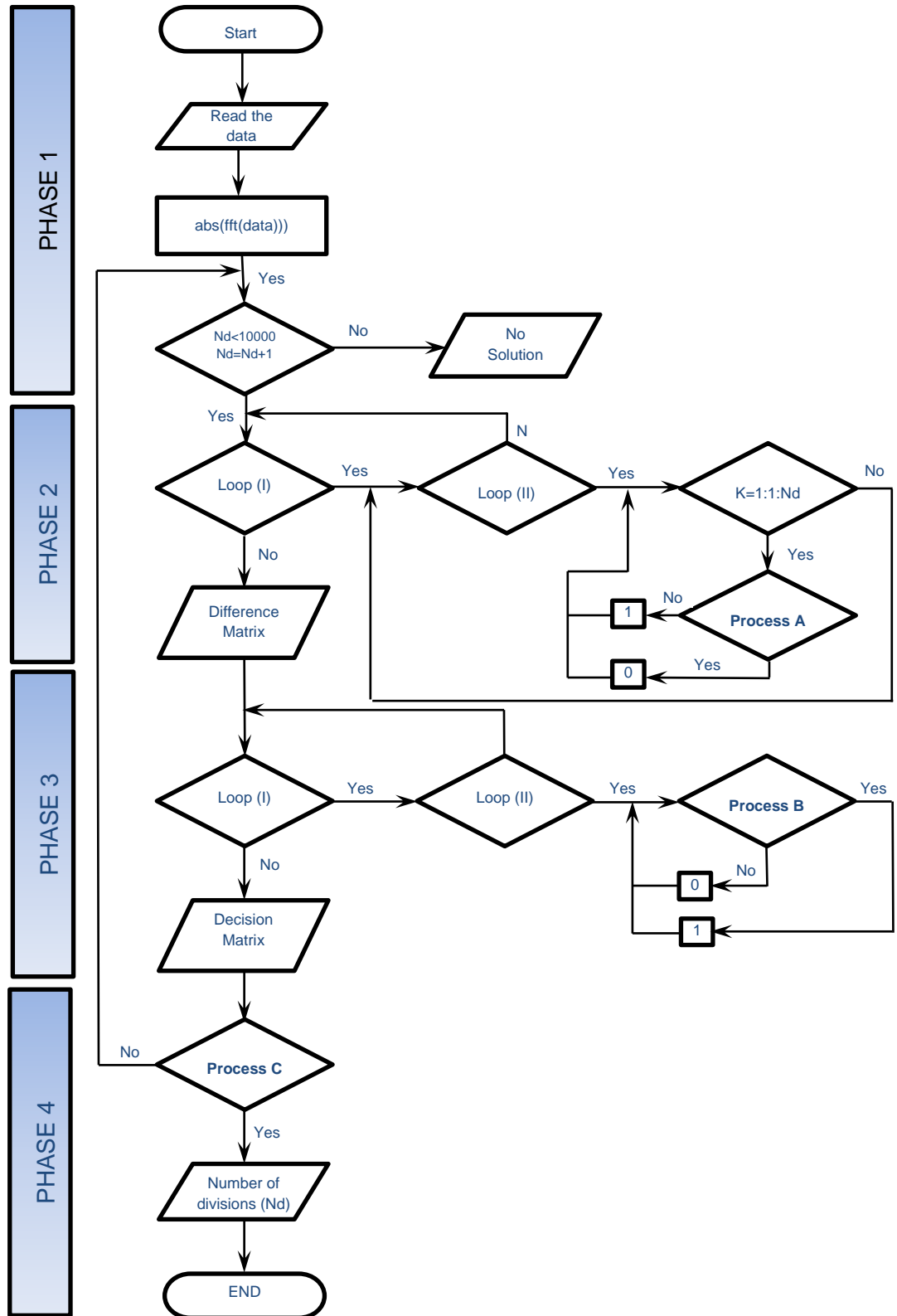
Figure 41: Block diagram of the proposed FS2FI algorithm.

Having followed the previous training process to select the number of divisions, and having set up the benchmark spectral data for each fault condition, a rule-based MATLAB program was developed for online fault detection and identification using the proposed algorithm. The program converts the real-time signal to a FFT spectrum, and then divides the spectrum into a number of segments equal to Nd. It compares the values of peak amplitudes, one peak at each segment, of the online signal with all corresponding peak amplitudes of benchmark spectra. Based on the results of the comparison, the program then produces a matrix that identifies the benchmark machine condition feature which is nearest to the online signal feature at each segment. The machine condition with the largest number of matching features is then identified as a possible solution. The performance of the FS2FI algorithm is evaluated in Section 5.5.

5.4 EXPERIMENTAL RESULTS

Nine datasets were collected to train the proposed algorithm at each of the three different rotational speeds (3600 RPM, 6960 RPM and 15650 RPM), one dataset for each machine condition. Eighteen different datasets were collected for verification and testing, two datasets for each machine condition. The output of the algorithm proposed in the previous section is that 6 segments (Nd=6) with 6 different features, one feature at each segment, will be sufficient to differentiate between all of the machine conditions addressed, at all rotational speeds.

5.4.1 AE FAULT SIGNATURES AT 3600 RPM.

Figure 42 shows the AE fault signatures measured at bearing (A) using the R6a AE sensor at 3600 RPM. Figure 42 (a) indicates the benchmark AE signature of the healthy condition, while Figure 42 (b) to Figure 42 (i) show the other eight AE signatures of the fault conditions addressed. The shapes of the AE signatures are different, and faults can be easily differentiated. The spectra below were utilised to train the proposed algorithm.
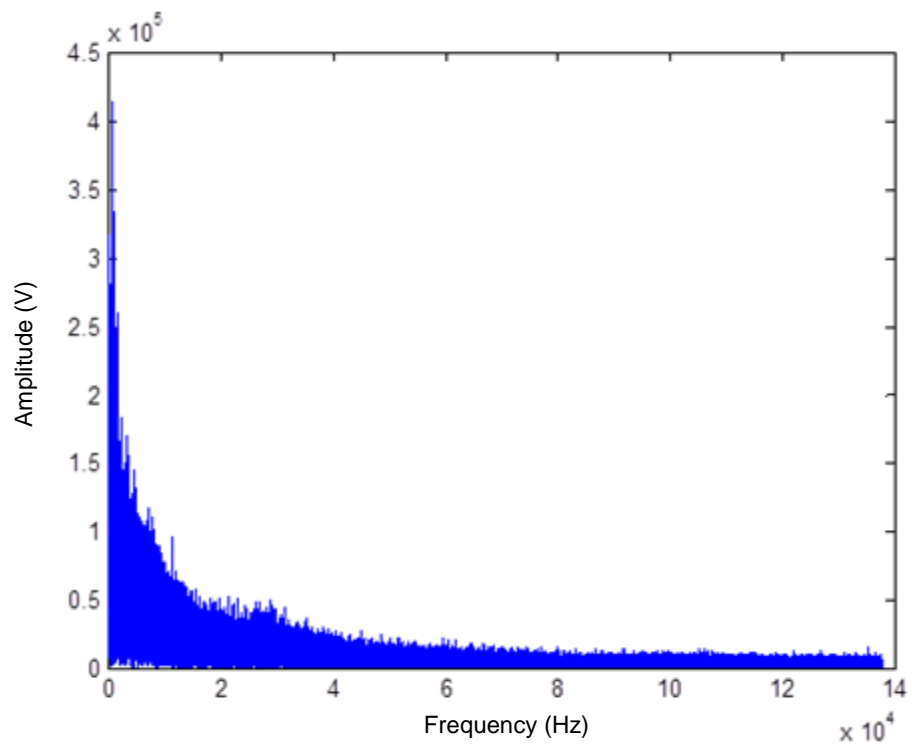
Figure 42 (a): MC 1
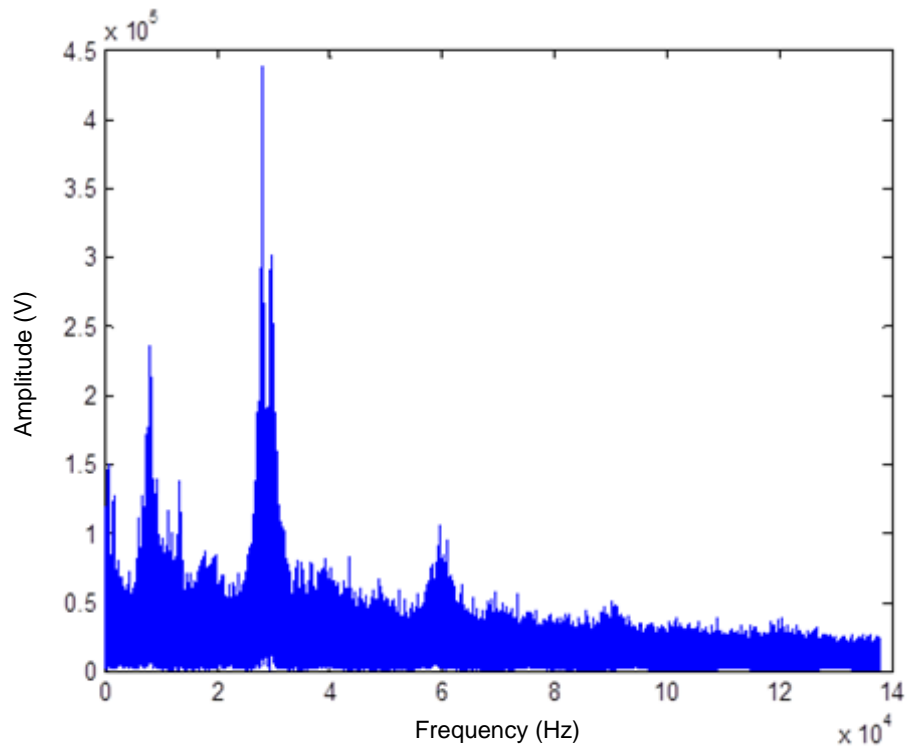
Figure 42 (b): MC 2
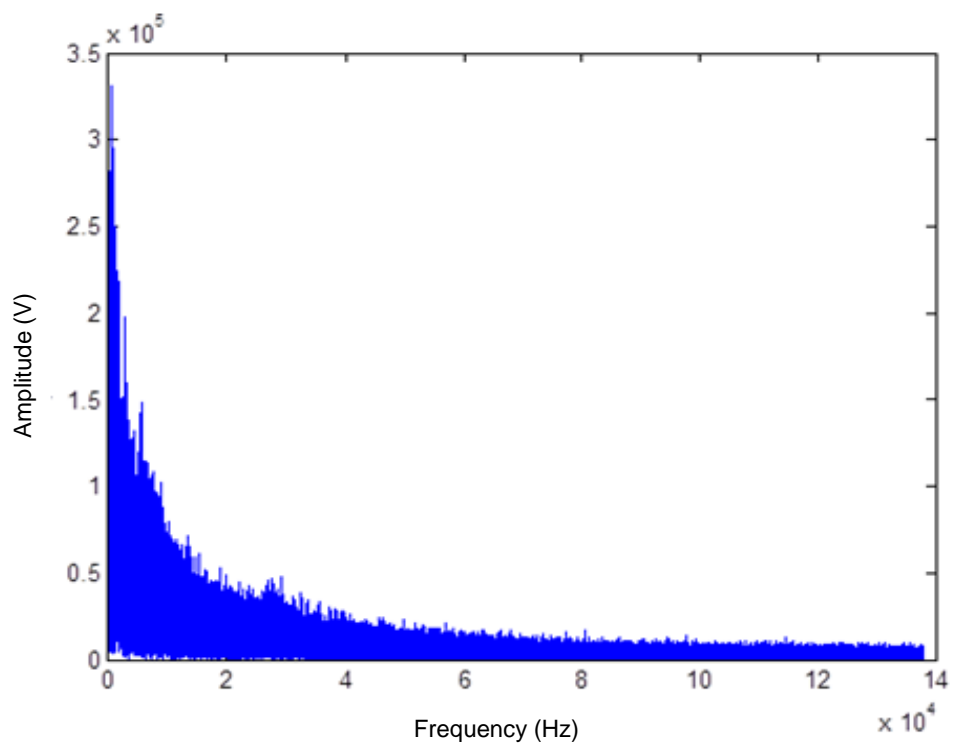
Figure 42 (c): MC 3

Figure 42 (d): MC 4

121
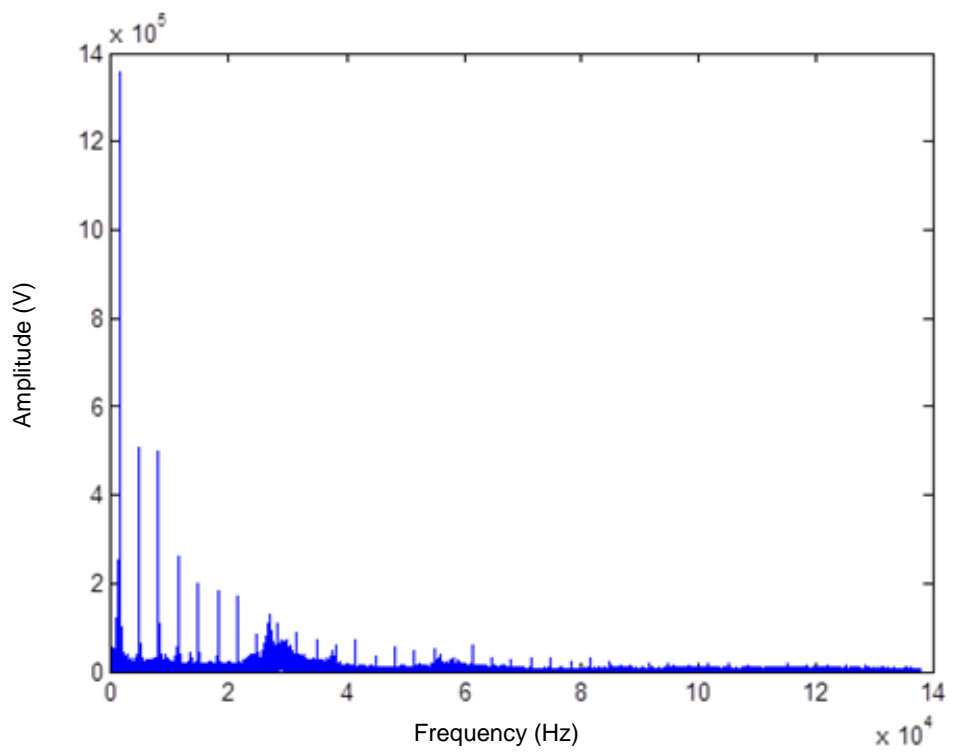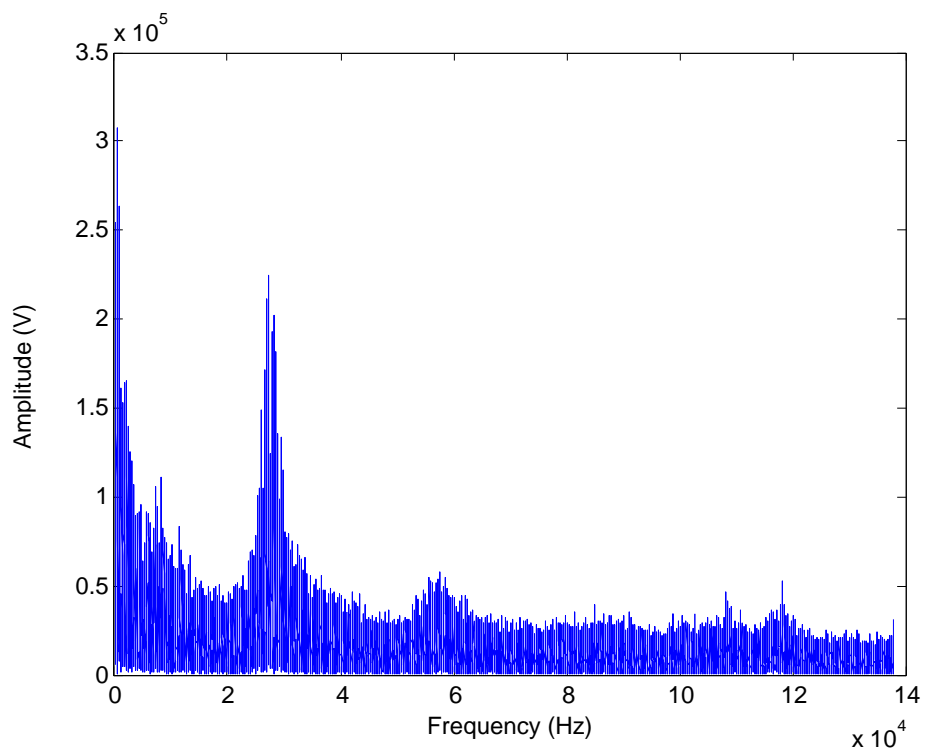
Figure 42 (e): MC 5
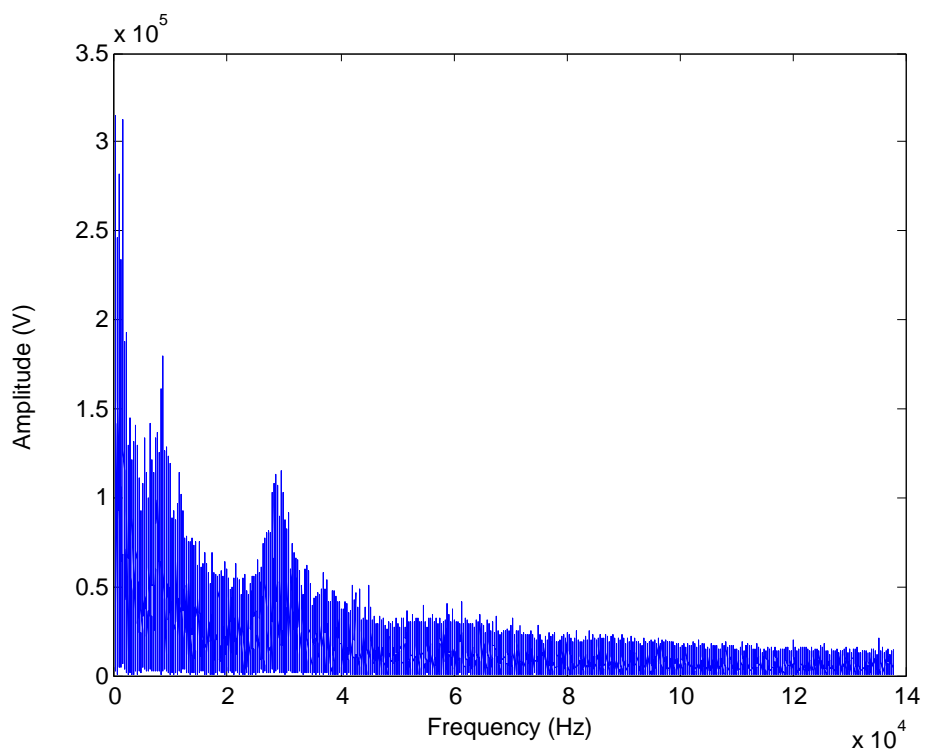
Figure 42 (f): MC 6
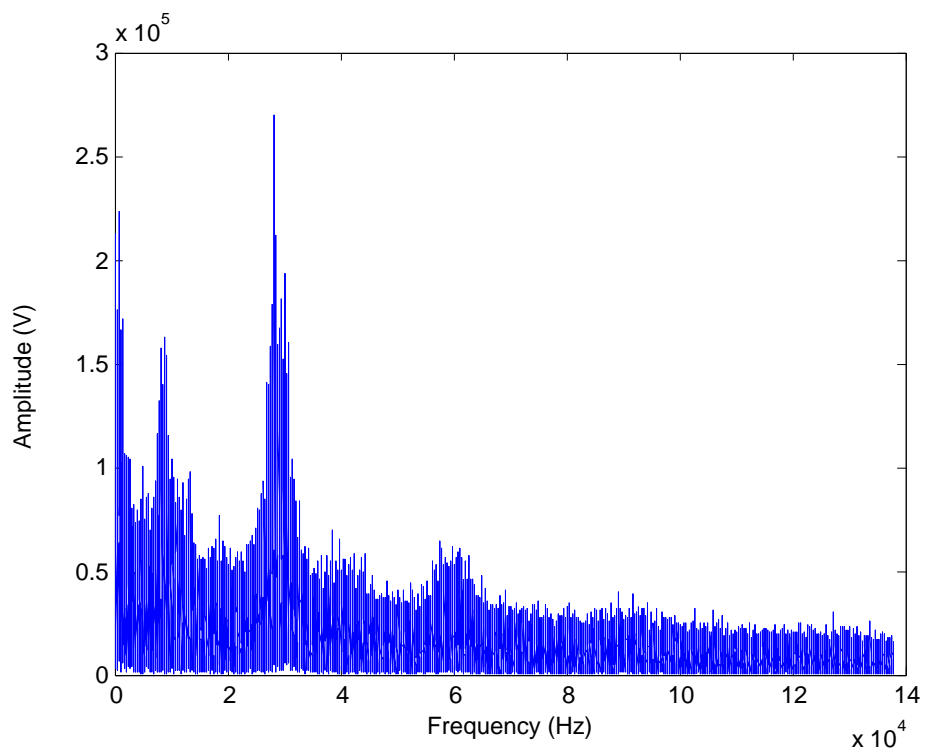
Figure 42 (g): MC #7
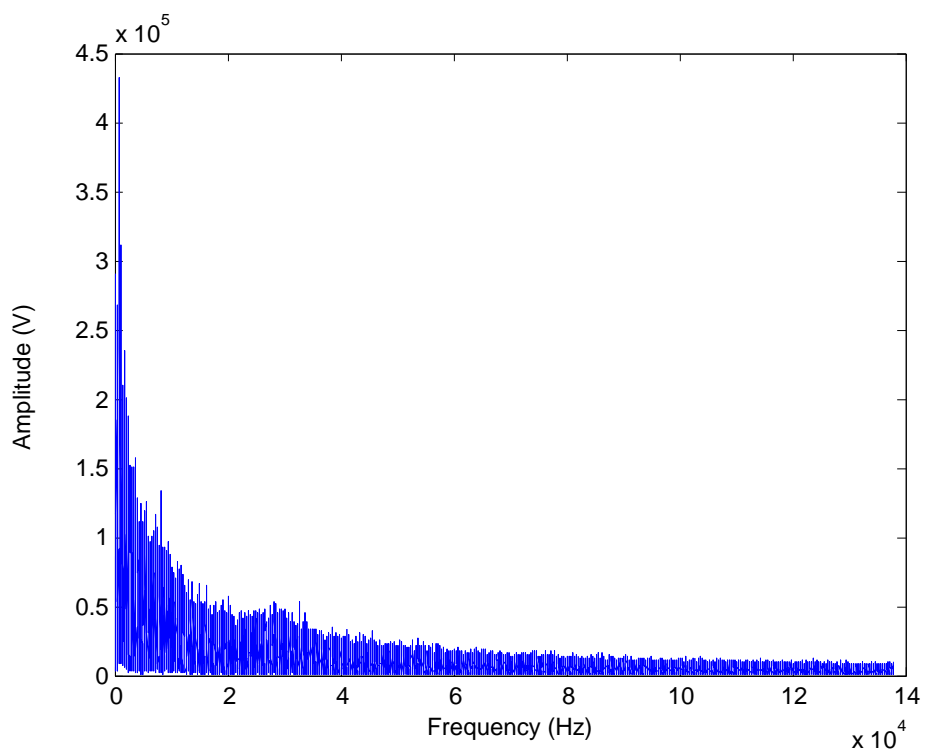
Figure 42 (h): MC 8
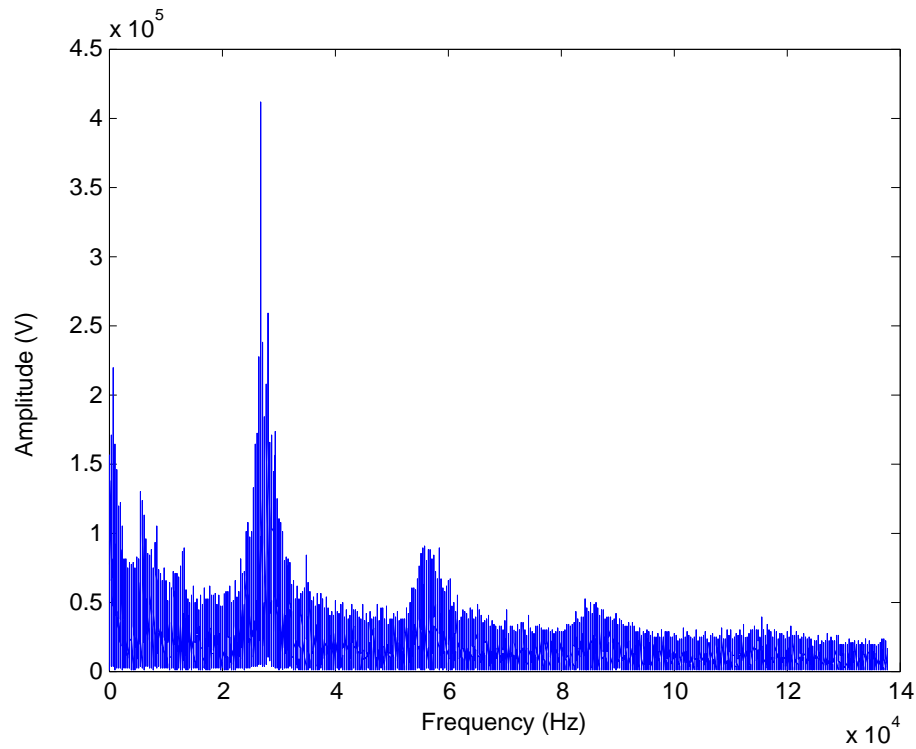
Figure 42 (i): MC 9

Figure 42: AE frequency spectra of the machine conditions addressed (2 - 140 kHz) at

3600 RPM.

5.4.1.1 THE MAIN FEATURES OF THE DIFFERENT FAULT CONDITIONS

Table 13 summarises the main features of the training FFT spectra shown in Figure 42.

These features were identified using the algorithm developed. The algorithm split the

frequency range (2 kHz to 122 kHz) into six equal segments and identified the peak

amplitude in each segment of each machine condition.

Table 10: Peak FFT amplitudes of all machine fault signatures at 3600 RPM.

Freq. (Hz)	2000-21999	22000-41999	42000-61999	62000-81999	82000-101999	102000-121999
MC 1	413,978	51,359	28,102	21,975	15,805	13,551
MC 2	236,606	437,280	105,488	95,593	50,017	39,043
MC 3	330,871	48,741	26,677	17,869	17,718	12,342
MC 4	1,355,991	169,398	73,699	62,552	34,691	17,713
MC 5	307,854	225,133	58,208	44,809	39,071	52,557
MC 6	314,974	115,601	51,088	41,403	25,244	19,937
MC 7	223,586	269,910	64,691	60,958	39,971	32,279
MC 8	432,541	57,189	33,496	20,335	16,255	13,810
MC 9	218,584	411,400	90,726	66,246	51,989	38,669

The algorithm compared all the peak segment amplitudes with their corresponding benchmark values, and then placed the outputs into the decision matrix shown in Table 14. The benchmark signals (MC 1 to MC 9) are shown in rows, and the validation signals (MC 1 to MC 9) are shown in columns. If the algorithm was not able to differentiate between two fault patterns, it set the corresponding matrix element to "0". Otherwise it set it to "1". For example, if MC 1 and MC 2 had been differentiable, the values of Decision matrix elements (1,2) and (2,1) would have been set to "1". Otherwise, they would have been set to "0". As can be seen, all machine conditions were successfully identified at this segment number (6 segments).

Table 11: Decision matrix ($M_{Decision}$). The benchmark signals (MC 1 to MC 9) are shown in rows, and the validation signals (MC 1 to MC 9) are shown in columns.

	MC 1	MC 2	MC 3	MC 4	MC 5	MC 6	MC 7	MC 8	MC 9
MC 1	0	1	1	1	1	1	1	1	1
MC 2	1	0	1	1	1	1	1	1	1
MC 3	1	1	0	1	1	1	1	1	1
MC 4	1	1	1	0	1	1	1	1	1
MC 5	1	1	1	1	0	1	1	1	1
MC 6	1	1	1	1	1	0	1	1	1
MC 7	1	1	1	1	1	1	0	1	1
MC 8	1	1	1	1	1	1	1	0	1
MC 9	1	1	1	1	1	1	1	1	0

5.4.2 AE FAULT SIGNATURES AT 6960 RPM

Figure 43 shows the AE signatures measured at bearing (A) using the R6a AE sensor at 6960RPM. Figure 43 (a) indicates the benchmark signature of the healthy condition (MC 1) while Figure 43 (b) to Figure 43 (i) show the signatures of the fault conditions (MC 2 to MC 9). The shapes of AE signatures are different and faults can be easily identified. The spectra below (benchmark fault signatures) were utilised to train the proposed algorithm.
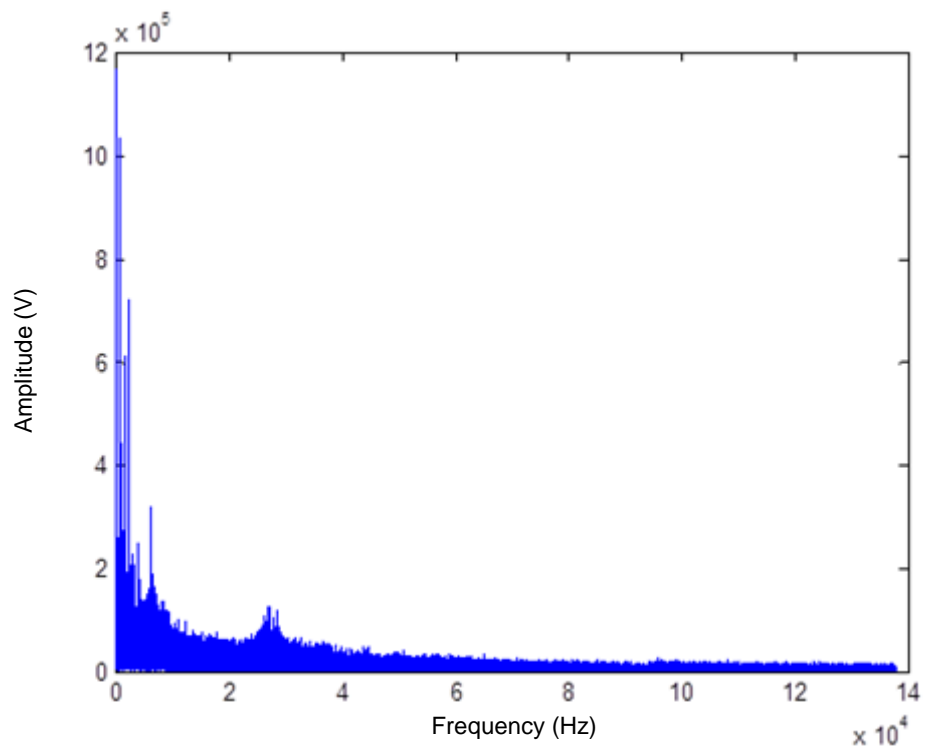
Figure 43 (a): MC 1
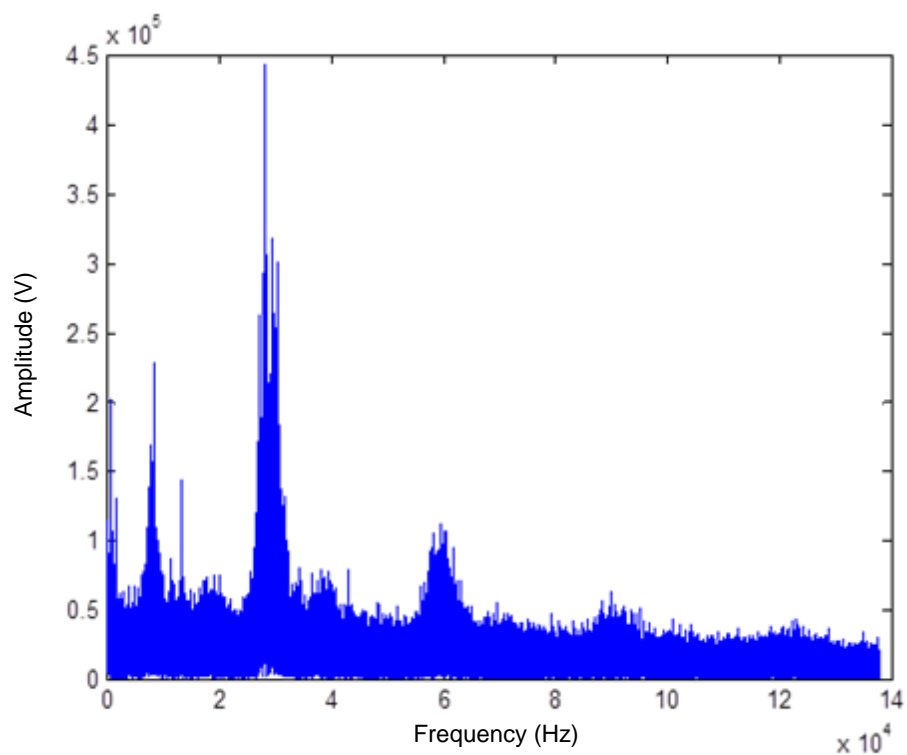
Figure 43 (b): MC 2
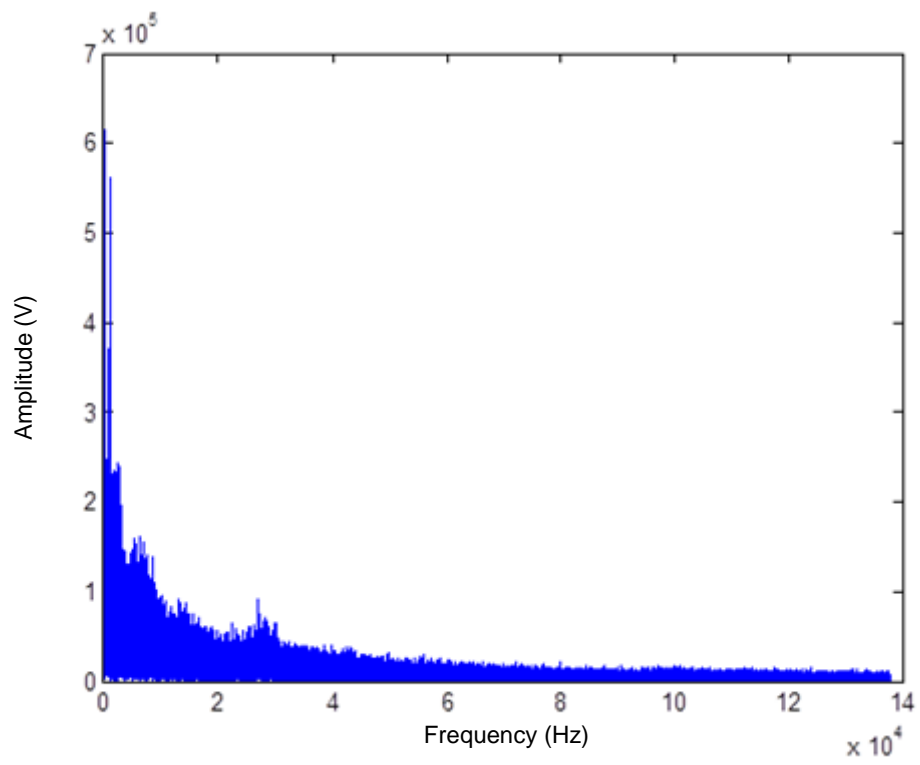
Figure 43 (c): MC 3
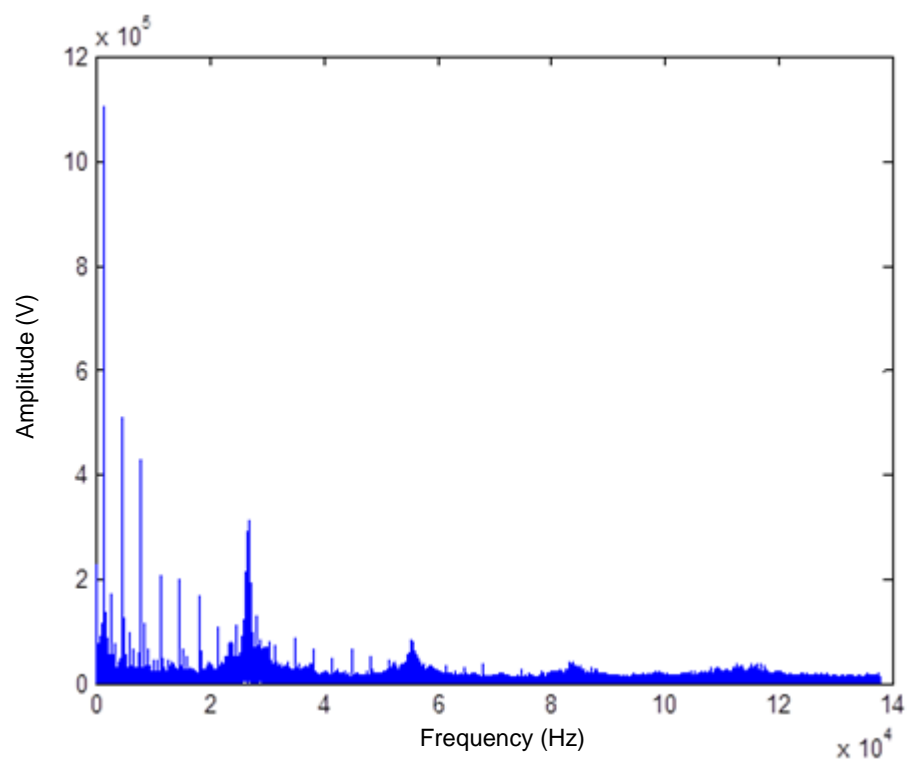
Figure 43 (d): MC 4
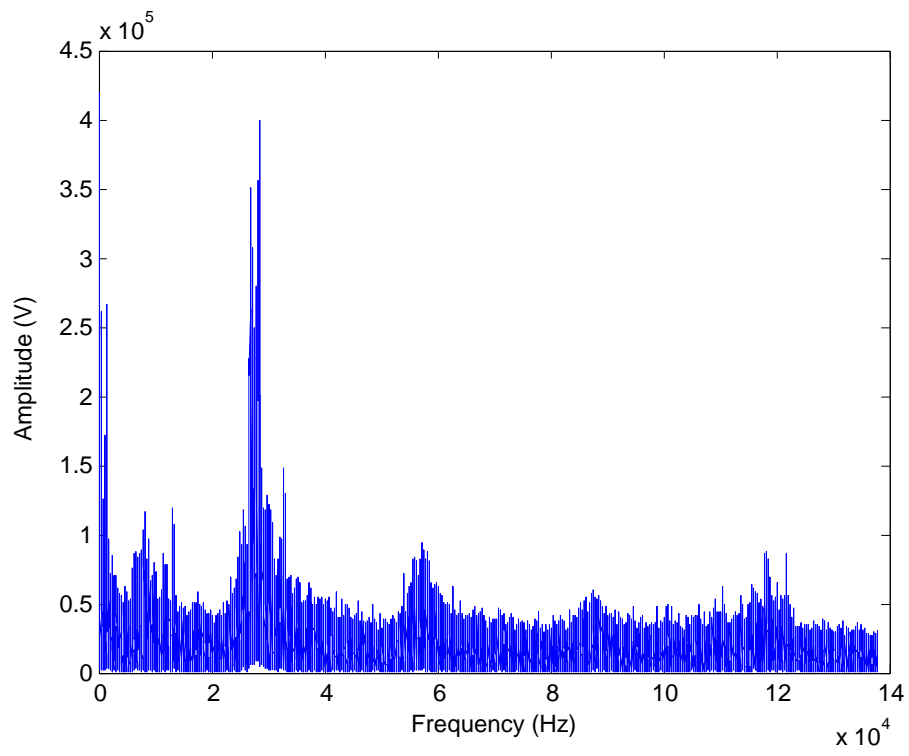
Figure 43 (e): MC 5
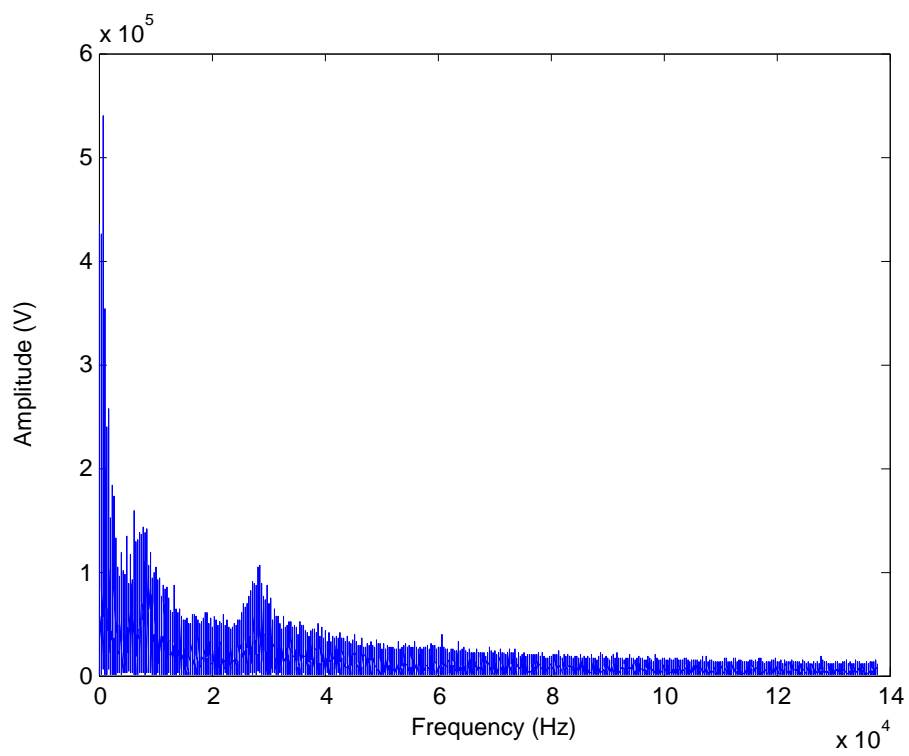
Figure 43 (f): MC 6
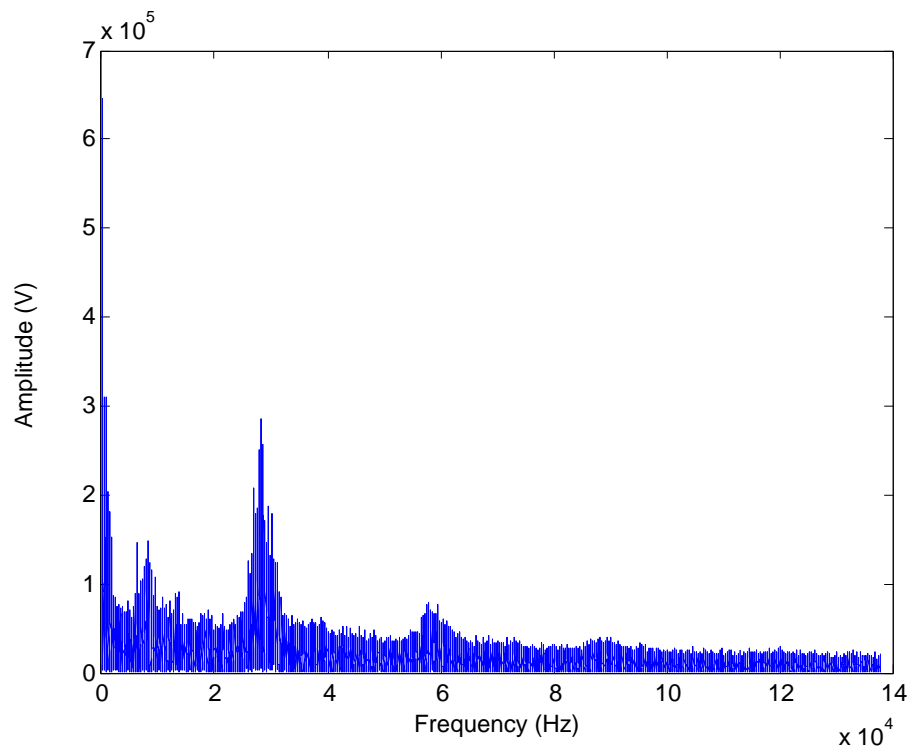
Figure 43 (h): MC 7

Figure 43 (g): MC 8

130
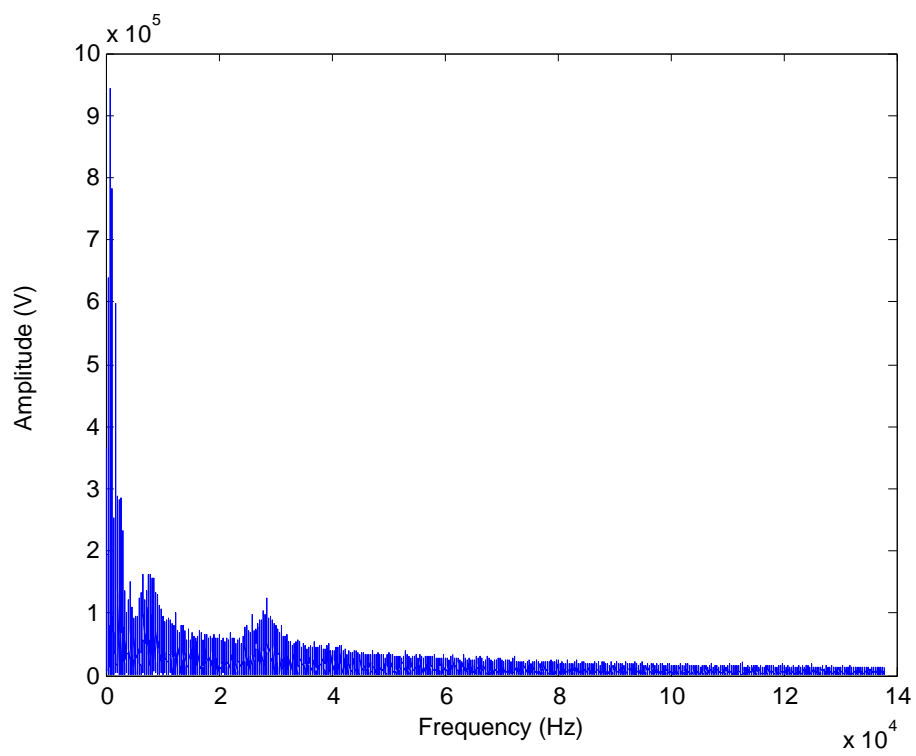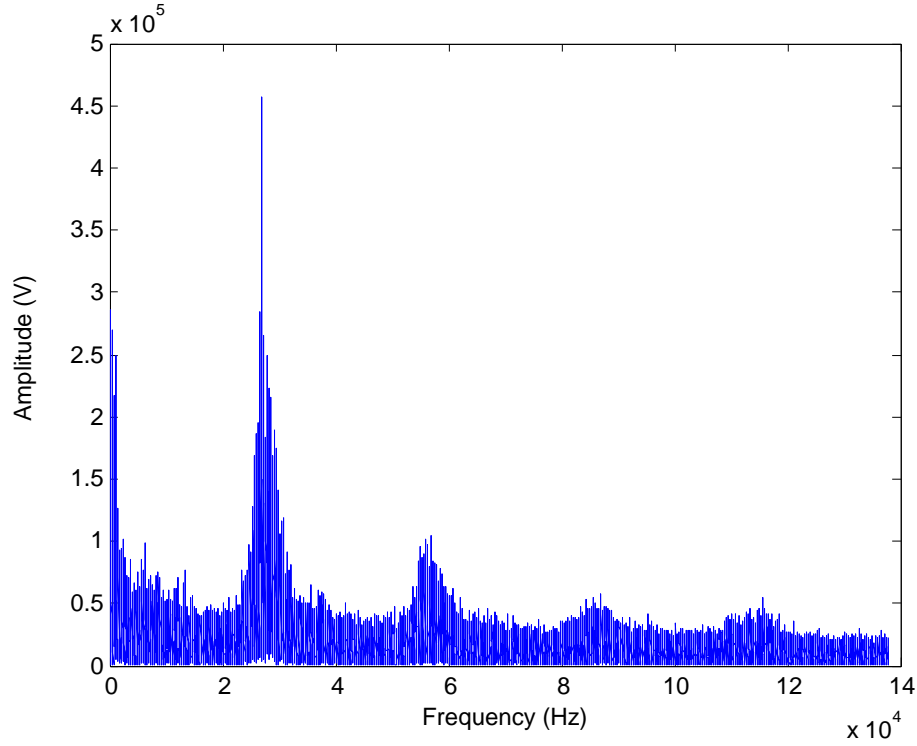
Figure 43 (i): MC 9

Figure 43: AE frequency spectra of the machine conditions addressed (2 -140 kHz) at

6960 RPM.

5.4.2.1 THE MAIN FEATURES OF THE DIFFERENT FAULT CONDITIONS

Table 15 illustrates the main features of the training FFT spectra shown in Figure 43.

These features were identified by the feature extraction algorithm which had been

developed. The program split the frequency range (2 kHz to 122 kHz) into six equal

segments, and identified the peak amplitude in each segment of each machine condition.

Table 12: Peak FFT amplitudes of all machine fault signatures at 6960 RPM.

Freq. (Hz)	2000-21999	22000-41999	42000-61999	62000-81999	82000-101999	102000-121999
MC 1	1,167,096	128,868	50,777	35,866	27,004	21,370
MC 2	228,564	442,717	113,155	107,457	64,208	43,737
MC 3	615,039	91,521	40,470	25,733	22,720	17,667
MC 4	1,104,438	315,618	87,224	40,506	44,697	38,060
MC 5	419,489	399,212	94,537	63,086	59,530	87,375
MC 6	539,553	106,488	42,832	39,763	23,470	18,269
MC 7	644,833	285,015	79,516	59,659	40,048	29,725
MC 8	942,486	123,970	49,047	34,752	25,396	21,895
MC 9	285,771	456,392	104,739	61,873	57,300	55,109

5.4.3 AE FAULT SIGNATURES AT 15650 RPM

Figure 44 shows the AE signatures measured at bearing (A) using the R6a AE sensor at 15650 RPM. Figure 44 (a) indicates the benchmark signature of the healthy condition (MC 1) while Figure 44 (b) to Figure 44 (i) show the signatures of the fault conditions addressed (MC 2 to MC 9). The shapes of AE signatures are different and faults can be easily differentiated. The spectra below (benchmark fault signatures) were utilised to train the proposed algorithm.
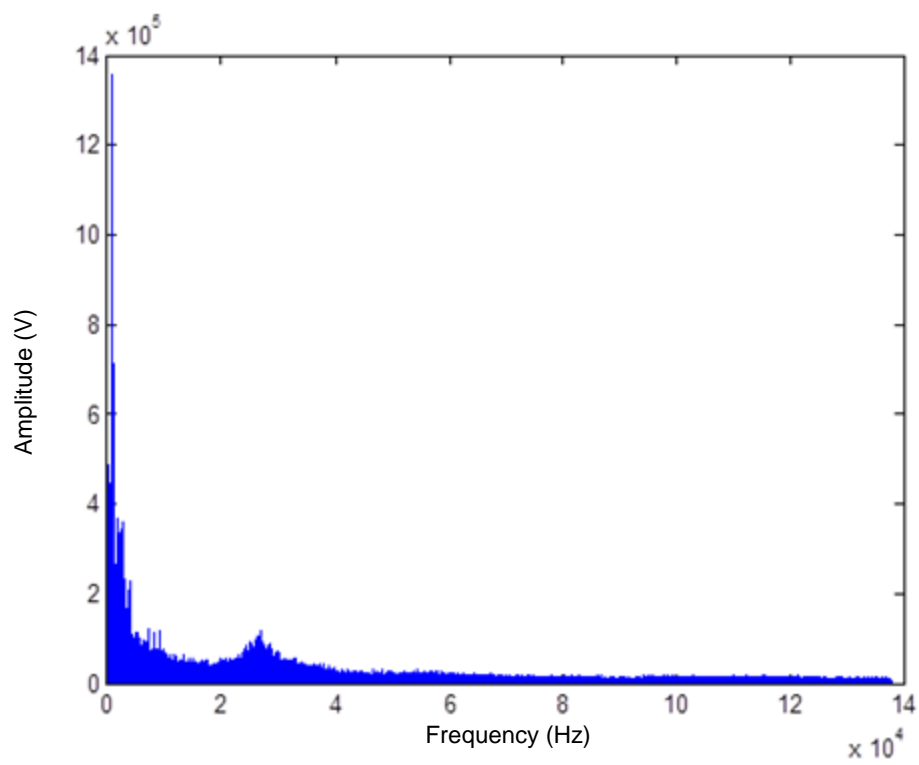
Figure 44 (a): MC 1
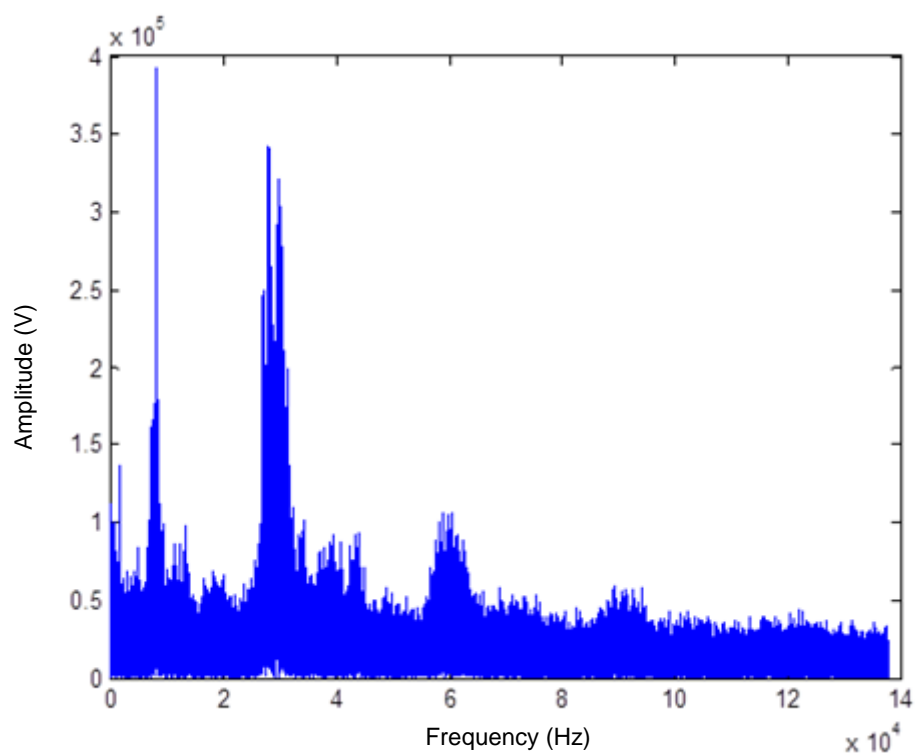
Figure 44 (b): MC 2
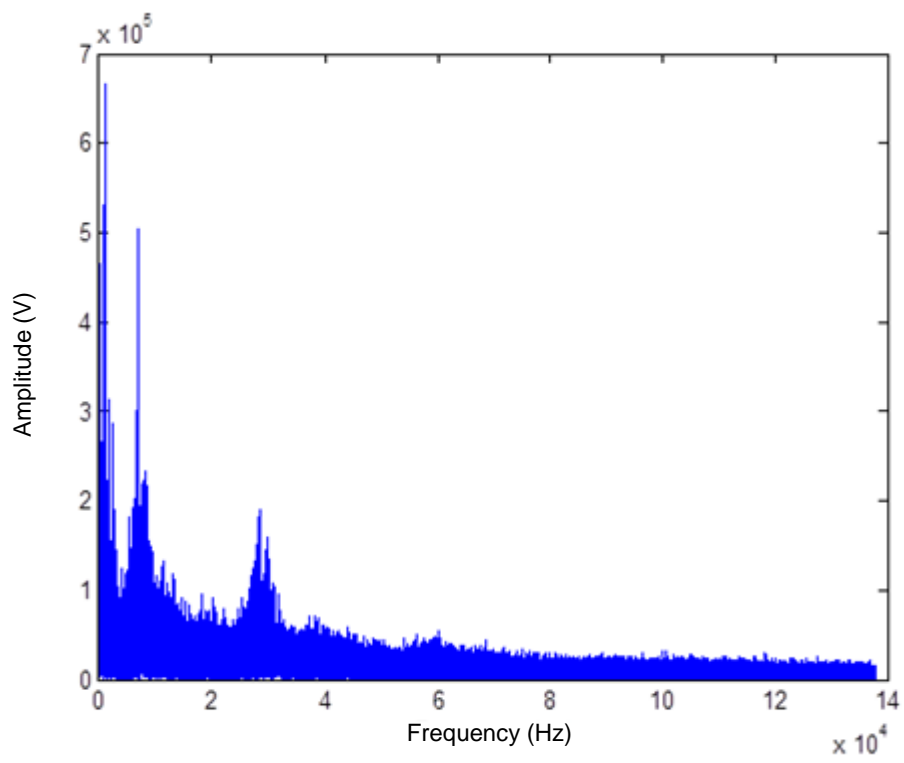
Figure 44 (c): MC 3
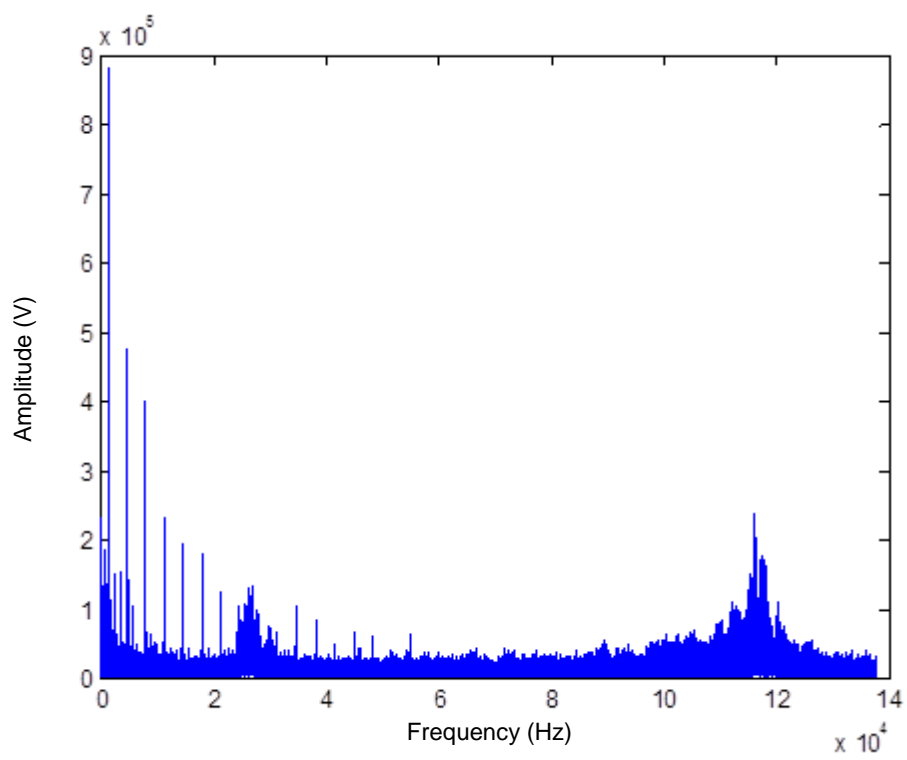
Figure 44 (d): MC 4
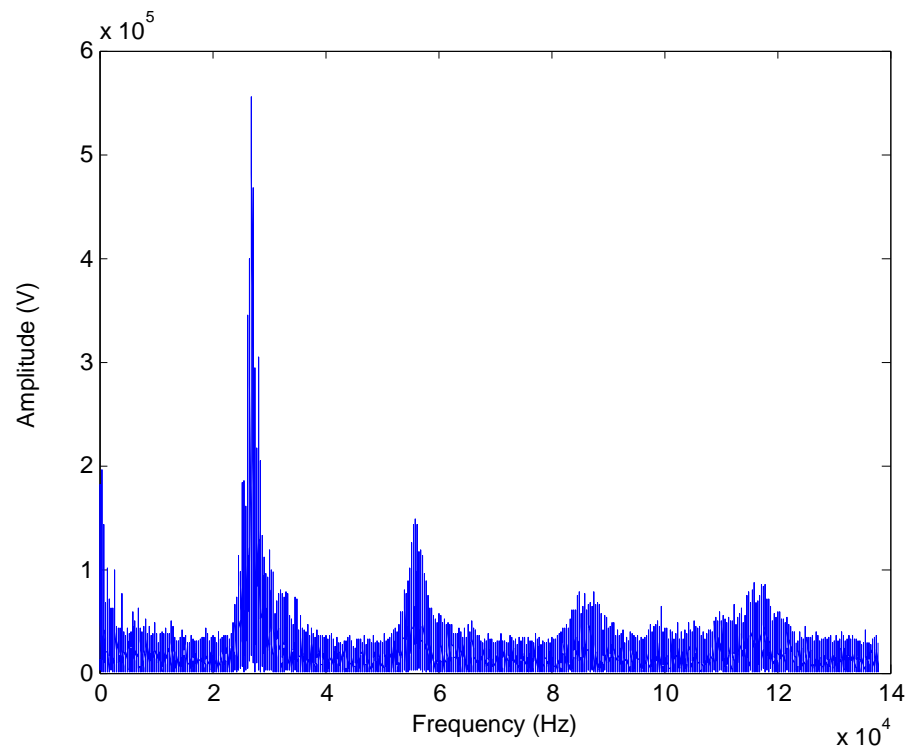
Figure 44 (e): MC 5

Figure 44 (f): MC 6

135
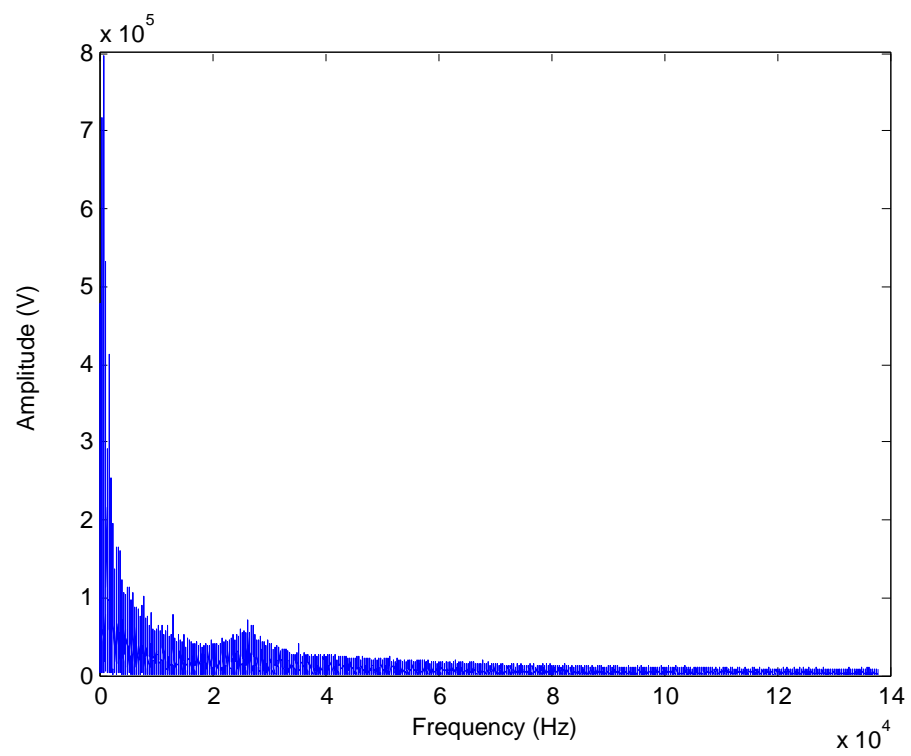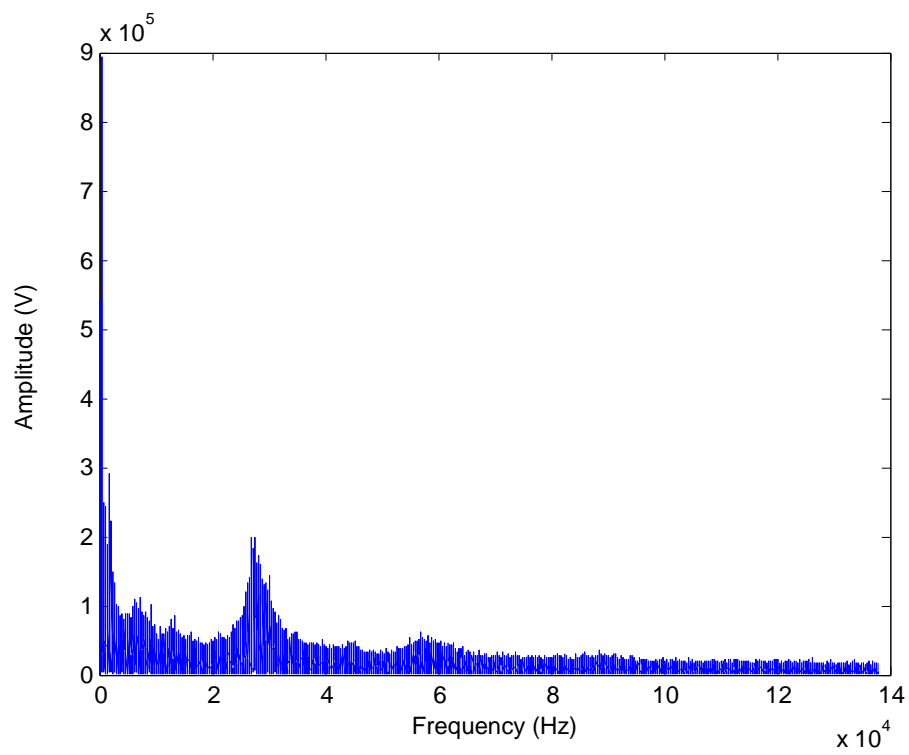
Figure 44 (g): MC 7

Figure 44 (h): MC 8

136
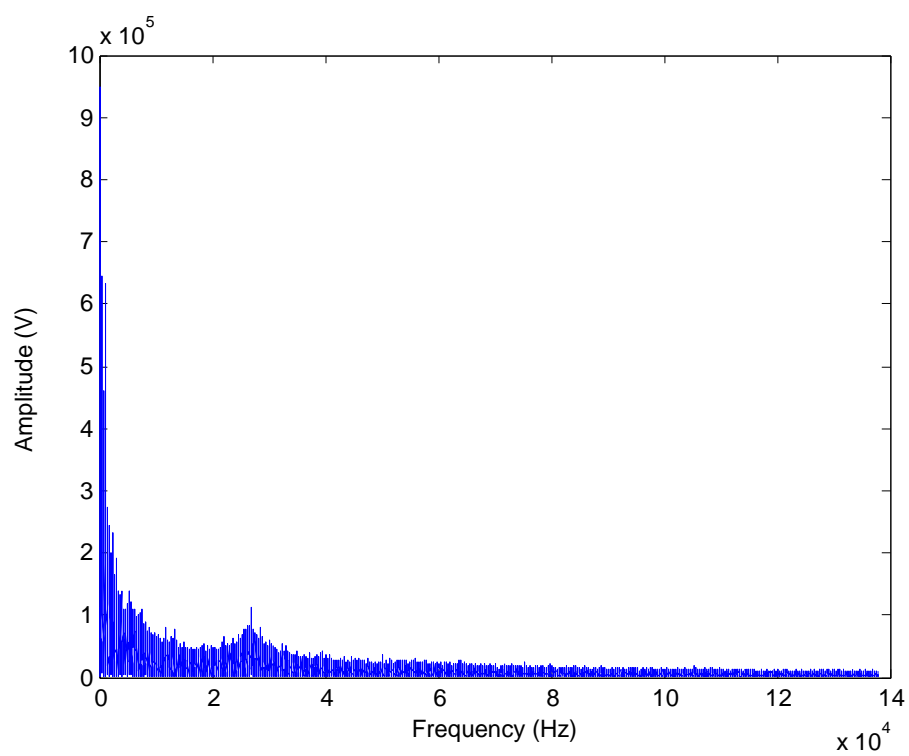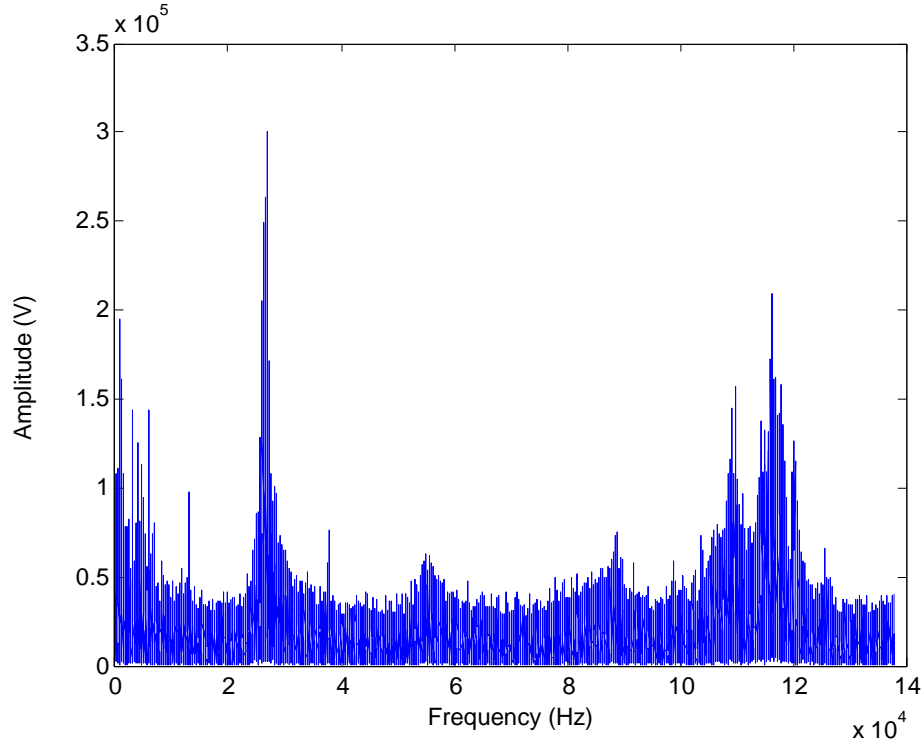
Figure 44 (i): MC 9

Figure 44: AE frequency spectra of the machine conditions addressed (2 - 140 kHz) at

15650 RPM.

5.4.3.1 THE MAIN FEATURES OF THE DIFFERENT FAULT CONDITIONS

Table 16 summarises the main features of the training FFT spectra shown in Figure 44.

These features were identified using the proposed features extraction algorithm. The

program split the frequency range (2 kHz to 122 kHz) into 6 equal segments and

identified the peak amplitude in each segment of each machine condition.

Table 13: Peak FFT amplitudes of all machine fault signatures at 15650 RPM.

Freq. (Hz)	2000-21999	22000-41999	42000-61999	62000-81999	82000-101999	102000-121999
MC 1	1,358,010	116,955	35,523	28,384	22,327	17,772
MC 2	392,496	342,191	106,436	107,199	59,807	43,562
MC 3	665,579	191,419	61,130	56,713	32,086	31,233
MC 4	880,479	133,362	65,893	44,021	57,739	237,854
MC 5	196,444	554,600	148,681	56,288	77,670	87,238
MC 6	796,217	70,775	27,424	19,632	15,536	12,681
MC 7	893,740	198,546	60,773	49,754	34,602	26,031
MC 8	949,661	112,431	37,095	27,830	20,019	17,606
MC 9	194,498	300,274	63,290	50,102	75,132	209,323

It can be observed that, at high operational speeds, the pattern shapes obtained significantly changed. The fault patterns of MC 3 clearly identify a pattern shape change, as a significant difference can be seen between the patterns at 15650 RPM and the patterns at 3600 and 6960 RPM.

5.5 PERFORMANCE EVALUATION

A sample of nine datasets was utilised to evaluate the detection accuracy of the algorithm developed. The spectra were divided into six segments based on the output of the proposed FS2FI algorithm. Hence, for each machine condition, six segments with six different peak FFT amplitude values were successfully identified by the proposed algorithm.

Table 17 shows the values of the peak amplitude spectral features of the nine machine conditions tested, at a rotational speed of 15650 (Dataset #2). These nine datasets were utilised for testing, and were different than those used for training (Table 16).

Table 14: Peak FFT amplitudes of the nine testing datasets at 15650 RPM.

Freq. (Hz)	2000-21999	22000-41999	42000-61999	62000-81999	82000-101999	102000-121999
MC 1	1,430,258	116,183	37,852	27,367	23,301	20,272
MC 2	187,027	420,193	97,205	91,691	77,619	45,597
MC 3	803,298	229,145	61,524	46,563	36,117	30,080
MC 4	890,325	155,739	81,476	48,524	61,330	281,009
MC 5	279,718	515,917	132,680	54,501	81,390	104,639
MC 6	816,365	73,517	30,237	22,441	17,349	15,172
MC 7	585,540	200,443	56,368	48,188	39,798	26,220
MC 8	915,350	103,044	38,767	28,527	23,067	16,007
MC 9	216,585	382,526	66,894	45,617	73,898	202,596

Table 18 presents the result of the FS2FI algorithm developed, and shows the comparison results between the peak FFT values of the testing dataset and all benchmark peak values at each corresponding segment. The algorithm selects the nearest benchmark peak value, and identifies to which machine condition it belongs. The algorithm only detects a fault if the majority of features are found to match one of the benchmark fault pattern features. The machine condition detected at each segment is shown in Table 18. With this segment size, seven conditions out of nine were successfully detected, yielding a detection accuracy of 63%. MC 1, 2, 4, 5, 6, 7 and 8 were successfully detected when the values of the FFT peak amplitudes shown in Table 17 were compared to the benchmark values shown in Table 16, thereby yielding a detection accuracy of 63%.

Table 15: Results of fault identification process using dataset # 2 at 15650 RPM (correct diagnoses in bold).

Freq. (Hz)	2000-21999	22000-41999	42000-61999	62000-81999	82000-101999	102000-121999	Detected MC
MC 1	*MC1*	*MC1*	MC8	MC8	*MC1*	*MC1*	MC1
MC 2	MC5	*MC2*	*MC2*	*MC2*	MC9	*MC2*	MC2
MC 3	MC6	MC7	*MC3*	MC4	MC7	*MC3*	--
MC 4	*MC4*	*MC4*	*MC4*	MC7	MC2	*MC4*	MC4
MC 5	*MC5*	*MC5*	*MC5*	*MC5*	*MC5*	*MC5*	MC5
MC 6	*MC6*	*MC6*	*MC6*	*MC6*	*MC6*	MC8	MC6
MC 7	MC3	*MC7*	*MC7*	*MC7*	*MC7*	*MC7*	MC7
MC 8	*MC8*	*MC8*	*MC8*	MC1	MC1	*MC8*	MC8
MC 9	MC5	MC2	MC4	MC4	*MC9*	*MC9*	--

The FS2FI algorithm which was developed yielded a 63% detection accuracy rate when using a single dataset for training. This low fault identification accuracy could be attributed to the methods utilised for pattern differentiation and pattern classification. The training is based on a single benchmark fault pattern for each machine condition which does not enable the algorithm to consider the variation in fault patterns over time. The algorithm only produces the minimum number of segments and features at which all fault patterns are identifiable. This adversely affects the fault identification performance as the algorithm does not consider the maximisation of differences between the features of fault signals, and its ability to accurately identify the fault pattern is limited. Moreover, the rule-based MATLAB program identifies faults on the basis of a minimum number of matching features, which in turn leads to false fault identifications.

5.6 MODIFIED FFT-BASED SEGMENTATION, FEATURE SELECTION AND FAULT IDENTIFICATION (MFS2FI) ALGORITHM

The previously proposed algorithm failed to detect all the machine conditions addressed. The main disadvantages of the FS2FI algorithm are that it produces the minimum possible number (count) of segments and features to differentiate between fault patterns, and also utilises the minimum possible number of matching features to identify the fault. Moreover, it does not investigate other possible solutions at different segment sizes, nor quantify the similarity between the fault signal patterns of different machine conditions at each segment.

Hence, a modified algorithm will be proposed in this section to improve detection accuracy, as well as to provide CM system developers with a measure for the differences between all of the machine condition features addressed, at each segment. This measure will be called "Confidence Level" (CL). The enhanced algorithm will determine the detection confidence level at each segment in order to give the ability to the developer of the CM system to weigh up the detection confidence level against the computing time and cost. The algorithm will be also utilised to evaluate the suitability of the spectral features of the AE technique for fault diagnosis in typical centrifugal equipment.

The modified algorithm (MFS2FI) automatically breaks down the main frequency domain range into smaller groups of frequencies. The computing time required for building, analysing, verifying and testing segments will increase with every additional segment the algorithm creates. For this application, based on the distribution of the majority of FFT peaks, the segment sizes (S) were selected from a range of 1,000 to 119,000 Hz, though more specifically starting from 2000 Hz with a loop step of 1000 Hz.

5.6.1 EMULATED MACHINE CONDITIONS

The same experiments were carried out in the laboratory except for those which emulated bearing lubrication problems. To acquire more data samples, the emulation time was increased from 40s to 177s. When attempting to carry out experiments over a longer duration to emulate the lack of bearing lubrication faults, a complete failure of a bearing occurred due to rapid overheating at this high rotational speed. For the purposes of risk management, as a lack of lubrication causes fusing, the four experiments that emulated the lack of lubrication fault over a time period of 177 seconds at this high rotational speed were excluded.

Five machine conditions were emulated at the highest rotational speed (15,650 RPM), as shown in Table 19. The R6a sensor which was directly positioned above bearing (A) gave the highest reading at 15,650 RPM. Hence, as the experiment was designed to have one AE sensor only, the bearing (A) R6a sensor was selected for its proper installation and high AE readings.

Table 16: Machine health conditions.

	Bearing (A)	Bearing (B)	Leakage
MC 1	Healthy	Healthy	No
MC 2	Healthy	Healthy	Yes
MC 3	Outer race defect	Healthy	No
MC 4	Healthy	Outer race defect	No
MC 5	Outer race defect	Outer race defect	No

Five experiments were conducted at a rotational speed of 15,650 RPM. The first experiment emulated the healthy condition while the other four experiments emulated four different fault conditions as shown in Table 19. The experiments described in Section 5.2 were carried out in the laboratory except for those which emulated bearing

lubrication problems. To acquire more data samples and datasets, the emulation time was increased from 40s to 177s. When attempting to carry out experiments over a longer duration to emulate the lack of bearing lubrication faults, a complete failure of a bearing occurred due to rapid overheating at this high rotational speed. For the purposes of risk management, as a lack of lubrication causes fusing, the four experiments that emulated the lack of lubrication fault over a time period of 177 seconds at this high rotational speed were excluded.

The data were sampled using the high speed NI DAQ board at a sampling rate of 1 MS/s for a time period of 1 s each. For each of the five machine conditions, a Machine Fault Signature Set (MFSS) was collected. Each MFSS includes 10 datasets (DS# 1 to DS# 10) collected at a fixed time interval of 13s (one set every 13 seconds). Each dataset has a size of 1×10^{6} samples at 1MHz. The first datasets for the five machine conditions were recorded 60 seconds after the blower reached its full rotational speed. 50% of the 50 datasets (DS) were used for both training and validation (DS# 1, 3, 8, 10 for training and DS# 5 for validation) while the rest were used for testing.

5.6.2 DETAILED ALGORITHM DESCRIPTION

The main focus of the MFS2FI algorithm is to improve the feature selection and comparison processes to better classify fault patterns. Figure 45 illustrates a simplified flowchart that describes the fault identification process using a modified FFT-based segmentation and fault identification algorithm. It starts by transforming all training and real-time signals from time-domain to frequency-domain (FFT spectrum). Each spectrum is segmented into a number of equal segments *(Segmentation of spectra)*. Each segment represents a specific range of frequencies, and has its own unique maximum amplitude value. The minimum and maximum benchmark amplitude values at each segment are calculated on the basis of the training FFT spectra of each machine condition. The

maximum FFT amplitude value at each segment of the real-time signal is selected as a unique feature of this frequency division *(Feature selection)*. The selected feature is then compared to the corresponding minimum and maximum benchmark threshold amplitude values for all machine fault conditions. The algorithm selects the machine condition with the nearest maximum benchmark amplitude value and inserts the machine condition number at the matrix element that corresponds to the segment number. The algorithm also selects the machine condition with the nearest minimum benchmark amplitude value and inserts the machine condition number at the matrix element that corresponds to the segment number. The result of the comparison is inserted into a matrix split into two halves: the first half includes the results of a comparison between the real-time signal and the maximum benchmark amplitude values of all machine conditions, and the second half includes the results of a comparison between the signal and the minimum benchmark amplitude values. This matrix was purposely built up to facilitate the counting of matching features and hence the classification of machine fault patterns. This step is iteratively repeated from the first segment to the last segment in order to build up a comparison matrix that considers all segments (feature comparison). The comparison matrix is then utilised to count up the total number of features that match each machine condition (quantification of matching features).

The fault identification process is based on the machine condition that mostly appears (or counted) in the matrix. For example, if an FFT spectrum with 100 segments and 100 features is utilised for fault identification, and the results of the quantification of matching features show 40 features matching with MC 1, 35 features matching with MC 2, and 25 features matching with MC 3, then the algorithm will identify MC 1 as a possible solution as it has the largest number of features matching this machine fault condition.

In order to improve the certainty of the detection and avoid false alarms, the algorithm is set to not identify the machine condition if the minimum difference between the first and second numbers of matching features was smaller than the CL. The CL is a numerical parameter whose maximum value increases with the number of segments or frequency divisions. The CL gives a real measure of the similarity between two fault spectra and its maximum value is equal to double the number of segments. Hence, the larger the number of segments, the better the level of fault identification certainty will be.
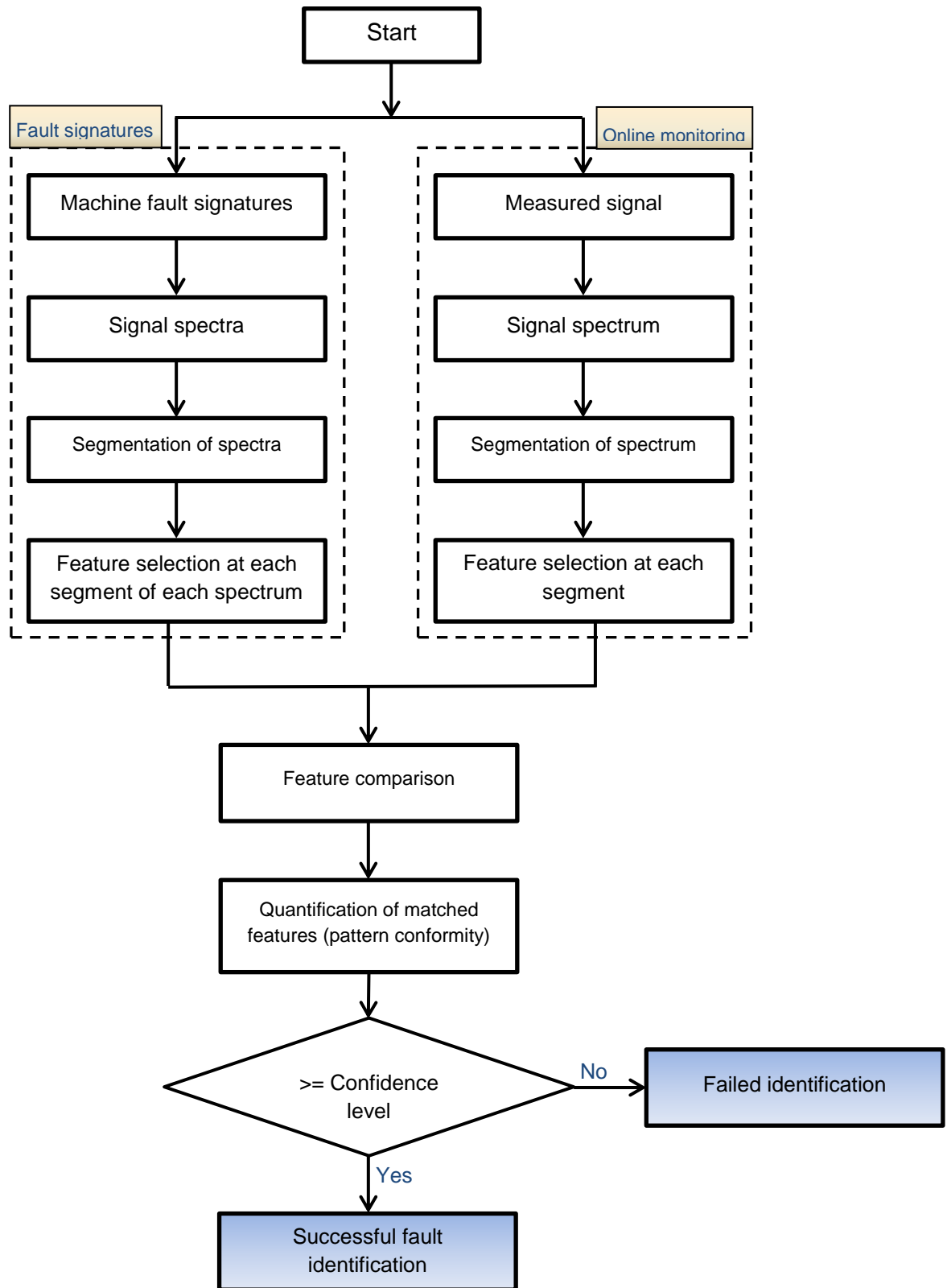
Figure 45: Illustrative flowchart of the MFS2FI algorithm.

The detailed MFS2FI algorithm shown in Figure 46 starts by reading twenty training datasets, four datasets (DS# 1, 3, 8 and 10 from MFFS sets) for each machine condition for each of the five machine conditions. Then it converts the signal from time domain to frequency domain. Depending on the step value of loop "S", the algorithm determines the maximum number of different segment sizes (SNmax) and the segment size (S). The algorithm utilises the value of S to determine the maximum number of segments (Kmax). SN and K are both needed for the nested loops as the number of iterations of SN loop ranges from 1 to SNmax, while the number ranges from 1 to Kmax for K. The frequency spectra and S, SNmax, K and Kmax values are then utilised to produce five FSelection matrices through processes A, B and C. The algorithm flowchart shown in Figure 46 illustrates the main six processes, namely processes A, B, C, D, E and F. The full MATLAB (Version12) code is given in Appendix C.
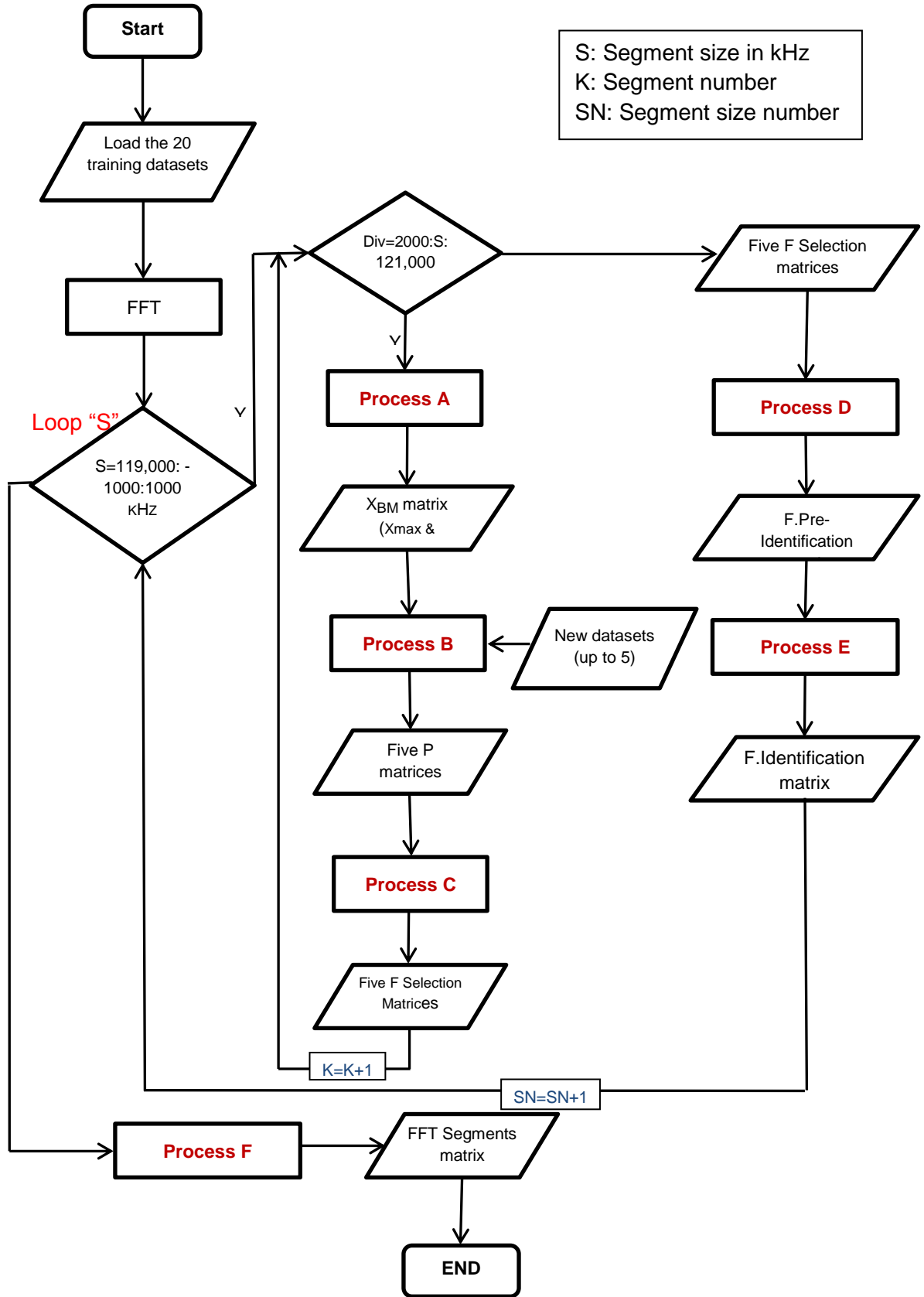
Figure 46: Detailed flowchart of the MFS2FI algorithm.

The role of each process can be described as follows:

Process A (first training cycle – X_{BM} matrix): this process takes the following inputs: number of machine health conditions (NoC), segment size (S), frequency range value (Div). It then calculates the segment number (K). It also calculates the maximum FFT segment amplitudes ($X_n(NoC,K)$) of all training signals of all machine conditions. X_n is the number of training signal and ranges from 1 to 4. Then the process outputs the maximum and minimum amplitude values of each machine condition, and puts them all in $X_{max}(NoC,K)$ and $X_{min}(NoC,K)$ matrices. These two matrices are combined and considered as a benchmark threshold matrix for all the machine conditions addressed. Table 20 shows the output of Process A at K=1:2, and NoC=1:5 using the MATLAB equations shown below as equations (7), (8) and (9).

$$X_{max}(NoC,K) = \max(X_n(NoC,K)) \qquad (7)$$

$$X_{min}(NoC,K) = min(X_n(NoC,K)) \qquad (8)$$

$$X_{BM} = \begin{pmatrix} X_{max} \\ X_{min} \end{pmatrix} \qquad (9)$$

Table 17: Benchmark matrix (X_{BM}) using the training datasets number DS# 1, 3, 8 and

10.

(a)X_{BM} – Xmax section.

	Segment 1 (S$_1$) (2 :108.99 kHz)	Segment 2 (S$_2$) (109 : 121.99 kHz)
MC 1	19795	492
MC 2	14934	457
MC 3	55989	270
MC 4	85780	15621
MC 5	36788	6008

(b)X_{BM} – Xmin section.

	Segment 1 (S$_1$) (2 : 108.99 kHz)	Segment 2 (S$_2$) (109 to 121.99 kHz)
MC 1	13312	446
MC 2	9422	451
MC 3	45789	163
MC 4	79350	9871
MC 5	23106	4749

Process B (Percentage difference – P matrix): This process takes the X_{BM} matrix, in addition to a maximum of five new AE datasets (online data), as inputs. Firstly the process segments the new signals into K segments and then calculates the percentage differences between the amplitudes of the corresponding new segmented datasets (maximum of 5 datasets per patch) and the benchmark threshold amplitudes at each segment; finally, it yields five percentage matrices (P_1 to P_5) where the subscript indicates the dataset number (or the machine condition number in this example). The percentage differences are calculated for all the segments using MATLAB equation (10). $NoDS_T$ is the total number of new datasets i.e. 5. In this example, the $NoDS_T$ is equal to the total number of machine conditions (NoC_T).

For example, a new dataset for machine condition 2 (MC 2) is converted into frequency domain, and then segmented into (K) segments. The maximum amplitude at each segment (X) is calculated, and then compared to all corresponding benchmark threshold amplitudes (X_{BM}). The results of this comparison are put into the P_2 matrix.

$$P_{NoC}\ (1:2*NoC_T\ (or\ NoDS_T), K) = \frac{X_{NoC}(1:2*NoC_T,K) - X_{BM}(1:2*NoC_T,K)}{X_{BM}\ (1:2*NoC_T,K)}\ \% \tag{10}$$

Process C (Feature selection – FSelection matrix): This process builds a matrix portioned into two equal sections called FSelection for each machine condition. For the first section, the algorithm compares the maximum amplitudes of each of the five new AE datasets to all X_{BM} elements at each segment (K) yielding five FSelection matrices. Then, if the amplitude of the new dataset does not fit inside the range X (from X_{max}(NoC, K) to X_{min}(NoC, K)), the algorithm sets the corresponding FSelection element value to zero. Otherwise, the value is set according to the machine condition number which has the amplitude value within its amplitude interval. The machine condition numbers are 1, 2, 3, 4 and 5.

The second section of the FSelection matrix is calculated by selecting the minimum absolute percentage of the corresponding P matrix (P_1 to P_5). The corresponding FSelection matrix element is set to 1, 2, 3, 4 or 5 according to the machine condition number (NoC) using MATLAB equations (11) and (12), while the other matrix elements are set to "zero". The minimum absolute percentage is selected to limit the selection to only one pattern (one machine condition) for each dataset of the five new datasets. Hence, the corresponding matrix elements are set to zero when two or more identical P matrix elements are observed. A zero value means that this segment cannot be utilised to identify this fault.

Table 21 shows a FSelection matrix calculated using DS# 5 from MFSS# 5 (MC 5), at a segment size of 109 kHz. The matrix shown in Table 21 includes the results of the comparison between the MC 5 verification dataset and X_{BM}. The results show that the signal features of the verification dataset match all the four benchmark signal features of machine condition 5. It can be seen that '5' appears against MC 5 in both segments and matrix sections, and that it appears four times, each time showing a correct identification at this segment.

$$M_{FSelection_{NoC}}(\ RowIndex(\min(P_{NoC}(1:NoC_T,K)),K) =$$
$$RowIndex(\min(P_{NoC}(1:NoC_T,K)) \tag{11}$$

$$M_{FSelection_{NoC}}(RowIndex(\min(P_{NoC}(NoC+1:2*NoC_T,K)),K) =$$
$$RowIndex(\min(P_{NoC}(NoC+1:2*NoC_T,K)) \tag{12}$$

Table 18: FSelection matrix used for the fault identification of MC 5 using a segment size of 109 kHz (dataset number DS#5 from MFSS#5).

		Segment 1 (S$_1$)	Segment 2 (S$_2$)
	MC 1	0	0
	MC 2	0	0
Section I	MC 3	0	0
	MC 4	0	0
	MC 5	5	5
	MC 1	0	0
	MC 2	0	0
Section II	MC 3	0	0
	MC 4	0	0
	MC 5	5	5

Process D (Fault pre-identification – FPre-Identification matrix): This process merges the five FSelection matrices into one matrix, defined as a Pre-decision matrix, by counting up the features that match the benchmark threshold of each machine condition using MATLAB equation (13). This can be achieved by counting up every occurrence of the machine condition number in the matrix. The matrix consists of five columns (one for each validation signal), and five rows (one for each benchmark signal).

Table 22 illustrates the outputs of process D at different segment sizes using DS#5. DS#5 includes 5 datasets; one dataset for each machine condition (MFSS#1 to MFSS#5), and it was randomly selected to verify the algorithm. X_{NoC} indicates the number of the MFSS from which the validation kdataset was selected. For a 100% correct diagnosis, the maximum number of matching features should exist at the matrix diagonal elements. Thus, as presented in Table 22, the MC 1 signal cannot be diagnosed when S = 109 Hz, while it can be correctly diagnosed when S = 1 kHz.

$$M_{FPre-Identification}(\text{NoC}, X_{NoC}) = count(NoC, Range: M_{FSelection_{NoC}}(:,:)) \qquad (13)$$

Table 19: FPre-Identification matrix using five DS#5 from MFSS#1 to MFSS#5; one dataset for each machine condition (largest number of matching features in bold).

(a) FPre-Identification matrix using a segment size of 109 kHz.

Validation / Benchmark	MC 1	MC 2	MC 3	MC 4	MC 5
MC 1	**2**	1	0	0	0
MC 2	**2**	**3**	0	0	0
MC 3	0	0	**2**	0	0
MC 4	0	0	0	**4**	0
MC 5	0	0	1	0	**4**

(b) FPre-Identification matrix using a segment size of 1 kHz.

Validation / Benchmark	MC 1	MC 2	MC 3	MC 4	MC 5
MC 1	**180**	62	16	4	41
MC 2	63	**153**	17	2	20
MC 3	20	23	**169**	2	20
MC 4	3	2	2	**198**	2
MC 5	35	31	28	0	**163**

Process E (Fault identification – F-Identification matrix): This part of the modified algorithm checks whether the patterns of the different machine conditions are recognised. The calculation is based on the Pre-Identification matrix, where the algorithm selects the machine condition with the largest number of features (mostly appeared in each of the matrix column). If the selection is correct and the machine condition is successfully detected at this set of segments, the algorithm sets the value of the corresponding zero F-Identification matrix element to "1" using MATLAB equation (14); in the meantime, other column element values remain as zero. Otherwise, the algorithm does not change the value of the corresponding F-Identification matrix element.

Table 23 shows the output of Process E at the same two segment sizes. The table consists of five columns (one for each validation signal), and five rows (one for each benchmark

signal). It can be seen that at a segment size of 109 kHz (Table 23(a)), the identification of MC 1 is not possible, whilst at 1kHz (Table 23(b)), all machine conditions are successfully identified.

$$M_{F-Identification}\ (NoC, X_{NoC}) =$$

$$1\ \ if\ (RowIndex\big(\max((M_{FPre-Identification}(1:NoC\ _T, X_{NoC})))\big) = NoC) \tag{14}$$

Table 20: F-Identification matrix for DS#5. The identifiabilty of each machine condition is shown in bold ("1" for identified machine conditions and "0" for non-idnetified conditions).

(a) F-Identification matrix using a segment size of 109 kHz.

Validation \ Benchmark	MC 1	MC 2	MC 3	MC 4	MC 5
MC 1	**0**	0	0	0	0
MC 2	0	**1**	0	0	0
MC 3	0	0	**1**	0	0
MC 4	0	0	0	**1**	0
MC 5	0	0	0	0	**1**

(b) F-Identification matrix using a segment size of 1 kHz.

Validation \ Benchmark	MC 1	MC 2	MC 3	MC 4	MC 5
MC 1	**1**	0	0	0	0
MC 2	0	**1**	0	0	0
MC 3	0	0	**1**	0	0
MC 4	0	0	0	**1**	0
MC 5	0	0	0	0	**1**

Process F (Detectability of faults at all segment sizes - FFT segment matrix): The FFT segment matrix is built to show the detectability of different machine conditions at all segment sizes (S). These segment sizes are then passed to the fault identification part of the algorithm for machine condition identification. This matrix is calculated on the

basis of the previously calculated F-Identification matrix by placing the diagonal elements in a column matrix with a size of (5 x 1), and then inserting the simplified matrix into the FFT Segment matrix. Hence, the final dimensions of the matrix will be (5 x S). Table 24 illustrates an FFT segment matrix for five machine health conditions (NoC = 1:5) at all segment sizes (Sn=119:1 kHz), given a frequency range of 2 kHz to 122 kHz, as considered in the analysis.

The segment sizes with all row elements equal to 1 are to be passed to the fault identification algorithm for fast machine condition detection. The fault identification algorithm which is part of the fault MFS2FI algorithm will then select the most appropriate segment size depending on the overall confidence level (CL) required. The confidence level is defined as the difference between the number of matching features of the detected machine condition (correct diagnosis), and the second highest number of features matching other machine conditions.

The overall confidence level is calculated on the basis of the Pre-Identification matrix shown in Table 22, and is defined as the smallest confidence level number in a five confidence level array (one confidence level for each training dataset). The maximum confidence level value is equal to twice the number of segments (K), and an overall confidence level of zero means that the number of segments is not sufficient for differentiating between the frequency spectra. For example, Table 22 (b) shows that the values of the differences between the sums of the machine condition features (MC 1, MC 2, MC 3, MC 4 and MC 5) are calculated as 117 (180-63), 91, 141, 194 and 122, respectively. The maximum number of the confidence level at 1 kHz is 238 (2 x 119 segments). Hence, the smallest number of these CF values is 91, which is considered to be the overall confidence level. The larger the overall confidence level number, the better. Hence, although the computing cost at a segment size of 1 kHz is relatively high,

the accuracy of the solution at this segment size is best. Table 24 indicates that 108 kHz is the largest interval at which a correct diagnosis can be made.

Table 21: Transposed FFT segment matrix for the machine conditions addressed using DS#5 from all MFSS The results show the identifiability of each machine condition at different segment sizes ranging (in descending order) from 119 to 1 kHz (correct diagnosis of the entire machine conditions in bold).

Testing DS / Segm. size	MC 1	MC 2	MC 3	MC 4	MC 5	CL
119 kHz	0	1	0	1	0	0
118 kHz	0	0	1	1	1	0
117 kHz	1	1	0	1	1	0
116 kHz	1	1	0	1	1	0
115 kHz	0	1	0	1	1	0
114 kHz	0	1	1	1	1	0
113 kHz	0	1	1	1	1	0
112 kHz	0	1	1	1	1	0
111 kHz	0	1	1	1	1	0
110 kHz	0	1	1	1	1	0
109 kHz	0	1	1	1	1	0
108 kHz	**1**	**1**	**1**	**1**	**1**	**1**
107kHz- 2 kHz	**1**	**1**	**1**	**1**	**1**	**2-40**
1 kHz	**1**	**1**	**1**	**1**	**1**	**91**

Figure 47 gives a detailed example of how the algorithm produces the different matrices using the main processes at a segment size of 109 kHz. In this figure, only one of the five new datasets was processed; the other four datasets should be processed in order to complete the Pre-Identification and Identification matrices.

Process A outputs the benchmark matrix (X_{BM}) that includes the maximum and minimum benchmark amplitudes at each segment for all of the machine conditions addressed. X_{BM} and X_1 matrices are used to calculate the P_1 matrix in process B. The process should finally output five P matrices for five new datasets (X1, X2, X3, X4 and X5). Process C compares X_1 matrix to X_{BM} matrix in order to calculate the first section of the FSelection matrix. The second section of the FSelection matrix is calculated by selecting the minimum percentage at each segment. In section II of the FSelection matrix, number "1" in segment S_1 means that the first segment identifies MC_1, and number "2" in segment S_2 means that the second segment identifies MC_2. Process C should finally produce five FSelection matrices, one matrix for each dataset (or machine condition) using X1, X2, X3, X4 and X5 datasets. Process D calculates five Fault Pre_Identification matrices based on the pre-calculated FSelection matrices. In this process, the algorithm sums the matching features of each machine condition, for example, two features for MC_1, two features for MC_2, and no features for others. Process E selects the machine condition with the largest number of features, and puts it into the corresponding cell of the F-Identification matrix. Finally, the F-Identification matrix is transformed into a column matrix, and is inserted into its corresponding FFT Segment matrix column to give the full detection information at this segment size.

X_{BM}

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} S_1 & S_2 \\ 19795 & 492 \\ 14934 & 457 \\ 55989 & 270 \\ 85780 & 15621 \\ 36788 & 6008 \end{bmatrix} X_{max}$$

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} 13312 & 446 \\ 9422 & 451 \\ 45789 & 163 \\ 45789 & 9871 \\ 23106 & 4749 \end{bmatrix} X_{min}$$

Benchmark thresholds

$DS_5 (X_1)$

$$MC_1 \begin{bmatrix} S_1 & S_2 \\ 12930 & 470 \end{bmatrix}$$

$$\frac{X_1(1,k) - X_{BM}(:,k)}{X_{BM}(:,k)} \%$$

	NoC
MC_1	1
MC_2	2
MC_3	3
MC_4	4
MC_5	5

$$if \left(\begin{array}{c} X_1(k) \leq X_{BM}(NoC,k) \\ X_1(k) \geq X_{BM}(5+NoC,k) \end{array} \right) \\ \{F\ Selection_1(NoC,k) = NoC\}$$

$P_1 (\%)$

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} S_1 & S_2 \\ 34.68 & 4.47 \\ 13.42 & 2.93 \\ 76.91 & 74.20 \\ 84.93 & 96.99 \\ 64.85 & 92.17 \end{bmatrix}$$

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} 2.87 & 5.42 \\ 37.23 & 4.24 \\ 71.76 & 189.00 \\ 83.71 & 95.24 \\ 44.04 & 90.10 \end{bmatrix}$$

F Selection 1

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} S_1 & S_2 \\ 0 & 1 \\ 2 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$ Section I — Based on benchmark thresholds

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$ Section II — Based on P matrix

$NoC = Row\ index(min(P_1(:,k)));\ if(NoC>5),\ NoC=NoC-5$

Pre-Identification (NoC, DS_1)=Count(NoC, Range: All FSelection matrix

Pre-Identification Matrix

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} DS_1 & DS_2 & DS_3 & DS_4 & DS_5 \\ 2 & \vdots & \vdots & \vdots & \vdots \\ 2 & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

if (Row index(max(:,DS1))==1) { Identification (1, DS1)=1;}

Identification Matrix

$$MC_1, MC_2, MC_3, MC_4, MC_5 \begin{bmatrix} DS_1 & DS_2 & DS_3 & DS_4 & DS_5 \\ 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Figure 47: Example of how the MFS2FI algorithm processes a new dataset (DS5(X1) from MFSS#1, using a segment size of 109 kHz. FFT AE spectra were segmented into two segments (K= S1 and S2).

5.7 PERFORMANCE EVALUATION OF THE MODIFIED ALGORITHM

The fault identification performance of the MFS2FI algorithm was assessed using 25 different datasets. Five datasets for each of the five emulated machine conditions were collected using the experimental setup presented in chapter 4.

Figure 48 shows two samples of the twenty five AE FFT frequency spectra: the first figure (a) shows a healthy machine condition signature of MC 1, while the second figure (b) shows a faulty machine condition signature of MC 4. It can be observed that the

maximum amplitude of MC 1 in segment 1 is 14,413, while the maximum amplitude of MC 4 in segment 1 is 85,780. Due to the large difference in maximum amplitude values between these two machine conditions, a segment size of 119 kHz (from 2 kHz to 121 kHz) can be utilised for successful machine fault classification. Otherwise, if the difference between the maximum ampltidue values is small, the FFT spectra should be segmented into a larger number of divisions in order to increase the number of FFT features, and in turn, improve fault classification performance.
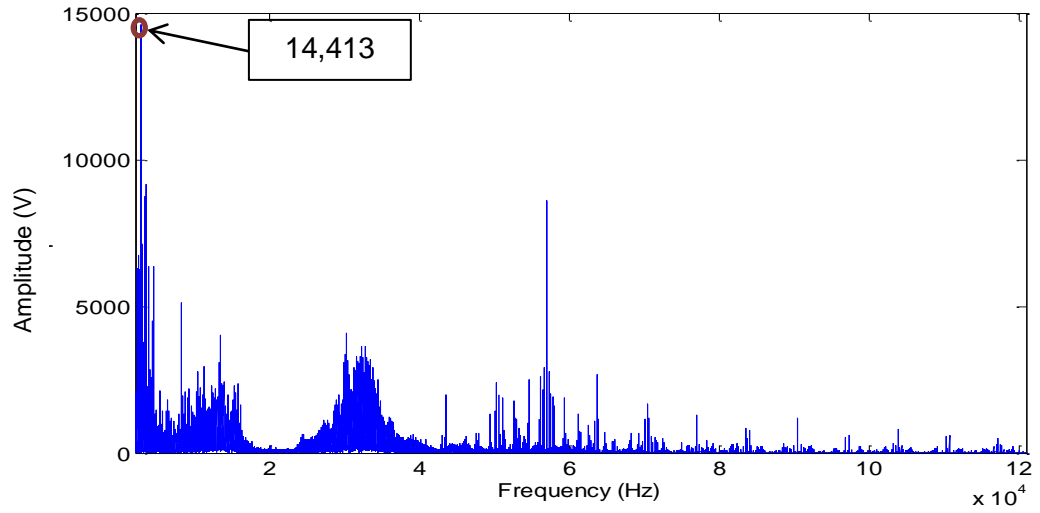
(a)Machine condition 1 (MC1).
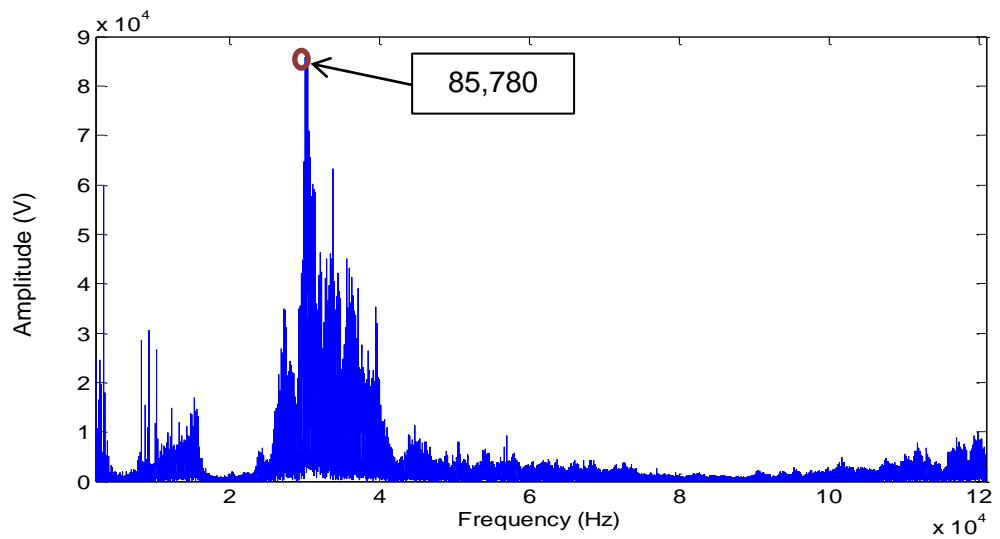
(b)   Machine condition 4 (MC4).

Figure 48: AE FFT spectra of machine conditions 1 and 4.

Table 25 presents the number of features matching each machine condition, and summarises the results of the testing process. 50% of the data (five datasets for each machine condition) were utilised for testing. The results presented in the table show that the MFS2FI algorithm managed to detect all of the machine conditions in issue with a

detection accuracy of 100%. The overall confidence level at 1 kHz is much higher than the overall confidence level at 108 kHz as the differences between the minimum number of matching features at 1 and 108 kHz are 92 and 1, respectively. However, all machine health conditions were successfully detected at both segment sizes with a success rate of 100%.

Table 22: Fault Pre-Identification matrices for 25 datasets at two different segment sizes (largest number of matching features in bold).

Segment size:		1 kHz					108 kHz				
Testing dataset#:		2	4	6	7	9	2	4	6	7	9
Healthy	**MC 1**	*177*	*182*	*170*	*182*	*165*	*3*	*3*	*3*	*3*	*2*
	MC 2	56	63	65	63	61	1	1	1	1	0
	MC 3	21	18	17	18	23	0	0	0	0	0
	MC 4	2	3	6	3	3	0	0	0	0	0
	MC 5	36	37	35	37	37	0	0	0	0	1
Leak	MC 1	68	65	64	65	59	1	0	0	0	0
	MC 2	*166*	*157*	*165*	*157*	*165*	*2*	*4*	*4*	*4*	*3*
	MC 3	21	21	24	21	24	0	0	0	0	0
	MC 4	3	2	4	2	3	0	0	0	0	0
	MC 5	33	33	37	33	31	0	0	0	0	0
Impeller	MC 1	11	13	20	13	18	0	0	0	0	0
	MC 2	20	17	15	17	15	0	0	0	0	0
	MC 3	*165*	*172*	*187*	*172*	*176*	*2*	*2*	*4*	*2*	*3*
	MC 4	1	2	1	2	2	0	0	0	0	0
	MC 5	31	29	27	29	26	1	1	0	1	0
Belt	MC 1	1	4	2	4	3	0	0	0	0	0
	MC 2	4	2	0	2	4	0	0	0	0	0
	MC 3	0	2	2	2	1	0	0	0	0	0
	MC 4	*198*	*203*	*190*	*203*	*189*	*3*	*4*	*4*	*4*	*4*
	MC 5	2	0	2	0	0	0	0	0	0	0
Both	MC 1	29	36	26	36	20	0	0	0	0	0
	MC 2	24	20	21	20	18	0	0	0	0	0
	MC 3	20	22	19	22	11	0	0	0	0	0
	MC 4	7	3	2	3	2	0	0	0	0	0
	MC 5	*165*	*167*	*177*	*167*	*172*	*3*	*4*	*4*	*4*	*3*

5.8 DISCUSSION

The FS2FI algorithm segments the frequency spectra of benchmark and on-line signals into a number of segments which is based on the minimum number of features required to differentiate between all machine fault signatures. This is carried out by identifying the segment size at which all fault patterns are identifiable. The maximum FFT amplitude value at each segment is employed as a unique feature, and is compared with all the corresponding maximum FFT amplitude values of other machine conditions.

During the training and verification processes, all machine fault patterns were successfully recognised. This was achieved by segmenting the frequency spectra into six equal divisions (section 5.5). Single faults occurring in one component, as well as multiple faults occurring in multiple components, were utilised to accurately evaluate the performance of the algorithm.

Early detection of lubrication faults can extend the life of bearings, and hence reduce breakdowns. The AE technique was employed to detect the onset of bearing lubrication faults as it provides indicative data for friction and impact. On the basis of the performance evaluation results presented in section 5.5, the algorithm managed to detect all the lubrication faults emulated (MC 2, MC 7 and MC 8). However, the FS2FI algorithm failed to detect all the machine conditions in issue as it yielded a relatively low detection accuracy of 63%. This detection accuracy is considered low in comparison to other existing feature selection techniques such as Fisher score, ANN and SVM. The low detection accuracy mainly resulted from the fact that the algorithm was not properly trained as it only considers a single training dataset. Moreover, the algorithm does not produce multiple solutions with a different number of feature differences that can improve fault identification performance, and give full control to developers of CM systems. Another drawback is that the certainty of the results is undetermined as the

algorithm does not quantify the number of matching features between the on-line signal and the patterns of the benchmark signals.

In order to counteract the relatively low detection accuracy and avoid the drawbacks of the MFS2FI algorithm, a modified and enhanced algorithm (MFS2FI) was developed and introduced in section 5.6. The number of datasets was increased in order to improve the training and testing processes. Table 24 clearly shows that the MFS2FI algorithm yielded a detection accuracy of 100%, and that all machine conditions were successfully identified at a segment size of 108 kHz, though the overall confidence level was small at this large segment size (CL=1 feature).

Despite successful detection of all machine conditions in question at this large segment size, the small confidence level may adversely affect the certainty of fault classification. It can be observed that the overall confidence level is best at 1 kHz segment size (CL= 91 features). Therefore, although the computing time and cost are relatively high at 1 kHz, the use of this segment size is recommended with a view to improving the certainty of the detection of centrifugal compressor faults. In large scale CM systems, a trade-off between the computing time and the confidence level is required in order to find the most suitable segment size that provides a good confidence level at the lowest possible computing time and cost. However, the development and computing time of the proposed MFS2FI algorithm will remain small in comparison to AI-based classifiers (see chapter 6).

5.9 SUMMARY

This chapter has described a novel FFT-based segmentation and fault identification algorithm (MFS2FI) which is easy to implement, automated, non-AI, fast, and systematic. The algorithm is automated in that it identifies the best number of segments at which all fault patterns could be accurately recognized. However, the algorithm still

needs to be fed with machine fault signatures which should be obtained experimentally. For full scale LNG compressors, damage should be added deliberately to the equipment in order to collect signatures of common faults or one could wait for naturally occurring faults to occur and record there cause and the data. The more fault signatures collected the more accurate results the algorithm will produce.

This algorithm was developed to overcome the drawbacks of existing fault classifiers. The drawbacks of existing classifiers, as reported in the literature, are the relatively low accuracy of the non-AI classifiers, the complexity of the NN architecture design, the sensitivity to design parameters, and the high computing and development times of AI-based classifiers, even if some optimisation techniques are applied to ANN [80] (see section 2.5).

The MFS2FI algorithm yielded a fault identification accuracy of 100% and a good overall confidence level of 91 differences between the machine condition patterns identified (91 features out of 238). A potential benefit of the MFS2FI algorithm is that it can be used by non-specialist engineers as it does not require any detailed knowledge or experience of AI methods, making it easy to implement. The robustness of the algorithm proposed is investigated in Chapter 6. The investigation includes the impact of noise disturbance and machine degradation on the accuracy of fault detection and identification.

The scope of this chapter was extended further by investigating the suitability of the AE technique for the detection of machine faults of high speed centrifugal equipment. It demonstrated its effectiveness in identifying the machine conditions in issue.

Chapter 6. Assessment of robustness and performance comparison with standard FFT and neural network pattern recognition classifiers

6.1 INTRODUCTION

The fault classification performance of FFT-based fault identification algorithms varies according to data window time length, location and sampling frequency. Although longer data windows and higher sampling frequencies provide better information, they may cause difficulties in recognising fault patterns on-line due to the high computing time needed for training and testing [128]. A performance evaluation of fault identification algorithms should be carried out using datasets from new experiments and with different operating conditions in order to check their ability to classify faults.

As there are numerous existing fault detection algorithms, a comparative study should be carried out to check the competitiveness of new algorithms. Algorithms can be compared by evaluating their detection accuracy, robustness, computing and development time, and costs needed for successful online fault identification.

Therefore, the main contribution of this chapter is to investigate the robustness of the algorithm proposed in chapter 5. This will be achieved by evaluating performance changes arising from variation of data window time length, location, and rotational speed. Moreover, a comparative study will be carried out in this chapter to compare the

effectiveness of the proposed MFS2FI algorithm with other methods, namely the basic FFT classifier and the neural network-based classifier (see section 6.3).

6.2 ROBUSTNESS ASSESSMENT

A sensitivity analysis was performed to properly evaluate the robustness of the proposed algorithm. The sensitivity analysis introduced in this section includes the sensitivity of the results to data window size, location, and normal changes in rotational speed and machine degradation.

6.2.1 VARIATION OF DATA WINDOW LENGTH

The performance of the proposed MFS2FI algorithm was tested iteratively over 76 window lengths (or signal time lengths) using 10 datasets of two randomly selected fault conditions (MC 2 and MC 5). The full length of each of the measured signals which was utilised in the previous section was considered.

Figure 49 shows an illustrative figure of a variation in data window length. In order to extract the component frequencies at 122 kHz, the minimum sampling rate must be equal to or greater than 244 k Samples/s, based on Nyquest theory. Hence 250 K Samples, 10 K samples and 1000 K Samples were selected as the minimum window size, size increment, and maximum window size, respectively.

Table 26 presents the feature count in addition to the CL value of the first three window sizes, the latter of which are samples. It also shows the numbers of matching FFT features between the machine fault signatures (threshold values) and the signals which were utilised for testing at different window lengths.

Table 23: Fault Pre-Identification matrices and CL factors for 10 datasets using the first three window lengths. The segment size is 1 kHz (correct diagnosis in bold).

Machine condition		Leak (MC 2)					Both (MC 5)				
Data Set (DS) #		2	4	6	7	9	2	4	6	7	9
From data sample no. 1 to 250 k	MC 1	65	65	62	58	59	18	19	17	16	12
	MC 2	**141**	**155**	**144**	**145**	**152**	19	17	17	11	16
	MC 3	8	4	7	5	7	7	10	3	4	5
	MC 4	25	24	22	25	23	20	33	18	20	23
	MC 5	29	28	24	26	24	**173**	**150**	**185**	**181**	**173**
	CL	76	90	82	87	93	153	117	167	161	150
From data sample no. 1 to 260 k	MC 1	61	62	73	59	65	12	22	21	24	24
	MC 2	**135**	**147**	**145**	**145**	**154**	12	24	6	10	11
	MC 3	6	8	5	9	10	2	8	3	3	3
	MC 4	25	29	22	28	21	32	33	14	18	16
	MC 5	32	25	23	24	29	**178**	**155**	**193**	**167**	**178**
	CL	74	85	72	86	89	146	122	172	143	154
From data sample no. 1 to 270 k	MC 1	55	51	65	46	57	21	25	15	19	8
	MC 2	**143**	**156**	**150**	**156**	**142**	10	24	9	5	15
	MC 3	6	4	9	8	9	4	8	7	9	8
	MC 4	25	24	22	23	24	23	31	14	11	13
	MC 5	33	25	30	25	24	**173**	**156**	**186**	**171**	**185**
	CL	88	105	85	110	85	150	125	171	152	170

On the basis of the results shown in Table 26, the two machine fault conditions were successfully detected at all window lengths. Due to the impracticality of presenting the results for 76 window lengths, the prediction confidence level will be analysed statistically. The statistical mean (μ), standard deviation (σ), and maximum and minimum values of CL were calculated to evaluate and summarise the effect of the window length on the detection performance of the proposed fault identification algorithm using equations (15) and (16).

$$\mu = \frac{1}{N}\sum_{i=1}^{N} Xi \qquad (15)$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Xi - \mu)^2} \qquad (16)$$

where

X= CL Value

N= Number of samples

The mean, standard deviation, maximum and minimum values of the 76 overall Confidence Levels (CL) are 101.17, 12.79, 134 and 59, respectively. According to the Empirical rule for normal data [129], approximately 99.7 % of data lies within $\mu \pm 3\sigma$. Thus the CL value at any window length will fall in the range of 101.17 $\pm$ 38.37. An overall CL of 101.17 means that there is a minimum difference between the features of any two machine fault patterns of more than 49% (101.17/238). However, the worst value of the overall CL (59) is still good for the classification of all of the fault patterns. The overall CL value should preferably be greater than 40 feature differences, and greater than 20% of the total count of features (being equal to double the segment count).
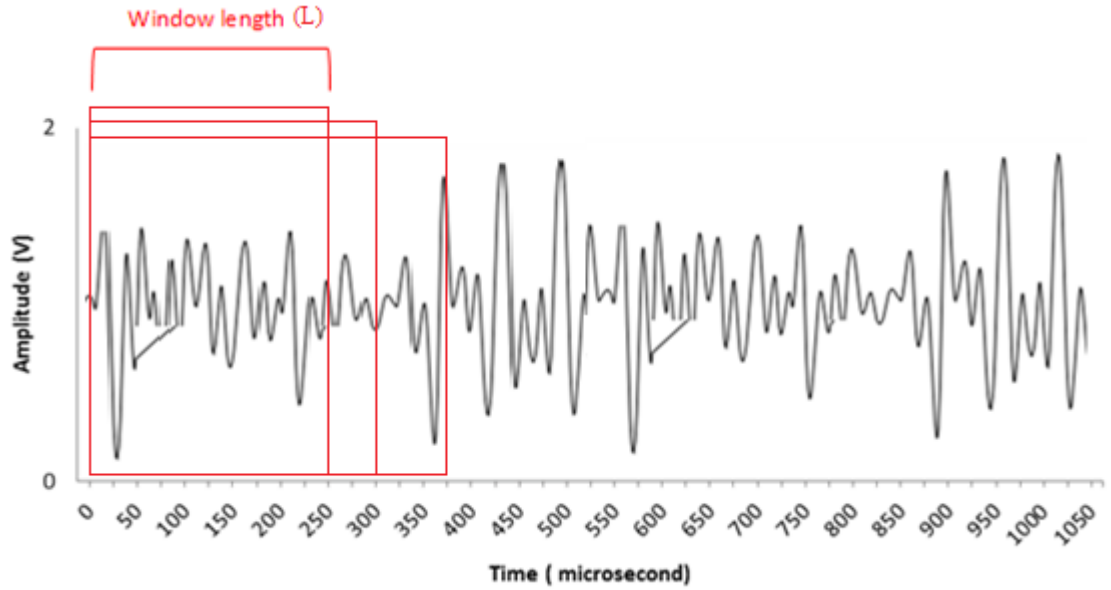
Figure 49: Illustrative figure for the variable data window length analysis at L = 250 kHz, 300 kHz and 375 kHz.

6.2.2 SLIDING DATA WINDOW ANALYSIS

As the measured signal changes over time during the data collection process, different datasets with different data window lengths and positions can be extracted from the basic signals captured. Thus, this section will study the effect of changing the window position on the performance of the proposed detection algorithm, on the basis that the signal was captured for a period of 1 second, at a sampling rate of 1 MHz, and with a data size of 1 mega samples. MC 2 and MC 5 represent the minor and the major fault conditions and hence they were selected for this sliding window analysis. MC 2 represents the leak fault and has slight variations in comparison to the healthy condition time signal; while MC 5 represents combined bearing faults.

Applying a sliding window, a window size of 250 k Samples was moved iteratively over the one million data samples. This was carried out by shifting the window by a time step (or offset) of 10 k samples.

Figure 49 depicts a sample of the analysed results of three sliding windows each with a

window length of 250 milliseconds and a sample data size of 250 K. All the machine conditions addressed were successfully detected at all window positions. Table 27 shows the numbers of matching FFT features between the benchmark machine fault signals (threshold values) and the signals which were utilised for testing at different data window positions.

Table 24: Fault Pre-Identification matrices for 10 datasets using the first three moving window positions. The segment size is 1 kHz (correct diagnosis in bold).

Machine condition		MC 2					MC 5				
Data set #		2	4	6	7	9	2	4	6	7	9
From data sample no. 1 to 250 k	MC 1	64	69	52	65	68	15	29	24	17	18
	MC 2	**131**	**142**	**164**	**145**	**144**	18	23	11	14	13
	MC 3	8	6	7	6	8	5	7	2	3	4
	MC 4	33	31	27	29	19	30	32	20	17	18
	MC 5	25	26	29	24	20	**171**	**139**	**185**	**179**	**178**
	CL	67	73	112	80	76	141	107	161	162	160
From data sample no. 10 k to 260 k	MC 1	61	72	54	71	56	19	24	17	16	17
	MC 2	**144**	**146**	**152**	**148**	**149**	17	22	12	15	14
	MC 3	6	5	6	8	8	9	8	3	5	4
	MC 4	32	33	27	27	26	27	36	23	19	21
	MC 5	31	22	29	25	29	**173**	**150**	**190**	**182**	**182**
	CL	83	74	98	77	93	146	114	167	163	161
From data sample no. 20 k to 270 k	MC 1	64	72	65	64	68	17	23	20	12	17
	MC 2	**140**	**150**	**151**	**151**	**144**	17	22	10	17	14
	MC 3	7	6	5	7	8	6	7	2	5	4
	MC 4	33	30	26	29	25	21	33	18	17	19
	MC 5	24	25	28	24	25	**173**	**153**	**189**	**193**	**180**
	CL	76	78	86	87	76	152	120	169	176	161

The complete set of results is very large and it is impractical to present such a large table. Therefore, the results are considered in statistical terms. The statistical mean (σ), standard deviation (μ) and the sample maximum and minimum values of CL were calculated to evaluate and summarise the effect of the window location on the performance of the proposed fault identification algorithm.

The mean, standard deviation, maximum and minimum values of CF are 87.28, 12.93, 120 and 54, respectively. Again, assuming 99.7% of data lies within $\mu \pm 3\sigma$, then the overall CL value across the one million data samples will fall in the range of 87.28 ± 12.93. The minimum value of CL (54) is good enough for a proper classification as it is still greater than both 40 feature differences and 20% of the total count of features (being equal to double the segment count).
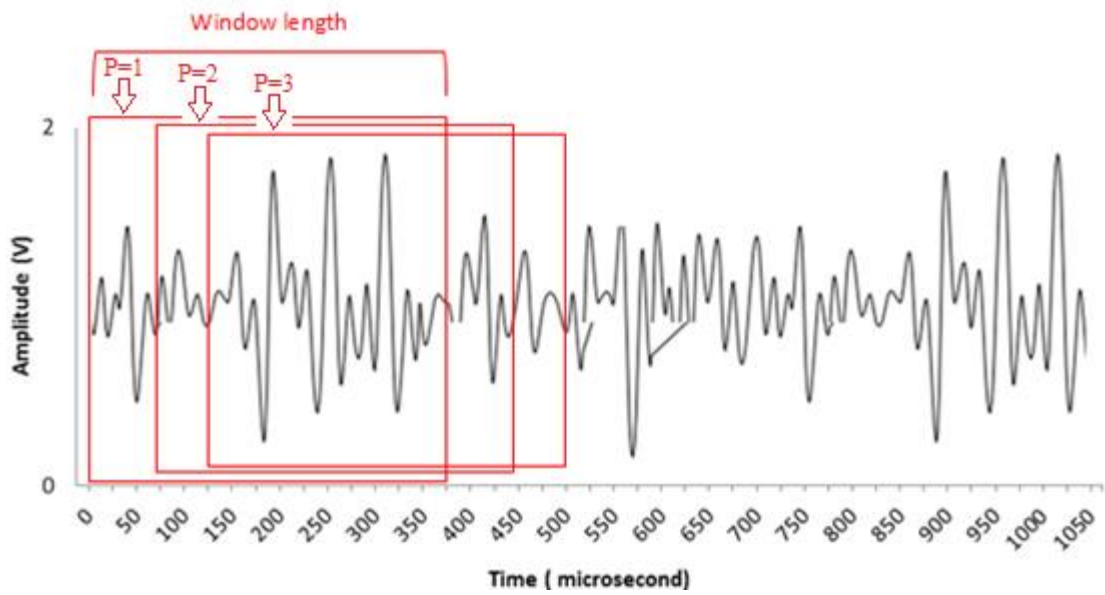
Figure 50: Illustrative figure for the moving window technique at three different data window positions (P= 1, 2 and 3).

6.2.3 VARIATION OF ROTATIONAL SPEED AND NOISE DISTURBANCE

A new experiment was carried out to check whether the algorithm has the ability to detect faults using datasets from new experiments and under different operating conditions. MC 1 and MC 3 were selected as case studies to check the effect of speed variation on the results of the proposed algorithm. The compressor was disassembled, bearings were installed, and the compressor was reassembled. The machine speed was reduced from

15,650 RPM to 14,480 RPM (-7.4%) before the data collection, on the basis that the normal operating rotational speed would be 15,650 RPM.

Table 28 shows the Pre-Identification matrix for MC 1 and MC 3 using data window lengths of 250 milliseconds and 1000 milliseconds at a reduced rotational speed of 14,480 RPM. It also compares the count of matching features at the reduced speed and the normal speed (see Table 22). It can be observed that the numbers of matching features and the overall CL significantly decreased with the reduction in rotational speed. The overall CL is the minimum difference between the highest number of matching features with the correct signal and the second highest number of matching features with the wrong signal. However, faults were successfully detected with a good overall confidence level of 60 in comparison to 117 at the normal operating speed. The change in signal time length was investigated and the results showed that this change had a very limited effect on the certainty of fault identification.

Table 25: The Pre-Identification matrices for MC1 and MC3 faults using signal lengths of 250 milliseconds and 1 second at a reduced rotational speed of 14480 RPM (correct diagnosis in bold). The count of the matching features at the normal operating rotational speed is in bracket.

	Signal length of 250 milliseconds	Signal length of 1 second @ 15650 RPM	Signal length of 250 milliseconds	Signal length of 1 second @ 15650 RPM
	MC1		MC3	
MC 1	**99**	**87 (180)**	7	34 (16)
MC 2	11	19 (63)	6	16 (17)
MC 3	21	11 (20)	**118**	**94 (169)**
MC 4	2	6 (3)	1	4 (2)
MC 5	8	24 (35)	9	28 (28)
CL		63 (117)		60 (141)

The results presented in Table 28 demonstrate the robustness of the MFS2FI algorithm as all the selected machine conditions were successfully identified at a reduced operating rotational speed of 14480 RPM at different signal lengths.

As noise affects the performance of fault identification performance, this section investigated the impact of noise coming from nearby mechanical equipment on the performance of the algorithm proposed. Any disturbance to a system (or noise coming from the mechanical equipment) usually creates "sound energy" resulting from an increase of friction between bearing components. According to the Shock-Pulse technique, these sounds produce a broadband noise (white noise) that contains all frequencies. . Therefore, a "broadband noise" was added to the MC#1, MC#3 and MC#4 time signals (dataset # 4) using "randn" MATLAB function in order to investigate the impact of noise on the fault identification performance of the MFS2FI. The "randn" function returns an array of numbers that are randomly drawn from a standard normal distribution whose mean is "0" and standard division is "1" (number of elements utilised: 1 Million, range: -5.2 to 4.99). The amplitudes of noise were calculated by multiplying the output of "randn" function by 5% of the maximum FFT amplitudes of the fault spectrum in issue. Based on the results shown in Table 29, the algorithm managed to identify all of the machine conditions in issue, but with smaller CL factor values (lowest CL= 29).

Table 26: The Pre-Identification matrices for noised MC1, MC3 and MC4 faults signals using DS#4 (correct diagnosis in bold). The count of the matching features of the original signals is in bracket.

	MC1	MC3	MC4
MC1	**109(182)**	22(13)	3(4)
MC2	59(63)	30(17)	2(2)
MC3	25(18)	**77(172)**	0(2)
MC4	2(3)	1(2)	**192(203)**
MC5	54(37)	48(29)	3(0)
CL	50(119)	29(143)	189(199)

6.2.4 SUMMARY OF ROBUSTNESS

The MFS2FI was checked for sensitivity using three different analyses. The analyses include the variation in data window length, the moving data window and the variation in rotational speed. In addition, the ability of the MFS2FI algorithm to identify degradation outside of the datasets for which it was trained was investigated.

The results showed that the performance of the proposed MFS2FI was not affected by the variations in data window length and the position of the data window as it successfully managed to identify all fault patterns at 76 different data window lengths and 76 data window positions (see Table 26 and Table 27). The algorithm was able to differentiate a minimum of 54 differences between the fault patterns of MC 2 and MC 5, which represents more than 22% of the total number of features. It also managed to identify degradation using new experimental datasets. However, the numbers of the overall CL significantly decreased with the reduction in rotational speed, from 117 differences to just 60 differences (see the full results in Table 28).

6.3 PERFORMANCE COMPARISON WITH OTHER METHODS

A MATLAB code was developed to evaluate the performance of the proposed classifier by comparing its results with the results obtained from applying two different existing

pattern classification methods to the same datasets, namely the non-AI FFT classifier and the ANN FFT classifier. Section 6.4 shows the results of the comparison between the proposed algorithm and a standard FFT classifier, while the results of the comparison between the proposed algorithm and the most utilised ANN-based classifier for pattern recognition are presented in section 6.5.

6.4 COMPARISON WITH A STANDARD FFT CLASSIFIER

The standard classifier developed in this chapter only considers the maximum peak amplitude feature in each FFT spectrum as a unique feature. Then, it compares the maximum peak amplitude of the testing signal with the benchmark peak amplitude range of each machine condition. The minimum and maximum benchmark threshold values were calculated on the basis of training datasets # 1, 3, 8 and 10 (for more information about the training datasets, see section 5.6.1). Columns 2 and 3 show the benchmark threshold ranges [Xmin, Xmax], while column 4 shows the maximum FFT amplitude values of the validation dataset number 5 (DS 5).

Table 30 compares the detection performance of the proposed algorithm with the detection performance of a standard FFT classifier. The classifier yielded a maximum detection accuracy of 60%, while the proposed algorithm yielded 100% detection accuracy at a segment size of 108 kHz. Moreover, the confidence level of the results of the proposed algorithm is measurable and controllable, while it is non-controllable in FFT classifiers.

Table 27: Benchmark threshold values of non-segmented frequency spectra (Xmin and Xmax matrices).

Machine condition	Xmin FFT amplitude	Xmax FFT amplitude	FFT amplitude DS#5	FFT classifier	Proposed algorithm
				Detectability	Detectability
MC1	1.33E+04	1.98E+04	1.29E+04	0	1
MC2	9.42E+03	1.49E+04	1.10E+04	1	1
MC3	4.58E+04	5.60E+04	3.71E+04	0	1
MC4	7.94E+04	8.58E+04	8.50E+04	1	1
MC5	2.31E+04	3.68E+04	2.45E+04	1	1
Performance				60%	100%

6.5 COMPARISON WITH A NN-BASED CLASSIFIER

This section aims to benchmark the performance of the proposed algorithm against NN-based classifiers. The most popular and well proven NN architecture, training algorithm, activation function, and error calculation method utilised for pattern recognition, were employed to carry out this comparison. Multi-Layer Perceptron (MLP) architecture was utilised in combination with the Scaled Conjugate Gradient-based (SCG) supervised learning algorithm and the Sigmoid Activation Function (SAF). The performance of the network was evaluated using the Mean Squared Error (MSE) quantitative measure at different neurons and hidden layer numbers. The following sections present the literature survey and the detailed comparison results.

6.5.1 INTRODUCTION TO NEURAL NETWORK

An Artificial Neural Network (ANN) is a mathematical or computational paradigm that is inspired by the information processing approach of the biological nervous system. It is made up of a large number of interconnected nodes (neurons). ANNs are adaptive, flexible, and configurable, and through a learning process, can be customised for specific applications, such as data classification, trend detection or pattern recognition. A trained

neural network performs as an expert system, and can answer a set of "what if" questions. The advantages of ANN techniques are (a) adaptive learning based on the data given, (b) self-organisation during learning, (c) real-time operation if the ANN computations are carried out fast and in parallel, and (d) fault tolerance if information redundancy is present [130].

The basic element of ANNs is the neuron. Artificial neurons have several inputs in addition to one output, as shown in Figure 51 , and they have two modes of operation: the training mode and the using mode. In the training mode, the neuron can be trained to fire an action for particular input patterns. In the using mode, when a taught input pattern is recognised at the input stage, its corresponding output becomes the current output of the network. If the input pattern is not recognised as one of the taught input patterns, the firing rules are utilised to determine whether or not to fire an action.
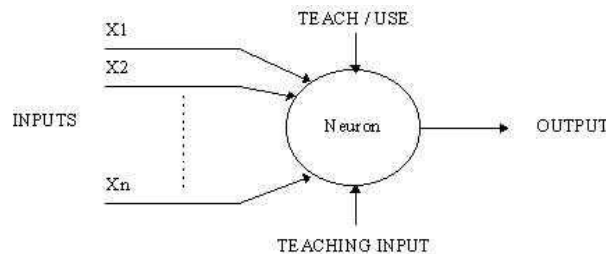
Figure 51: A simple Neuron [130].

A more advanced neuron is the McCulloch and Pitts model (MCP) [130] [131]. The difference between simple and advanced neurons is that the inputs of the latter are 'weighted'. Weighting is the process of multiplying an input by a factor which varies depending on the importance of this input. In decision-making, the effect that each input has is dependent on its weight. These weighted inputs are then summed and compared with a pre-set threshold value. The neuron fires an action if the sum of the weighted inputs exceeds the threshold value. Otherwise, the neuron will not fire. The firing rule is

an important concept in neural networks, and it determines whether or not a neuron should fire for any given input pattern. The decision is based on all input patterns, not only on the patterns on which the neuron was trained [130] [131]. Of course, it is important to cover a wide range of inputs during training so that the NN can give valid results.

6.5.2 ARCHITECTURE OF NEURAL NETWORKS

The major types of NN networks are Feedforward, Self-Organize, Recurrent, Stochastic, Modular and others such as Feedback (dynamic), NeuroFuzzy and Instantaneous Trained. The Feedforward neural network is the most widely used model, where signals are allowed to travel one way only, from input to output. Due to the absence of feedback signals, the output of each layer is not affected by the output of other layers. The Feedforward technique is the most widely used type in pattern recognition [132] [133] [134].

The major architectures of Feedforward neural networks can be divided into single-layer and multi-layer. Single-layer networks are based on the Completely Connected Perceptron architecture (CCP), while Multi-layer networks are based on the Multi-Layer Perceptron architecture (MLP), as shown in Figure 52. MLP is the most common neural network model, and it comprises three layers, namely the "input layer", the "hidden layer", and the "output layer". The increase in MLP hidden layers increases the complexity of the network as well as the difficulty of training [130] [131].
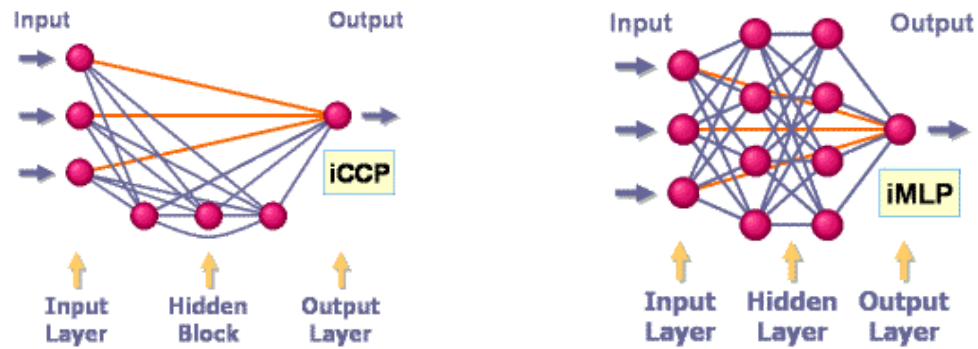
Figure 52: CCP and MLP Feedforward architecture of neural networks.

6.5.3 LEARNING PROCESS

Learning is one of the most important factors that affect the performance of ANNs. The behaviour of an NN network depends on the learning paradigm and the applied transfer and activation functions that determine the neuron's output

There are two major categories of learning methods utilised for adaptive neural networks: *unsupervised learning* and *supervised learning*. In unsupervised learning, the model is not provided with the historical system data during the training, and hence the network firstly self-organises the data it is presented with, and then detects their emergent properties.

Usually unsupervised learning is performed on-line and is utilised in data clustering. Hebbian and competitive learning rules are the major paradigms of unsupervised learning. In supervised learning, the model is provided with historical input and output data. The network utilises this data to calculate and then minimise errors in connection weights, which is a multivariate function that depends on the weights of connections in the network; the outcome is a minimum error between desired and computed values.

Paradigms of supervised learning include error-correction learning, stochastic learning and reinforcement learning. The most popular and robust NN learning algorithm is

"backpropagation", which is usually utilised in combination with a supervised error-correction learning rule. With backpropagation, the input data is repeatedly passed to the NN. In each cycle, the actual output of the NN is compared with the desired output, and an error is computed. In order to minimise this error, the connection weights are adjusted on the basis of the value of the returned (Back propagated) error [130] [135] [136]. This process is known as "training".

The following algorithms are the most common backpropagation algorithms and are utilised to either recognise patterns (discriminant analysis), or to approximate functions (regression): Levenberg-Marquardt, BFGS Quasi-Newton, Resilient Backpropagation, Scaled Conjugate Gradient, Conjugate Gradient with Powell/Beale Restarts, Fletcher-Powell Conjugate Gradient, Polak-Ribiére Conjugate Gradient, One Step Secant, and Variable Learning Rate Backpropagation.

Numerous tests were carried out in MATLAB (version 2012) to determine the fastest and most accurate algorithms for different applications. Although the Resilient Backpropagation (RP) algorithm is the fastest algorithm for pattern recognition problems, the Scaled Conjugate Gradient (SCG) algorithm is preferred as its performance does not degrade as quickly as that of RP when the error is reduced; in addition, the computing speed of SCG is almost as fast as RP for pattern recognition problems [136].

The performance of the SCG is affected by training, validation and testing ratios, random sets of data division, and the values of random weight initialisation (initial guess values). When a different combination of training patterns or initial weight values is used, the NN produces different classification results. Recent studies have shown that a proper selection of weight initialisation can significantly enhance the training process [137] [138].

The principal and most popular NN quantitative performance metric is the Mean Squared Error (MSE). The goal of the quantitative performance measure is to compare signals by describing the degree of similarity along with the level of error between them [139].

The three major transfer and activation functions are (a) linear (or ramp), (b) threshold, and (c) sigmoid. The linear activation function sets the output activity proportionally to the total weighted output. The threshold transfer function sets the output to be greater than or less than the threshold value. While the sigmoid transfer function continuously varies the output, but not linearly as the input changes. Although all three are considered rough approximations, the sigmoid transfer function is judged to have a greater resemblance to real neurons than linear or threshold transfer functions [140].

On the basis of the survey carried out on the NN architecture, learning methods, performance and also the most popular paradigm for pattern recognition, this study will employ the MLP model with a view to recognising the different fault patterns of the machine conditions addressed. The MLP will be used in combination with the SCG-based supervised learning algorithm and the Sigmoid Activation Function. The performance of the network will be measured using the MSE quantitative measure.

6.5.4 NN-BASED FAULT CLASSIFICATION PROCESS

Based on the literature presented in the previous sections, it can be concluded that the detection accuracy of NNs depends on (a) the training algorithm, (b) the architecture of the network (number of hidden layers and number of neurons), (c) the type of the activation function, (d) the values of the initial random weights, (e) the data division for training and validation, and (f) the training, validation and testing ratios.

Figure 53 illustrates the process employed to find a near-optimal fault classification result for the addressed machine fault conditions within the lowest possible computing time and within an acceptable overall development time.

The algorithm starts by reading all 50 training, verification and testing datasets. The selection of the NN architecture is carried out through the selection of the numbers of hidden layers and neurons. The NN network is trained by the backpropagation learning algorithm. After the random selection of the NN connections weights, the backpropagation learning algorithm is utilised to compute new sets of corrected weights in order to minimise the MSE error function. The learning process can be broken down into four steps: (a) feed-forward computation, (b) backpropagation to the output layer, (c) backpropagation to the hidden layer, and (d) weight updates. The learning process stops when the value of the mean squared error function has become sufficiently small. The error is the difference between input and output values [141].

The training of the back propagation algorithm is based on the minimisation of the mean squared value of the instantaneous error, as shown in equation (17) [142] [143].

$$E_{(MSE)} = \frac{1}{2}\sum_{q=1}^{n}(d_q - y_q)^2 \tag{17}$$

where d_q represents the desired network output for the qth input pattern and y_q is the actual output of the neural network. The NN weight update is carried out according to equation (18).

$$\Delta W_{ij} = -K\frac{\partial E}{\partial W_{ij}} \tag{18}$$

where k is a constant of proportionality, E is the MSE error function and W_{ij} represents the weights of the connection between neuron j and neuron i. The weight adjustment is iteratively repeated until the difference between the NN output and the actual output is within an acceptable tolerance.

The time taken to learn the NN and identify the pattern (computing time) in addition to the fault classification performance are calculated to evaluate the overall performance of the network. The fault classification performance is the ratio between the number of correctly identified fault patterns and the total number of testing patterns (25 patterns). The number of training times, neurons, and hidden layers are changed (within specific limits) to investigate the performance of a large number of multi-layer perceptron NN configurations. A near-optimal fault classification result is then selected on the basis of the highest pattern identification performance and lowest computing time.

In this study, the investigation of the near-optimal solution is carried out in three nested loops as shown in Figure 62. The first loop examines the rise of retraining cycles, the second loop investigates the increase of neurons number, while the third loop investigates the increments in the number of hidden layers.
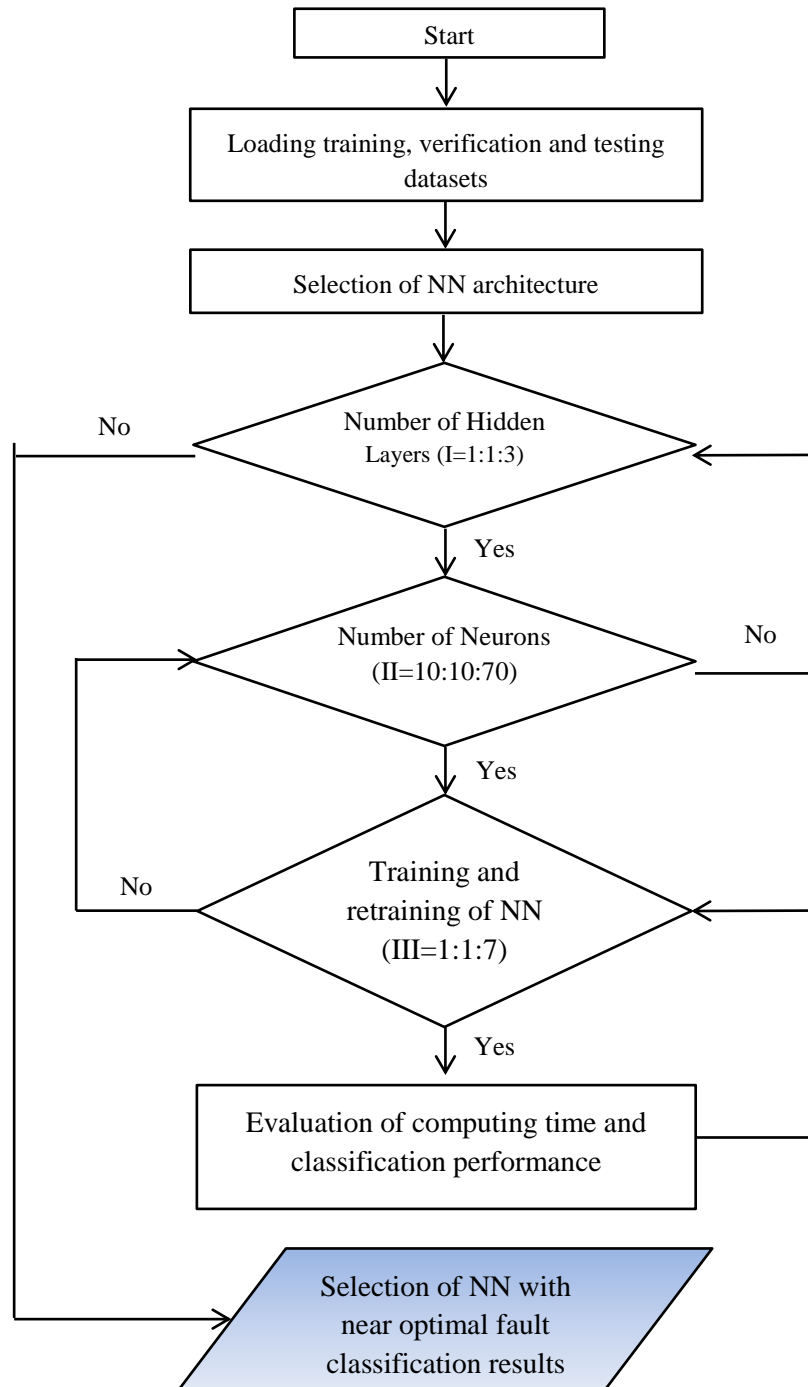
Figure 53: NN-based fault classification process.

In order to apply the same comparison criteria to both the MFS2FI algorithm and NN, the same training, validation and testing datasets and ratios were utilised in the training, verification and testing processes of both MFS2FI and NN. 50% of the data was utilised for training and validation, while the remainder was utilised for testing (see section

5.6.1). Combined NN configurations (hidden layer and neuron numbers) were also tested to investigate the effect of these combinations on the results, in terms of pattern recognition accuracy and computing time. As the NN results are sensitive to the initial values of weights, and in order to allow the reproduction of the same classification results every time the simulation is run, the weight initialisation (the initial guess values of the weights) for each run was controlled using the Random Number Generation (RNG) MATLAB function. This function helped to ensure an identical sequence of random numbers at the beginning of each simulation, just as if the MATLAB had been restarted. The NN performance was measured using the principal and most common quantitative measure, MSE.

6.5.5 PERFORMANCE COMPARISON RESULTS

The comparison is based on the fault identification performance and the times and costs related to both development and computing. The development time is the time taken by the CM system developer to investigate different NN architectures in order to optimise the fault identification performance and the computing time. The computing time includes the time taken by the PC to train (or retrain) the NN network, and to classify 25 fault patterns.

Table 31, Table 32, Table 33 and Table 34 show the results of the NN-based fault classification at different MLP NN configurations. As shown in Table 31, two NNs, each with a single hidden layer of 60 neurons and 70 neurons respectively, produced a 100% classification performance when retrained 6 times with total computing times of 784.7 and 1780.2 seconds, respectively. A 100% classification performance means that all of the 25 fault test patterns were successfully classified. A two hidden-layer neural network with 70 neurons in each layer produced a classification performance of 100% from the first run with a total computing time of 83.1 seconds as shown in Table 32.

In Table 33, the simulations showed that three NNs, each with three hidden layers (20 – 20 – 20 neurons, 60 – 60 – 60 neurons and 70 – 70 – 70 neurons, respectively) produced a 100% classification performance in 462.6, 330.6 and 501.7 seconds, respectively.

Simulations with different configurations of hidden layer numbers and neuron numbers were run to investigate the effect of decreasing the number of neurons in the second and third layers, as shown in Table 34. These results show that layers with similar numbers of neurons are more effective in terms of performance as none of the other configurations produced a 100% classification performance.

From the above results, it can be concluded that an NN with one hidden layers of 60 neurons is best for the classification of the machine fault patterns in question as the NN manged to classify all machine conditions in 57.2 seconds only. This investigation was conducted to identify a near-optimal NN configuration. It is apparent that the development time of an efficient NN-based CM system is lengthy and that the optimised configuration may not work for different patterns as the results may differ from one set of fault patterns to another; in turn, this may be due to the trial-and-error correlation learning method used in the training of NN.

Table 28: Classification results for a single-layer NN with different numbers of neurons (100% detection accuracy in bold).

Training cycle number	1	2	3	4	5	6	7	Best acc.	Total time
One hidden layer - 10 neurons									
Classification performance (%)	64	20	68	72	40	56	72	72	
Computing time (s)	73.9	8.4	10.5	6.3	11.4	10.9	6.5		127.9
One hidden layer - 20 neurons									
Classification performance (%)	20	84	36	20	68	20	80	84	
Computing time (s)	121.5	13.6	10.7	3.5	13.2	2.2	16.7		181.4
One hidden layer - 30 neurons									
Classification performance (%)	20	44	88	84	64	72	36	88	
Computing time (s)	43.5	17.8	29.2	22.5	12.1	34.7	5.7		165.5
One hidden layer - 40 neurons									
Classification performance (%)	92	80	84	96	60	84	88	96	
Computing time (s)	420.6	33.8	42	36.6	27	42.3	44.1		646.4
One hidden layer - 50 neurons									
Classification performance (%)	84	76	88	80	84	88	80	88	
Computing time (s)	491.3	41.4	45	22.8	56.9	47.6	26.6		731.6
One hidden layer - 60 neurons									
Classification performance (%)	88	80	92	92	96	80	**100**	100	
Computing time (s)	424.2	37.4	50.8	64.5	106	44.6	57.2		784.7
One hidden layer - 70 neurons									
Classification performance (%)	92	92	76	56	80	76	**100**	100	
Computing time (s)	1416.3	65	27.2	62.5	81.4	43.8	**84**		1780.2

Table 29: Classification results for a two-layer NN with different numbers of neurons (100% detection accuracy in bold).

Training cycle number	1	2	3	4	5	6	7	Best acc.	Total time
Two hidden layers - 10 neurons each									
Classification performance (%)	36	32	28	28	20	80	20	80	
Computing time (s)	45.2	4.5	8.4	4	8.4	10.9	7.8		89.2
Two hidden layers - 20 neurons each									
Classification performance (%)	44	60	60	36	84	28	80	84	
Computing time (s)	69.5	25.9	20.2	16	23.4	4.4	53.2		212.6
Two hidden layers - 30 neurons each									
Classification performance (%)	88	84	32	56	68	92	92	92	
Computing time (s)	492.2	47.6	5.7	41.9	32.3	32.9	33.1		685.7
Two hidden layers - 40 neurons each									
Classification performance (%)	40	52	48	20	88	80	88	88	
Computing time (s)	37	27.1	17.5	8.6	43.2	35.1	74.9		243.4
Two hidden layers - 50 neurons each									
Classification performance (%)	56	84	84	88	80	80	80	88	
Computing time (s)	21.8	41.8	34.6	88.8	55.7	79.9	60.6		383.2
Two hidden layers- 60 neurons each									
Classification performance (%)	92	96	84	96	92	80	84	96	
Computing time (s)	58.2	60.9	76.4	96.9	53.1	58.7	86.9		491.1
Two hidden layers - 70 neurons each									
Classification performance (%)	**100**	80	88	76	96	84	80	100	
Computing time (s)	83.1	68.4	69	57.15	68.7	93.7	52.5		492.55

Table 30: Classification results for a three-layer NN with different numbers of neurons (100% detection accuracy in bold).

Training cycle number	1	2	3	4	5	6	7	Best acc.	Total time
Three hidden layers - 10 neurons each									
Classification performance (%)	44	52	40	24	4	20	20	52	
Computing time (s)	61	17.1	7.1	3.5	4.1	8	2.4		103.2
Three hidden layers – 20 neurons each									
Classification performance (%)	**100**	64	48	92	92	48	20	100	
Computing time (s)	462.6	38.6	15.7	39.2	23.7	9.8	18.2		607.8
Three hidden layers - 30 neurons each									
Classification performance (%)	56	76	96	36	20	84	32	96	
Computing time (s)	29.5	24.6	24.4	11.5	3.6	34.5	8.3		136.4
Three hidden layers - 40 neurons each									
Classification performance (%)	52	84	48	84	84	80	80	84	
Computing time (s)	42	166.8	12.6	42	119.6	56.1	59.1		498.2
Two hidden layers - 50 neurons each									
Classification performance (%)	88	80	80	84	80	76	60	88	
Computing time (s)	47.3	67.9	60.3	48.4	52.7	97.4	17.1		391.1
Three hidden layers- 60 neurons each									
Classification performance (%)	88	96	76	**100**	68	84	92	100	
Computing time (s)	65.4	81	49.3	134.9	70.9	89.8	64.3		555.6
Three hidden layers - 70 neurons each									
Classification performance (%)	96	84	84	76	**100**	20	92	100	
Computing time (s)	236.4	80.7	69.8	66	48.8	88.4	132.8		722.9

Table 31: Classification results of NNs with combined configurations.

Training cycle number	1	2	3	4	5	6	7	Best acc.	Total time
Three hidden layers: 10 - 20 - 10 neurons									
Classification performance (%)	40	20	56	64	48	40	60	64	
Computing time (s)	318.2	12.3	49.7	36.1	7.9	11.6	18.5		454.3
Three hidden layers: 40 - 30 - 20 neurons									
Classification performance (%)	80	80	80	72	20	92	36	92	
Computing time (s)	85.5	57.7	90.2	48.6	25.6	51.6	51.6		410.8
Three hidden layers: 70 - 30 - 10 neurons									
Classification performance (%)	80	76	68	60	96	52	52	96	
Computing time (s)	74.8	171.8	114.7	86	68.2	68.7	14.5		598.7
Two hidden layers: 70 - 30 neurons									
Classification performance (%)	28	96	84	92	84	24	56	96	
Computing time (s)	19.2	79.9	75.5	56.6	75.9	7.2	25.3		339.6
Two hidden layers: 40 - 10 neurons									
Classification performance (%)	84	84	44	20	32	28	20	84	
Computing time (s)	713.9	47.3	15.5	50.1	8.3	7.1	19.2		861.4
Two hidden layers- 40 - 20 neurons									
Classification performance (%)	32	76	92	36	80	92	80	92	
Computing time (s)	104.1	47.6	117.1	10.7	31	54	38		402.5
Two hidden layers: 60 - 20 neurons									
Classification performance (%)	84	80	76	80	72	96	80	96	
Computing time (s)	117.4	44.1	34.5	69.2	57.7	101.9	44.2		469

6.5.6 DISCUSSION

The results presented in Table 31 to Table 34 show that a suitable neural network with optimised design parameters can successfully detect all of the machine conditions in question. The main drawback of NN is that the results are very sensitive to various design parameters; a large number of trials (196) is therefore required, and consequently a long development time.

These results support the findings of the literature review presented in Chapter 2. In respect of the computing time, NN managed to detect all machine conditions with a minimum computing time of 57.2 seconds. This was achieved using an NN with one hidden layers of 60 neurons. Another set of data may require a network with a different NN design configuration, and may produce results showing different computing and development times.

In the context of parallel processing, parallelising neural network training will significantly decrease the execution time of each run. G. Dahl et al. demonstrated that the utilisation of 8 parallel computing nodes speeds up the training of NNs by a factor of 11 [144]. However, if this technique is applied to an NN, it will not only decrease its execution time, but will also decrease the execution time of the proposed algorithm, and hence the proposed algorithm will still have lower execution and development times.

6.5.7 SUMMARY OF THE COMPARISON

The MFS2FI algorithm demonstrated a number of advantages over NN-based and non-AI classifiers, such as simplicity of implementation, low development time, and low computing time; it was seen to yield a fault identification accuracy of 100% in only 3.5 seconds in comparison to 57.2 seconds for NN. The non-complexity of the proposed algorithm gives it a significant advantage over AI-based technique as it can be implemented across all FFT spectra by non-specialised engineers. Unlike NN, the MFS2FI algorithm provides good measures of pattern similarity (CL) and results certainty (overall CL), and thus gives CM system developers control over the selection of segment size, and hence guarantees certainty of results.

6.6 OVERALL SUMMARY

The robustness of the MFS2FI algorithm in classifying faults was investigated in this chapter by evaluating the fault identification performance of the algorithm at 76 different signal lengths ranging from 250 milliseconds to 1 second, and at 76 different locations by gradually sliding a data window of 250 milliseconds in length across the one second data sample with iterative increments of 10 milliseconds (see section 6.2). The results demonstrate the robustness of the proposed algorithm as they managed to classify all of the machine conditions addressed at different data window positions and at different signal lengths. The variation in data window length and position has a limited effect on the certainty of the fault identification process as the lowest percentage difference between fault patterns was still larger than 20% (for full results, see sections 6.2.2 and 6.2.3). New experiments were conducted to evaluate the performance of the MFS2FI algorithm to identify degradation outside of the datasets for which the algorithm was trained. The algorithm demonstrated its ability to consider the degradation of machine condition as it successfully managed to identify the selected machine conditions at a reduced rotational speed using datasets from new experiments.

The MFS2FI was compared with a detailed study of a standard FFT classifier and an NN-based classifier (sections 6.3, 6.4 and 6.5). The results demonstrated a number of advantages over NN-based fault classification methods as it successfully yielded a 100% fault identification accuracy using a set of 25 testing fault patterns in only 3.5 seconds (computing time) in comparison to 57.2 for NN. The computing time included the training and fault classification times.

It can be observed that NN is more than sixteen times more computationally intensive than the proposed algorithm which in turn, unlike NN, provides a pattern conformity measure (CL). The proposed FFT-based fault identification algorithm is easy to implement,

systematic, and requires low computing and development times, all of which make it favourable for CM of centrifugal equipment. The simplicity of the implementation of this algorithm on a new system is a significant advantage as it does not require any detailed knowledge or experience. Figure 54 displays and highlights the results of the comparison between both fault classification methods.
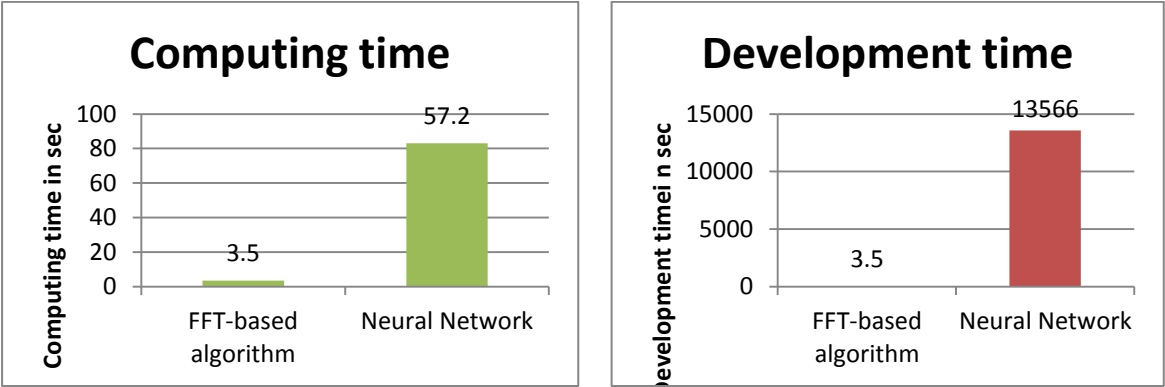
Figure 54: Performance comparison results between the proposed MFS2FI algorithm and a Neural Network-based method.

Chapter 7. Characterisation of major fault detection features and techniques

<hr />

7.1 INTRODUCTION

In chapter 5, a new frequency domain feature selection algorithm (MFS2FI) was developed for detecting faults in rotating equipment. The MFS2FI was tested on data from a centrifugal compressor (in the laboratory). Then, in chapter 6, the algorithm was tested for robustness, and compared with a neural network fault classifier. The results of the performance and robustness analyses demonstrated the effectiveness of the developed algorithm for fault identification.

In chapters 5 and 6, the MFS2FI algorithm was applied to acoustic emission (AE) signals only; this was because the literature review had suggested that they would be the most promising candidate signals. In this chapter, two further aspects of fault identification in centrifugal compressors are investigated in two ways: 1) using two additional signal types – vibration (acceleration) and process variable (pressure); and 2) the application of other FDI approaches from the literature, namely RMS, Crest Factor, Energy and Maximum Amplitude. This will facilitate the comparison of the AE technique with such other approaches and, most importantly, will make it possible to combine approaches to improve fault identification accuracy.

As demonstrated in the literature survey in this study, the most widely utilised condition monitoring techniques are AE, vibration and process information (such as pressure information). RMS, crest factor, energy and maximum amplitude are the major signal

features utilised for fault pattern recognition. As mentioned, it was suggested that the AE technique was more effective than other techniques for CM of rotating equipment. The literature survey also identified that the utilisation of multi-fault detection techniques maximises the accuracy of fault diagnosis, and provides an integrated system for the detection of rotating equipment faults.

In this chapter, an algorithm "characterisation table" is developed on the basis of experimental results in order to provide an integrated solution to fault detection challenges such as fault interference and centrifugal equipment noise. The results are presented for the 5 machine conditions discussed. (For more information about the five MCs, see section 5.6.1.) They are then compared for the purposes of characterising the best fault detection techniques and features. This will involve calculating and comparing 11 signal features using three different fault detection techniques, namely a) RMS, crest factor, energy, maximum amplitude and frequency spectral of vibration signals, b) RMS, crest factor, energy, maximum amplitude and frequency spectral of AE signals, and c) the average of pressure signals.

The main contribution of this chapter is to identify the pros and cons of the approaches investigated, to improve the CM performance of typical centrifugal equipment, and to avoid false alarms due to noise and fault interference; these aims are to be achieved through a fully integrated CM tool based on a multi-signal, multi-feature fault identification approach.

7.2 ACOUSTIC EMISSION ANALYSIS

As in chapter 5, the training datasets numbered DS#1, DS#3, DS#8 and DS# 10 were used for the determination of threshold values of different AE signal features. The AE signals were measured using a bearing (A) R6a AE sensor. A MATLAB program was

developed to calculate the four time domain features: RMS, Max. Amplitude, Crest Factor (CF) and Energy (E), as per equations (19), (20), (21) and (22) [145] [146]. The time domain signals were also converted to frequency domain signals, and the AE FFT features were selected using the MFS2FI algorithm. This gives a total of 5 "features" for fault identification.

$$X_{RMS} = \sqrt{\frac{\Sigma_1^N (X_n)^{\wedge}2}{N}} \tag{19}$$

$$CF = \frac{|X|_{max}}{X_{RMS}} \tag{20}$$

$$Max\ Amplitude = |X|_{max} \tag{21}$$

$$E = \sum_{-\infty}^{\infty} |X_n|^2 \tag{22}$$

For the time-domain results below, the maximum and minimum values of each parameter (feature) across the four training sets are used to set the threshold levels for fault identification.

RMS

Table 35 shows the RMS values for the training datasets. Across the top are the datasets and the minimum and maximum values, and down the side are the five machine conditions. It can be seen that, with the thresholds set to the minimum and maximum values, the RMS values can be used to detect MC 1, MC 2 and MC4. However, the RMS feature cannot be utilised to detect MC 3 and MC 5 because their threshold levels intersect, which in turn means that the faults are unidentifiable. The difference between the threshold ranges of MC 1 and MC 2 is small, and this may result in inaccurate fault detection.

Table 32: AE RMS values (in volt).

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	0.294	0.310	0.310	0.306	0.294	0.310
MC 2	0.342	0.352	0.325	0.321	0.321	0.352
MC 3	0.575	0.611	0.607	0.603	0.575	0.611
MC 4	2.824	2.834	2.552	2.468	2.468	2.834
MC 5	0.713	0.593	0.470	0.457	0.457	0.713

Maximum Amplitude

Table 36 shows the maximum signal amplitude threshold values of the training datasets. Based on these threshold values, the maximum amplitude feature can be used to separate all of the machine fault conditions addressed. However, again the separation of the threshold ranges of MC 1 to MC 2 and MC 3 to MC 5 is difficult as the maximum amplitude of MC 1 is very close to the minimum amplitude of MC 2, and the maximum amplitude of MC 5 is very close to the minimum amplitude of MC 3. These small differences (less than 1%) could adversely affect the accuracy of fault identification.

Table 33: AE maximum amplitude values (in volt).

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	1.347	1.445	1.389	1.273	1.273	1.445
MC 2	1.917	2.037	1.564	1.635	1.564	2.037
MC 3	4.622	4.494	4.520	5.015	4.494	5.015
MC 4	10.512	10.555	10.207	9.979	9.979	10.555
MC 5	4.465	3.638	3.521	3.588	3.521	4.465

Crest Factor

Table 37 shows the training crest factor threshold values for the datasets. The crest factor is equal to the RMS value divided by the maximum amplitude of the same signal. Based on the threshold values, the crest factor feature can be utilised to identify MC 1, MC 2,

and MC 4. This time series feature cannot be utilised for the detection of all of the addressed faults because of the intersection (overlap) between the crest factor intervals of MC 3 and MC 5. The difference between the crest factor values of MC 1 and MC 4 is small and this could negatively affect the accuracy of fault identification.

Table 34: AE crest factors.

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	4.585	4.662	4.486	4.165	4.165	4.662
MC 2	5.602	5.785	4.816	5.097	4.816	5.785
MC 3	8.044	7.351	7.449	8.319	7.351	8.319
MC 4	3.723	3.724	3.999	4.043	3.723	4.043
MC 5	6.262	6.137	7.493	7.845	6.137	7.845

Energy

Table 38 shows AE energy threshold values of the training datasets. On the basis of these threshold values, the energy feature can be utilised to differentiate between MC 1, MC 2 and MC 4. The main drawback is that the energy signal feature cannot be utilised to differentiate between MC 3 and MC 5 due to the intersection between their energy value intervals. The difference/ difference between the threshold values of MC 1 and MC 2 is small (in relative terms), and this could negatively affect the accuracy of fault detection.

Table 35: AE energy values (in Joule)

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	86,261	96,088	95,869	93,358	86,261	96,088
MC 2	117,079	124,004	105,439	102,940	102,940	124,004
MC 3	330,158	373,780	368,167	363,459	330,158	373,780
MC 4	7,973,092	8,031,877	6,514,243	6,093,149	6,093,149	8,031,877
MC 5	508,308	351,403	220,829	209,173	209,173	508,308

Figure 55 illustrates the differences between different time domain features. Although the energy feature is best by comparison to others, it is very difficult to employ it to differentiate between MC 3 and MC 5.
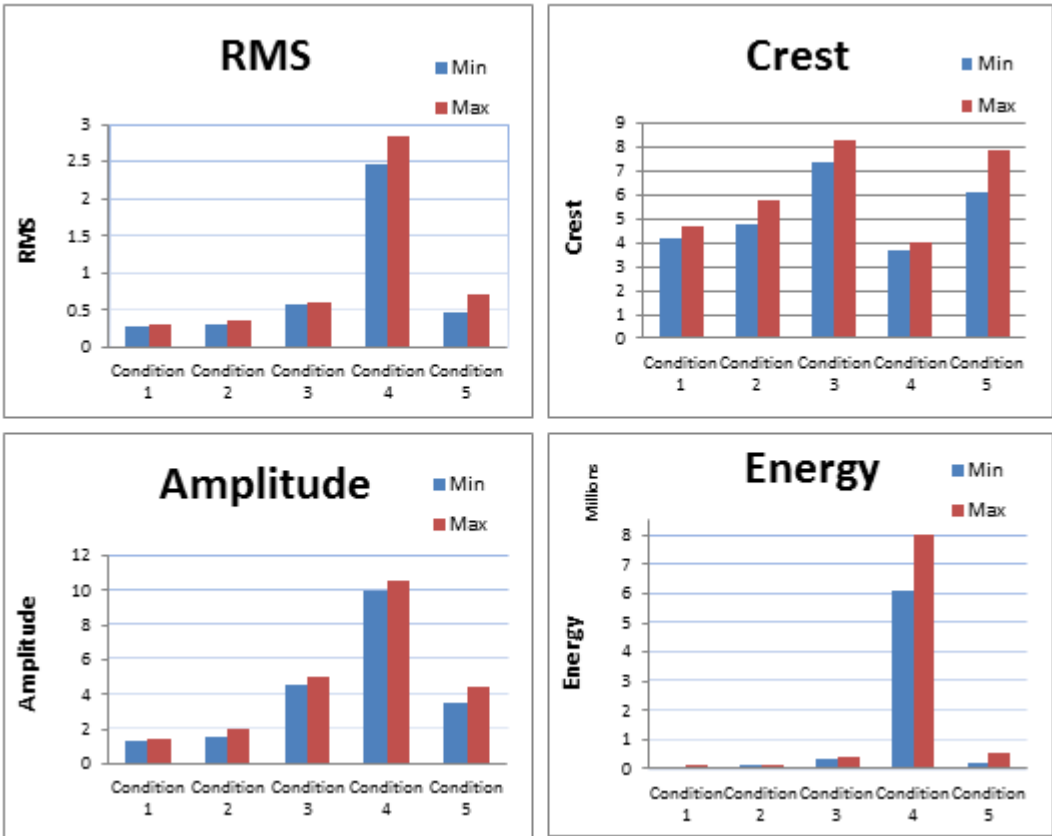
Figure 55: Graphical presentation for the AE RMS, amplitude, crest factor and energy values.

The FFT-based segmentation algorithm proposed in chapter 5 was utilised to investigate the suitability of AE spectral features for the detection of machine conditions using datasets number 1, 3, 8, 10 in the training cycle, and 5 in the validation cycle (see section 5.6.1 for full information about the datasets and the machine conditions emulated). Moreover, the algorithm investigates the segment sizes that can be used for efficient classification. The selection of the most suitable segment size depends on the overall

confidence level required for the results [14]. Table 39 shows the detection accuracy at different segment sizes where (0) means that the fault cannot be detected at this segment size, while (1) means that the fault can be detected. The addressed machine conditions were identified at the 108 kHz segment size with an overall confidence level of 1 out of 4. However, at this small overall confidence level, the accuracy of fault identification could be negatively affected as the pattern classification is carried out on the basis of the value of a single feature. The value of this feature (FFT amplitude) can vary during the operation of the machine due to noise, degradation or fault interference. At a segment size of 1 kHz, all of the machine conditions addressed were successfully identified with an overall confidence level of 91 out of 238 (for the detailed results, see Table 24). The accuracy of pattern classification at this large overall CL number is significantly better because at this segment size at least 91 feature differences would exist between any two fault patterns. The higher the value of the overall confidence level, the better the fault identification accuracy.

Table 36: Detectability of machine fault patterns using FFT AE spectra at different segment sizes (correct diagnosis in bold).

S	MC 1	MC 2	MC 3	MC 4	MC 5
121 kHz	0	1	0	1	0
120 kHz	0	1	0	1	0
119 kHz	0	1	0	1	0
118 kHz	0	0	1	1	1
117 kHz	1	1	0	1	1
116 kHz	1	1	0	1	1
115 kHz	0	1	0	1	1
114 kHz	0	1	1	1	1
113 kHz	0	1	1	1	1
112 kHz	0	1	1	1	1
111 kHz	0	1	1	1	1
110 kHz	0	1	1	1	1
109 kHz	0	1	1	1	1
108 kHz	**1**	**1**	**1**	**1**	**1**
107 2 kHz	**1**	**1**	**1**	**1**	**1**
1 kHz	**1**	**1**	**1**	**1**	**1**

7.3 VIBRATION ANALYSIS

Axial vibration was found to be more informative than radial vibration signals (see Appendix (E) for the detailed comparison). Below, axial vibration signals are analysed in the same way the AE signals were in section 7.2. The RMS, Max. Amplitude, Crest Factor and Energy are calculated in the time domain, and the frequency domain FFT features are extracted.

RMS

The RMS vibration threshold values of the training sets shown in Table 40 can only be used for the detection of MC 3. The RMS values of all the other machine conditions are in a very tight range, and this prevents the utilisation of this feature for the identification of centrifugal compressor machine conditions.

Table 37: Vibration RMS values (in volt).

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	1.663	1.655	1.633	1.629	1.629	1.663
MC 2	1.660	1.660	1.662	1.662	1.660	1.662
MC 3	1.872	1.853	1.828	1.823	1.823	1.872
MC 4	1.668	1.653	1.628	1.626	1.626	1.668
MC 5	1.685	1.668	1.636	1.630	1.630	1.685

Amplitude

The maximum amplitude threshold values of all vibration signals are almost equal, as shown in Table 41. Thus, this signal feature cannot be utilised to differentiate between the different machine faults.

Table 38: Vibration maximum amplitude values (in volt).

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	4.988	4.996	4.990	4.997	4.988	4.997
MC 2	4.983	4.978	4.997	4.972	4.972	4.997
MC 3	4.999	4.995	4.996	4.994	4.994	4.999
MC 4	4.993	4.997	4.992	4.997	4.992	4.997
MC 5	4.991	4.998	4.992	4.975	4.975	4.998

Crest Factor

The vibration crest factor threshold values of the training sets shown in Table 42 can only be utilised for the detection of MC 3. All other machine conditions have overlapping

tolerances, which prevents the use of this feature for the identification of compressor machine conditions.

Table 39: Vibration crest factors.

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	2.999	3.018	3.056	3.068	2.999	3.068
MC 2	3.003	2.999	3.008	2.992	2.992	3.008
MC 3	2.671	2.696	2.733	2.740	2.671	2.740
MC 4	2.994	3.023	3.066	3.073	2.994	3.073
MC 5	2.962	2.997	3.051	3.052	2.962	3.052

Energy

The calculated energy values of the training sets shown in Table 43 can be only utilised for the classification of MC 3. All other machine conditions have overlapping tolerances which prevent the use of this feature for the detection of centrifugal compressor machine condition.

Table 40: Vibration energy values 9 (in Joule).

	DS#1	DS#3	DS#8	DS#10	Min	Max
MC 1	922,397	913,554	889,123	884,541	884,541	922,397
MC 2	918,251	918,949	920,443	920,779	918,251	920,779
MC 3	1,167,622	1,144,503	1,114,004	1,107,512	1,107,512	1,167,622
MC 4	927,033	911,303	883,697	881,439	881,439	927,033
MC 5	946,480	927,465	892,592	885,805	885,805	946,480

Figure 56 illustrates the difference between different vibration time domain features. From the figure, it can be seen that the vibration time domain features can significantly detect MC 3, on the basis of a combination of RMS, CF and Energy. The features of the other machine faults interfered, and so could not be utilised for the detection of their corresponding faults.
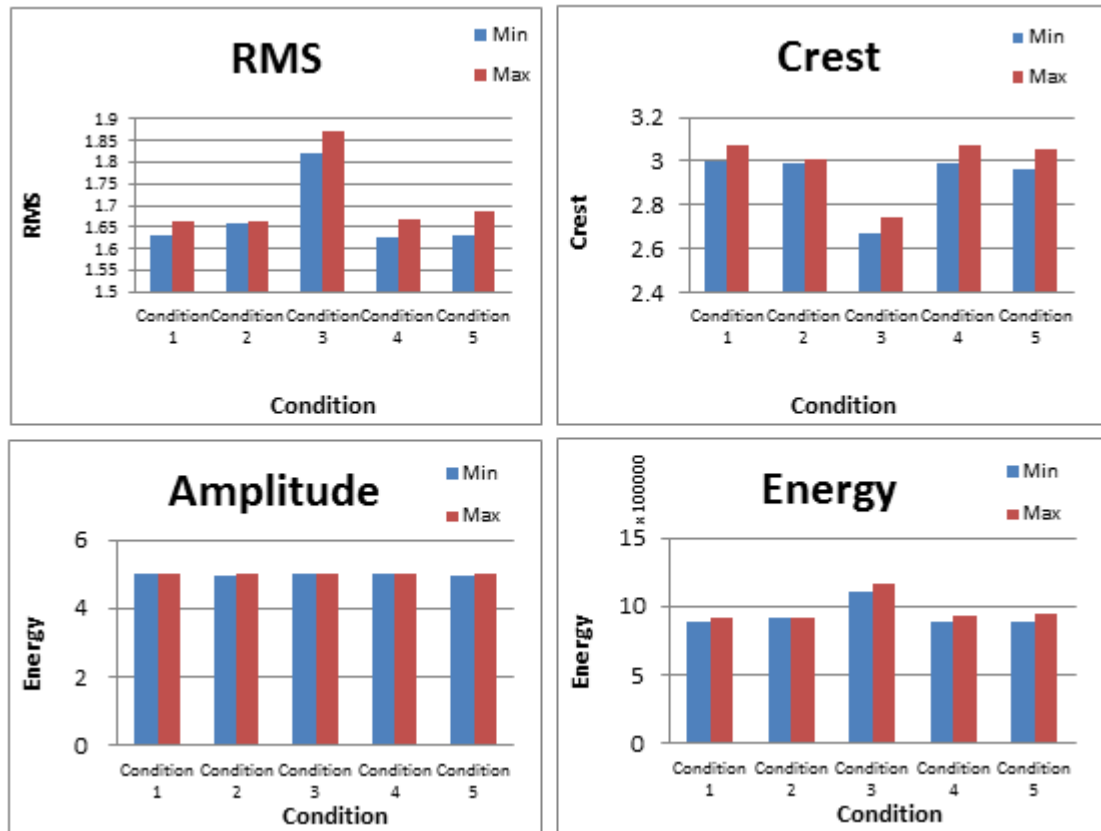
Figure 56: Graphical representation of the vibration RMS, amplitude, crest factor and

energy features.

The FFT-based segmentation algorithm proposed in Chapter 6 was utilised to investigate the suitability of the vibration spectral features for the identification of the machine conditions addressed. Table 44 shows the detectability of all of the addressed machine conditions at different segment sizes. The training was carried out using dataset numbers 1, 3, 8, 10 in the training cycle, and dataset number 5 in the validation cycle (for more information about the classification of datasets, see section 6.3). The addressed machine conditions were successfully identified with 100% detection accuracy at a segment size of 10 Hz and an overall confidence level of 72 out of 400.

Table 41: Segmented FFT vibration spectra (correct diagnosis in bold).

S	MC1	MC2	MC3	MC4	MC5
2000 Hz	0	1	1	1	1
1900 Hz	0	1	1	1	1
1800 Hz	0	0	1	0	1
1700 Hz	0	0	1	0	0
1600 Hz	0	0	1	0	0
1500 Hz	0	0	1	1	0
1400 Hz	0	0	1	1	0
1300 Hz	0	1	1	1	0
1200 Hz	0	1	1	1	0
1100 Hz	0	1	1	1	0
1000 Hz	1	1	1	1	0
900 Hz	**1**	**1**	**1**	**1**	**1**
800 Hz	0	0	1	0	0
700 Hz	1	1	1	1	0
600 Hz	1	0	1	0	1
500 Hz	1	0	1	1	1
400 Hz	1	0	1	1	1
300 Hz	**1**	**1**	**1**	**1**	**1**
200 Hz	**1**	**1**	**1**	**1**	**1**
100 Hz	**1**	**1**	**1**	**1**	**1**
10 Hz	**1**	**1**	**1**	**1**	**1**

Taking all of the above into account, the time-domain approaches will not work in isolation for vibration signals. The frequency domain methods are needed alongside the time-domain ones in order to identify all the MCs addressed.

7.4 PROCESS INFORMATION (PRESSURE) ANALYSIS

The RMS pressure technique was selected for investigation as it is one of the major process information techniques used for centrifugal equipment. The pressure information was collected using a pressure sensor installed in the air outlet tube. Four pressure readings were collected over a time period of four seconds for each machine condition. The average pressure was calculated on the basis of these four pressure samples. The average air outlet pressures for all MCs are shown in Table 45. The difference values between the average pressures of all MCs are less than 1.5%, except for the value of MC 2. Therefore, based on the analysis results, the pressure information can be only utilised for fault identification of MC 2.

Table 42: Air outlet RMS pressure (Correct diagnosis in bold).

	Average pressure (BarA)
MC 1	1.165
MC 2	**1.067**
MC 3	1.161
MC 4	1.157
MC 5	1.150

7.5 TESTING AND DISCUSSION

In this section, the algorithms developed and "trained" previously are tested on new (unseen) data. Fifty AE samples, fifty vibration samples, and fifty pressure samples were collected for all of the addressed machine conditions (see section 6.2). 44% of the samples were utilised to identify the benchmark thresholds, and 56% were utilised for the evaluation of detection accuracy.

A MATLAB code was developed on the basis of the benchmark thresholds shown in sections 7.2, 7.3, and 7.4 to evaluate the detection accuracy of a large number of signal

features and techniques using 30 datasets (60% of all datasets). The ability of the RMS and energy features of AE signals to detect MC 1, MC 2, MC 4 was proven, with a detection accuracy of 100%. The crest factor and amplitude features only detected MC 4, with a detection accuracy of 100%. The main drawback of the AE time-domain features is that MC3 and MC5 are always undetectable. The AE spectral feature proved its effectiveness over time-domain features as it was successfully utilised to detect all faults at any segment size smaller than or equal to 108 kHz, with a detection accuracy of 100%.

RMS, amplitude, crest factor and energy features of vibration signals proved their efficiency in detecting machine MC 3, with a detection accuracy of 100%. The main drawback of the vibration time-domain features is that MC 1, MC 2, MC 4, and MC 5 are undetectable. The vibration spectral features failed to detect all of the addressed machine conditions at all segment sizes except at 10, 100, 200, 300 and 900 Hz. The maximum overall confidence level was found at the smallest segment size (10 Hz). The vibration spectral feature technique proved its effectiveness over time-domain features as it successfully detected all of the addressed machine conditions at a segment size of 10 Hz, with a detection accuracy of 100%.

The pressure information proved its efficiency in detecting MC 2 with a detection accuracy of 100%. The main drawback of this technique is that the pressure information did not provide enough information for the detection of the other addressed machine conditions.

7.6 CHARACTERISATION TABLE

In this section, the combination of all of the above methods is investigated to provide a multi-fault detection technique for the CM of the centrifugal compressor. Eleven features were extracted for each machine condition. The characterisation table shown in Table 46

was built on the basis of the experimental results using both training and validation datasets. "Weak" means that the minimum difference between benchmark threshold of this machine condition and the benchmark thresholds of other machine conditions (or the value confidence level) is less than or equal to 10%; "Good" means that the minimum difference is greater than 10% but less than 20%; "Strong" means that the minimum difference is greater than 20% but less than 30% ; "Very Strong" means that the minimum difference is equal or greater than 30%; while "x" means that the signal feature failed to identify the correct machine condition.

A classification program was developed using MATLAB, based on the illustrated characterisation table, and was found to give a detection accuracy of 100%. As suggested by the literature, these results show that the use of multi-detection and multi-feature techniques has the potential to minimize false detections caused by fault interference and noise issues.

Table 43: Characterisation table for the classification of centrifugal compressor faults.

Technique	MC1	MC2	MC3	MC4	MC5
AE					
RMS	Weak	Weak	x	Very Strong	x
Amplitude	x	x	x	Very Strong	x
CF	x	x	x	Weak	x
Energy	Weak	Weak	x	Very Strong	x
FFT @S <90 kHz	Very Strong	Very Strong	Very Strong	Very Strong	Very Strong
Axial Vibration					
RMS	x	x	Good	x	x
Amplitude	x	x	x	x	x
CF	x	x	Weak	x	x
Energy	x	x	Good	x	x
FFT@ S= 10 Hz	Good	Good	Good	Good	Good
Pressure					
RMS	x	Very strong	x	x	x

7.7 SUMMARY

In this chapter, acoustic, vibration and pressure sensor data from the compressor have been analysed using five approaches, namely RMS, Crest Factor, Energy, Maximum Amplitude and FFT. This will allow comparison of the AE technique with the approaches stated and, most importantly, will make it possible to combine the approaches to improve the fault identification.

The lab-based industrial air compressor was employed to emulate five different centrifugal equipment machine conditions. The different techniques for fault detection of centrifugal equipment were investigated and compared in terms of their fault

identification ability, using data from the experiments. The AE time-domain features failed to identify the five addressed machine conditions, while the AE frequency-domain features managed to identify them all, with an identification accuracy of 100%. Moreover, the segment sizes with small overall confidence level values proved their effectiveness in detecting the addressed machine conditions. This means that these small overall confidence level values were enough for effective classification. However, the bigger the overall confidence level value, the better the elimination of false indications.

The vibration time-domain features failed to identify the five addressed machine conditions, while the vibration frequency-domain features managed to identify them all, at a segment size of 100 Hz, and with an identification accuracy of 100%. Although the overall confidence level values of vibration spectral features was relatively low in comparison with the confidence level values of AE spectral features, all of the addressed machine fault conditions were successful in their diagnoses, with a fault identification accuracy of 100%.

The pressure information was useful in the detection of air leakage problems (condition 2). For other machine conditions, the classification program failed to differentiate between other machine conditions with the use of the pressure information.

In conclusion, based on the results of this chapter, the characterisation table (see Table 46) provides the CM system developer with a full-capacity system for monitoring centrifugal equipment for the first time. The AE technique proved its effectiveness over vibration and pressure information techniques. In comparison to time-domain features, the FFT spectral features were best for the detection of high-speed centrifugal air compressor faults. The proposed characterisation table yielded a fault identification accuracy of 100%, and will improve the fault identification performance of a full-

capacity CM system for centrifugal equipment, as well as reducing the potential for false fault identification.

Chapter 8. Conclusion and future work

The ultimate, real-world, aim of this study was to suggest modifications to LNG production plants and their operations that could improve the profitability and availability of those plants, increase safety and reduce air pollution. Progress toward this has been achieved by using a combination of reliability and availability analysis and by developing (and assessing the performance of) an effective and non-complex CM system which, in the future, could form the foundation of a condition-based maintenance system. Taken together, these approaches allow improved availability with lower cost of maintenance.

The literature review on this subject showed that very little research has been reported on the use or development of C3MR liquefaction system redundancy schemes. It was also suggested in the literature that the accuracy of many available on-line CM methods for centrifugal equipment was inadequate, or that some methods could not be easily implemented as they required a long setting-up time and high development and computing costs.

Therefore, in chapter 3 of this study, the reliability and availability of the C3MR liquefaction system was analysed, and the introduction of redundancy was outlined on the basis of a detailed cost analysis. A Condition Monitoring (CM) system was proposed and experimentally verified in chapters 5, 6, and 7. The CM system was based on a novel Modified FFT-based Segmentation, Feature Selection and Fault Identification algorithm (MFS2FI). A robustness analysis for the proposed MFS2FI algorithm was performed, and its performance was compared with the performances of a standard FFT classifier, and an

NN-based classifier in chapter 6. The major fault detection techniques and signal features were characterised in respect of the fault identification of high speed centrifugal equipment in chapter 7.

To the best of the author's knowledge, the major contributions and findings of this study are novel, current, and offer the developers of offshore and onshore LNG production trains algorithms that have been verified in the laboratory environment which ultimately could improve the availability and profitability of LNG production plants.

The main findings of this study can be summarised as follows:

1. A reliability model was developed in order to analyse the reliability and operational availability of a typical FPSO LNG liquefaction system. The model was also extended to allow redundancy options to be considered. The results showed that the introduction of a 100% standby system could increase the reliability of the existing typical medium-sized C3MR LNG liquefaction systems (3MTPA) by around 15.7%.

2. The model was used to estimate the costs and benefits of partial and full redundancy of a typical FPSO liquefaction system. The results showed that the implementation of a 100% standby system on typical medium-sized FPSOs (3MTPA) has the potential to increase annual profit by around US$296 million (from US$1,190 million to US$1,485.98 million per year)..

3. An experimental compressor test-rig was built to emulate a number of compressor machine faults, and to generate data for CM studies. The test rig was commissioned, tested, calibrated and used in order to test and validate the algorithms discussed in this thesis.

4. A novel automated, non-AI, systematic, fast and easy-to-implement MFS2FI algorithm was proposed to improve the applicability, performance, development time and computing time of current CM systems. The algorithm was developed and validated on the lab test rig. The MFS2FI properly segmented the frequency spectra and yielded a detection accuracy of 100% at various segment sizes ranging from 1 kHz to 108 kHz. The best detection performance was achieved at the smallest segment size (1 kHz). At this segment size, the MFS2FI yielded the highest fault identification accuracy and the highest confidence level value.

5. The robustness of the MFS2FI was demonstrated using various signal time lengths and data window positions. It was found that changing the data window length and position did not have a significant effect on the fault identification accuracy of the algorithm. A comparative study of the MFS2FI and the standard FFT and NN-based classifiers showed that the proposed algorithm is more accurate than standard FFT classifiers, has a much shorter development time, and is less than sixteen times computationally intensive than NN-based classifiers as it yielded a 100% detection accuracy in only 3.5 seconds (57.2 seconds for NN).

6. An algorithm "characterisation table" was developed to combine information from several fault detection techniques and signal features, namely AE, vibration, air pressure, crest factor, energy factor, RMS, amplitude and spectral features (FFT). The AE spectral features demonstrated their effectiveness over the other techniques and signal features which were addressed, and they yielded a fault identification accuracy of 100%. This approach is considered to have good potential for the development of CM system for typical centrifugal equipment and for the improvement of the fault identification accuracy (compared with a single technique). The collective utilisation of a number of techniques and signal features

could decrease the potential for false alarms resulting from noise disturbance and fault interference.

Overall, based on the findings of this study, the introduction of a 100% standby system to the C3MR liquefaction system on FPSO alongside the implementation of condition monitoring as part of a condition based maintenance system could provide LNG production trains with significant improvements in system reliability and significantly reduced maintenance costs.

However, the lack of information about FPSO maintenance costs hindered the preventive maintenance cost estimations presented in this study. This study also limited itself to similar high-speed industrial centrifugal compressors, and was carried out in a controlled laboratory environment at a specific ambient temperature and operational times, and under specific conditions. An investigation should be carried out to apply the result of this study to similar industrial centrifugal compressors at different operating conditions, and to different types of centrifugal equipment. However, it is anticipated that the results will be applicable to a wide range of high-speed centrifugal equipment.

8.1 FUTURE WORK

Future work should build on the promising lab scale results, and begin applying the redundancy and the CM technique to real FPSO plants. For example, a key step would be to collect real data from compressors on a working FPSO, and evaluate the effectiveness of the proposed CM approaches on such real data over an extended period of time (and as faults develop).

Another future academic study which could emerge from the present one is the application of the research approach in this study to other rotating equipment. A number of outstanding real world implementation issues must be solved to pave the way for the

development of a truly general purpose solution for all types of FPSOs. These problems indicate a variety of research directions that need to be pursued to make such a system feasible. One such direction is the investigation of the reliability of other liquefaction systems utilised on FPSOs. Another possibility is the application of the proposed CM system to different equipment, and to increase the number of fault patterns. It would be also beneficial to investigate more signal analysis techniques such as wavelets, and to add the results of such investigations to the characterisation table developed.

References

[1]   "World Energy Outlook 2014," *International Energy Agency,* 2014.

[2]   "World Energy Outlook 2011," International Energy Agency, 2011.

[3]   "Annual Energy Outlook 2010," U.S. Energy Information Administration, 2010.

[4]   "Annual Energy Outlook 2013," U.S. Energy Information Administration, 2013.

[5]   J. Bukowski, Y. Liu, S. Boccella and K. Leo, "Innovations in natural gas liquefaction technology," in *International Gas Union Research Conference*, Seoul, October 2011.

[6]   The Energy Department's Fossil Energy Organization, USA, (available at http://energy.gov/fe/about-us).

[7]   S. Gowid, R. Dixon and R. Ghani, "Assessment of liquefaction systems' process performance on LNG floating export terminals," *International Conference on Mechanical, Automotive and Aerospace Engineering, 2-4 July, Kuala Lumpur, Malaysia (2013).*

[8]   C. Pil, M. Rausand and J. Vatn, "Reliability assessment of reliquefaction systems on LNG carriers," *Reliability Engineering & Systems Safety,* vol. 93, no. 9, September 2008.

[9]   Q. Li and Y. Ju, "Design and analysis of liquefaction process for offshore associated gas resources," *Journal of Applied Thermal Engineering,* vol. 30, no. 16, p. 417, January 2011.

[10]   A. Rienstra, " Predictive Maintenance for the Masses," Airbone Ultrasonic, SDT North America, 2005.

[11]   PuraDYN, (avilable at http://www.puradyn.com/products/ConditionBasedMaintenance-CBM.html).

[12]   I. Viktorov, Rayleigh and Lamb Waves: Physical Theory and Application, New York, USA: Plenum Press, 1967.

[13]   R. N. 2.79/288, "Flaring & venting in the oil & gas exploration & production industry," Internaitonal Association of Oil and Gas Producers (OGP), London, UK, 2000.

[14]   S. Gowid, R. Dixon and S. Ghani, "FFT-Based robust automated segmentation and features selection algorithm for condition based monitoring systems," *Journal of Applied Acoustics*

(ELSEVIER), Submitted Jan 2014.

[15] G. Yan and J. Yonglin, "LNG-FPSO: Offshore LNG solution," *Frontiers of Energy and Power Engineering China,* vol. 2, no. 3, pp. 249-255, 2008.

[16] T. N. O. I. A. (OLF), "10 Year operability survey of Norwegian FPSOs," 2011.

[17] M. Haid, "Small and Mid-Scale LNG FPSO," in *15th TGE Symposium* , Shanghai China, 2010.

[18] "World LNG report," International Gas Union, 2014.

[19] J. Bukowski et al, "Natural gas liquefaction technology for floating LNG facilities," Air Products and Chemicals Inc., 2013.

[20] M. Schiraldi, Reliability and Maintainability in Operations Management, INTECH (DOI: 10.5772/54161), 2013.

[21] Weibull.com, "The eMagazine for the reliability professionals," 2003. [Online]. Available: http://www.weibull.com/hotwire/issue26/relbasics26.htm. [Accessed 1 4 2016].

[22] OREDA 2009 Handbook, Offshore Reliability Data Handbook, 5th ed., vol. 1.

[23] M. Rausand and A. Hyland, System Reliability Theory: Models, Statistical Methods, and Applications, 2nd ed., New York: Wiley, 2004.

[24] E. Calixto, Gas and Oil reliability engineering: modeling and analysis, 1st Edition, USA: Gulf Professional Publishing, 2012.

[25] R. Peters, Reliable Maintenance Planning, Estimating, and Scheduling (1st edition), USA: Gulf professional publishing, 2014.

[26] T. Aven, "Availability evaluation of oil/gas production and transportation systems," *Relibaility Engineering,* vol. 18, no. 1, pp. 35-44, 1987.

[27] S. Cheng, B. Lin, B. Hsu and M. Shu, "Fault-tree analysis for liquefied natural gas terminal emergency shutdown system," *Expert Systems with Applications – Elsevier,* vol. 36, no. 9, pp. 11918-11924, 2009.

[28] Y. Gu and Y. Ju, "Effect of parameters on performance of LNG-FPSO offloading system in offshore associated gas fields," *Applied Energy,* vol. 8, no. 1, p. 417, January 2011.

[29] A. Sarkar, D. Behera, S. Kumar and M. Singha, "Reliability assessment of Rukhia Gas Turbine Power Plant in Tripura," *International Journal of Current Engineering and Technology,* vol. 2, no. 1, pp. 184-195, 2012.

[30] M. Miranda and O. Meira, "Life cycle assessment of turbomahcinery for offshore applications – updated with field data," *Proceding of the 37th Turbomachinery*

Symposium, Turbomachinery Laboratory, Texas A&M University, Texas, pp. 103-110, 2008.

[31] O. Thorsen, N. Coll and M. Dalva, "A survey of faults on induction motors in offshore oil industry, petrochemical industry, gas terminals, and oil refineries," *Industry Applications, IEEE Transactions,* vol. 31, no. 5, pp. 1186-1196, 1995.

[32] D. Yuhva and Y. Datao, "Estimation of failure probability of oil and gas transmission pipelines by fuzzy fault tree analysis," *Journal of Loss Prevention in the Process Industries,* vol. 18, no. 2, pp. 83-88, 2005.

[33] C. Caroni, "Modeling the reliability of ball bearings," *Journal of Statistics Education,* vol. 10, no. 3, 2002.

[34] L. Langston and G. Opdyke, "Introduction to gas turbines for non-engineers," *Global Gas Turbine News,* vol. 37, no. 2, 1997.

[35] E. Turbo, "Horizontal split compressors," [Online]. Available: http://www.elliott-turbo.com/Files/Admin/CentrifugalCompressors/Elliott-Centrifugal-Compressor-Specifications.pdf. [Accessed Jan 20017].

[36] "subseaworldnews," [Online]. Available: http://subseaworldnews.com/2014/03/25/bw-offshore-inks-contract-extension-for-sendje-berge-fpso/. [Accessed 8 1 2017].

[37] J. Pukite and P. Pukite, Modeling for reliability analysis: Markov modeling for reliability, maintainability, safety, and supportability analyses of complex systems, Wiley-IEEE Press, 1998.

[38] D. Stamatics, Introduction to risk and failures: tools and methodologies, Taylor & Francis Group, 2014.

[39] J. Andrew and C. Ericson, "Fault Tree and Markov Analysis Applied to Various Design Complexities," in *Proceedings of the 18th international system safety conference*, Texas, USA, September 2000.

[40] N. Fuqua, "The applicability of makov analysis method to the reliability, maintainability and safety," *Reliability Analysis Center,* vol. 10, no. 2, 2003.

[41] L. Ridley and J. Andrews, "Opitmal design of systems with standby dependencies," *Quality and Reliability Egineering International,* vol. 15, no. 2, pp. 103-110, 1999.

[42] R. Mobley, An introduction to Predictive Maintenance, Butterworth-Heinemann, 2002.

[43] ABB, ABB Article, [Online]. Available: http://new.abb.com/medium-voltage/service/maintenance/feature-articles/4-types-of-maintenance-strategy-which-one-to-choose. [Accessed Jan 2016].

[44] M. Vasili, T. Hong, N. Ismail and M. vasili, "Maintenance optimization models: a review and analysis," in *Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management*, Kuala Lumpur, 2011.

[45] D. Ghosh and S. Roy, "Maintenance optimization using probabilistic cost-benefit analysis," *Journal of Loss Prevention in the Process Industries,* vol. 22, no. 4, pp. 403-407, 2009.

[46] G. Ravnestad, S. Panesar, D. Kayrbekova and T. Markeset, "Improving periodic preventive maintenance strategies using condition monitoring data," *IFIP Advances in Information and Communication Technology, Springer,* vol. 384, pp. 260-267, 2012.

[47] E. Hurdle, L. Bartlett and J. Andrews, "Fault diagnostics of dynamic system operation using a fault tree based method," *Reliability Engineering and System Safety,* vol. 94, no. 9, pp. 1371-1380, 2009.

[48] M. Lampis and J. Andrews, "Bayesian belief networks for system fault diagnostics," *Quality and Reliability Engineering International,* vol. 25, no. 4, pp. 409-426, 2009.

[49] "Implementation strategies and tools for condition based maintenance at nuclear power plants," *IAEA-TECDOC-1551, International Atomic Energy Agency, May 2007.*

[50] S. Gowid, R. Dixon and S. Ghani, "Profitability, reliability and condition based monitoring of LNG floating platforms: A review," *Journal of Natural Gas Science and Engineering,* vol. 27, no. 3, p. 1495–1511, 2015.

[51] S. Katipamula and M. Brambley, "Methods for fault Detection, diagnostics, and prognostics for building systems: a review, Part I," *International Journal of HVAC&R Research (ASHRE),* vol. 11, no. 1, pp. 3-25, 2005.

[52] C. Byington, M. Roemer, M. Watson, T. Galie and C. Savage, "Prognostic Enhancements to Diagnostic Systems (PEDS) applied to Shipboard Power Generation Systems," *ASME Turbo Expo 2004: Power for Land, Sea, and Air,* vol. 2, pp. 825-833, 2004.

[53] T. Galka and M. Tabaszewski, "An application of statistical symptoms in machine condition diagnostics," *Mechanical Systems and Signal Processing,* vol. 25, no. 1, pp. 253-265, 2011.

[54] J. Yan, Machinery Prognostics and Prognosis Oriented Maintenance Management, Wiley, 2015.

[55] L. e. a. Ginzinger, "Model-based condition based monitoring of an auxiliary bearing following contact events," in *Proceeding of MOVIC*, Tokyo, 2010.

[56] G. Charles, R. Goodall and R. Dixon, "Model-based condition monitoring at the wheel–rail interface," *Vehicale system dynamics,* vol. 46, pp. 451-430, 2008.

[57] C. Ugechi, E. Ogbonnaya, M. Lilliy, S. Ogaji and S. Probert, "Condition-based diagnostic approach for predicting the maintenance requirements of machinery," *Engineering,* vol. 1,

pp. 177-187, 2009.

[58] "Dynamic modeling and analysis of a spur planetary gear involving tooth wedging and bearing clearance nonlinearity," *European Journal of Mechanics - A/Solids,* vol. 29, no. 6, pp. 1022-1033, 2010.

[59] V. Venkatasubramanian, R. Rengaswamy, S. Kavuri and K. Yin, "A review of process fault detection and diagnosis - Part III: process history based methods," *Computers and Chemical Engineering,* vol. 27, pp. 327-346, 2003.

[60] F. Shang, L. Jong and Z. Yabin, "Fault diagnosis system for reciprocating air compressor based on support vector machine," in *Proceedings of the 2009 international workshop on information security and applications*, Qingdao, China, November 2009.

[61] J. Miller, "Statistical signatures used with principal component analysis for fault detection and isolation in a continuous reactor," *Journal of Chemometrics,* vol. 20, pp. 34-42, 2006.

[62] S. Toprak and A. Iftar, "Fault Diagnosis on Hermetic Compressors Based on Sound Measurements," in *Proceedings of the 16th IEEE International Conference on control applications part of IEEE multi-conference on systems and control*, Singapore, Singapore, October 2007.

[63] C. e. a. Scheer, "Quality control using vibration and acoustic signatures," Available at http://www.aaende.org.ar/ingles/sitio/biblioteca/material/T-040.pdf, Accesed Feb. 2014.

[64] C. Boller, F. Chang and Y. Fujino, "Encyclopedia of Structural Health Monitoring," Chichester, John Wiley & Sons, 2009.

[65] Z. Orlowski, "Vibrodiagnostics of Steam Turbines," in *Proceedings of the Institute of Power Engineering*, Warszawa, 1989.

[66] C. Marcelo, J. Fossatti and J. Terra, Fault diagnosis of induction motors based on FFT, DOI: 10.5772/37419, 2012.

[67] El-ThaljiI. and E. Jantunen, "Fault analysis of the wear fault development in rolling bearings," *Engineering Failure Analysis, Elsevier,* vol. 57, p. 470–482, 2015.

[68] A. Minga, Z. Qina, W. Zhangb and F. Chua, "Spectrum auto-correlation analysis and its application to fault diagnosis of rolling element bearings," *Mechanical Systems and Signal Processing, Elsevier,* vol. 41, no. 1-2, pp. 141-154, 2013.

[69] J. Wang and H. Hu, "Vibration-based fault diagnosis of pump using fuzzy technique," *Measurement – Elsevier,* vol. 39, no. 2, pp. 176-185, 2006.

[70] H. Liao and Z. Huang, "The fault diagnosis for centrifugal compressor based on time series analysis with neutral network," in *Proceedings of the 3rd International Conference on*

Advanced Computer Theory and Engineering (ICACTE), Chengdu, China, Augut 2010.

[71]   G. Hafaifa, "Robust fuzzy fault detection and isolation approach applied to surge in centrifugal compressor modeling and control," *Fuzzy Information and Engineering,* vol. 2, no. 1, pp. 49-73, March 2010.

[72]   M. Elha, M. Almrabet, M. Rgeai and I. Ethiwesh, "A combined practical approach to condition monitoring of reciprocating compressors using IAS and dynamic pressure," *World academy of science, Engineering and technology,* vol. 39, no. 1, pp. 186-192, March 2010.

[73]   S. Zanoli , G. Astalfi and L. Barboni, "Applications of fault diagnosis techniques for a multi-shaft centrifugal compressor," in *Proceedings of the 18th Mediterranean Conference on Control & Automation*, Marrakech, Morocco, June 2010.

[74]   Z. H., "Fault Detection, Supervision and Safety of Technical Processes 2006," in *A Proceedings Volume from the 6th IFAC Symposium, SAFEPROCESS 2006*, Beijing, P.R. China, 2006.

[75]   S. Schulthesis, C. Lickteig and R. Parchewsky, "Reciprocating compressor condition monitoring," in *Proceedings of the 36th turbo machinery symposium*, Texas, USA, September 2007.

[76]   A. Rienstra, "Predictive maintenance for the masses," in *Airbone Ultrasonic, SDT North America, 2005*.

[77]   W. l. M. o. t. f. U. u. d. instruments, UE Systems INC, (available at http://www.uesystems.eu/applications/ultrasonic-bearing-mechanical-inspection).

[78]   D. Mba and R. Rao, "Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines; Bearings, Pumps,Gearboxes, Engines and Rotating Structures," *The Shock and Vibration Digest,* vol. 38, no. 1, pp. 3-16, 2006.

[79]   K. Miller and P. Mclntire, Non-destructive testing handbook (3rd ed.), USA: American Society of Non-destructive testing, 1994.

[80]   A. Saxena and A. Saad, "Genetic Algorithms for Artificial Neural Net-based Condition Monitoring System Design for Rotatin Mechanical Systems," *Advances in Soft Computing ,* vol. 34, pp. 135-149, 2006.

[81]   J. Mathews, Acoustic emission (3rd ed.), USA: Gordon and Breach Science Publishers Inc, 1983.

[82]   A. Al-Ghamd and D. Mba, "comparative experimental study on the use of Acoustic Emission and vibration analysis for bearing defect identification and estimation of defect size," *Mechanical Systems and Signal Processing,* vol. 20, no. 7, pp. 1537-1571, 2006.

[83] N. Tandon and B. Nakra, "Defect detection of rolling element bearings by acoustic emission method," *Journal of Acoustic Emission,* vol. 9, no. 1, pp. 25-28, August 1990.

[84] L. Rogers, "The application of vibration analysis and acoustic emission source location to on-line condition monitoring of anti-friction bearings," *Journal of Tribology International,* vol. 12, no. 2, pp. 51-58, April 1979.

[85] J. Schoess, "Development and application of stress-wave acoustic diagnostics for roller bearings," in *Proceedings of SPIE - The International society for optical engineering*, California, USA, June 2000.

[86] G. Neill, E. Brown, J. Steel, P. Standford and R. Reuben, "Detection of recirculation in pumps using acousticemission," in *Proceedings of the 11th International Conference on Condition Monitoring and Diagnostic Engineering Management*, Tasmania, Australia, June 1997.

[87] L. Alfayez and D. Mba, "Detection of incipient cavitation and determination of the best efficiency point for centrifugal pumps using acoustic emission," *Journal of Process Mechanical Engineering,* vol. 219, no. 1, pp. 327-344, 2005.

[88] J. Dane, "Ultrasonic measurement of unsteady gas flow," *Flow measurement instrumentation – Elsevier,* vol. 8, no. 3, pp. 183-190, April 1998.

[89] A. Puttmer, "New applications for ultrasonic sensors in process industries," *Ultrasonics,* vol. 46, no. 1, pp. 1379-1383, June 2006.

[90] A. Goodman, "Utilizing airborne/structure borne ultrasound for condition monitoring predictive maintenance inspections," Compressionjobs.com, March 2010.

[91] K. Gryllias, C. Yiakopoulos and I. Antoniadis, "Automated diagnostic approaches for defective rolling element bearings using minimal training pattern classification methods," *Engineering Asset Lifecycle Management - Proceedings of the 4th World Congress on Engineering Asset Management, WCEAM,* pp. 862-876, 2009.

[92] M. Samhouri, A. Al-Ghandoor, S. Alhaj-Ali and I. Hi, "An intelligent machine condition monitoring system using time-based analysis: neuro-fuzzy versus neural network," *Jordan Journal of Mechanical and Industrial Engineering,* vol. 3, pp. 294-305, 2009.

[93] P. Gupta and S. Wadhwani, "Feature selection by genetic programming, and artificial neural network-based machine condition monitoring," *International Journal of Engineering and Innovative Technology (IJEIT),* vol. 1, pp. 177-181, 2012.

[94] C. Scheer, U. Südmersen, W. Reimche and W. Bach, "Quality control using vibration and acoustic signatures," Rio de Janeiro - RJ – Brasil, 2003.

[95] W. Wang, Complex system maintenance handbook, Springer Series Relibaility Engineering,

2008.

[96]   T. Lima, A. Silva and T. Ludermir, "Selection and fusion of neural networks via differential evolution," in *Advances in Artificial Intelligence - IBERAMIA 2012*, Colombia, Springer, 2012, pp. 153-155.

[97]   Y. Chan and S. Gu, "Modeling of turbine cycles using a Neuro-Fuzzy based approach to predict turbine-generator output for nuclear power plants," *Energies,* vol. 5, no. 1, pp. 101-118, 2012.

[98]   V. Vapnik, "An Overview of Statistical Learning Theory," *IEEE Transactions on Neural Networks, 10(1999)..*

[99]   Q. Meng, Q. Meng and W. Feng, "A new application of support vector machine method: condition monitoring and analysis of reactor coolant pump," *25th International Congresson Condition Monitoring and Diagnostic Engineering, Journal of Physics:ConferenceSeries ,* vol. 364, 2012.

[100] Z. Zhao, L. Wang and H. Liu, "Efficient spectral feature selection with minimum redundancy," in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, Georgia, USA, 2010.

[101] L. Lin, "Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis," *Journal of Sound and Vibration,* vol. 234, p. 135–148, 2000.

[102] J. Qiu, J. Lee, J. Lin and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration,* vol. 289, pp. 1066-1090, 2006.

[103] M. Bozchalooi, "A smoothness index-guided approach to wavelet parameters election in signal de-noising and fault detection," *Journal of Sound and Vibration,* vol. 308, pp. 246-267, 2007.

[104] D. Wang, C. Shen and P. Tse, "A novel adaptive wavelet stripping algorithm for extracting the transients caused by bearing localized faults," *Journal of Sound and Vibration,* vol. 332, pp. 6871-6890, 2013.

[105] Z. Shen, N. He and L. Li, "An intelligent monitoring system with the capability of automated features selection," *Journal of Harbin Institute of Technology ,* vol. 42, no. 9, pp. 1495-1499, 2010.

[106] Tusiani and D. Micahel, "LNG: a non-technical guide," Pennwell Corporation, USA, 2007.

[107] A. Alabdulkarema, A. Mortazavia, Y. Hwanga, R. Radermachera and P. Rogersb, "Optimization of propane pre-cooled mixed refrigerant LNG plant," *Applied Thermal Engineering,* vol. 31, no. 6-7, pp. 1091-1098, 2011.

[108] K. Inc., (available at http://www.users.uswest.net/~kryopak/LNGprocesses.html), accessed on 12 May 2012.

[109] P. B. N. Z. Ltd, "Cost estimates for thermal peaking plant," 2008. [Online]. Available: http://large.stanford.edu/publications/coal/references/docs/thermal-peaking.pdf. [Accessed 1 11 2015].

[110] S. Gowid, R. Dixon and S. Ghani, "Optimization of reliability and maintenance of liquefaction system on FLNG terminals using markov modelling," *International Journal of Quality & Reliability Management (IJQRM),* vol. 31, no. 3, pp. 293 - 310, 2014.

[111] M. Xie, H. Kong and G. T.N, "Exponential approximation for maintained Weibull distributed component," *Journal of Quality in Maintenance Engineering,* 2000.

[112] Isograph software , (available at http://www.isograph-software.com/index.htm).

[113] L. W. News, "Skikda LNG unit shutdown for maintenance," 2014. [Online]. Available: http://www.lngworldnews.com/skikda-lng-unit-shut-down-for-maintenance/. [Accessed Jan 2015].

[114] R. Schoen, T. Habetler, F. Kamran and R. Bartheld, "Motor bearing damage detection using stator current monitoring," *IEEE Transactions,* vol. 31, no. 6, pp. 1274-1279, 1995.

[115] M. Norton and D. Karczub, Fundamentals of Noise and Vibration Analysis for Engineers - 2nd edition, Cambridge University Press, 2003.

[116] B. Granny and K. Starry, "Rolling element bearing analysis," *Materials Evaluation,* vol. 70, no. 1, pp. 78-85, 2011.

[117] R. Mobley, An introduction to Predictive Maintenance, USA: BUTTERWORTH HEINEMANN, 2002.

[118] T. Wang, Intelligent Condition Monitoring and Diagnosis Systems, Ios Press Inc, 2003.

[119] "Acoustic Emission Testing," in *Damage, Testing, Prevention and Detection in Aeronautics*, Vrije Universiteit Brussel, 2006, p. 10.15.

[120] S. Goldman, Vibration spectrum analysis: a practical approach, Industrial Press, Inc.; 2 edition, 1999.

[121] Y. Liua, L. Guob, Q. Wangc, G. And, M. Guob and H. Lianb, "Application to induction motor faults diagnosis of the amplitude recovery method combined with FFT," *Mechanical Systems and Signal Processing, Elsevier,* vol. 28, no. 8, p. 2961–2971, 2010.

[122] T. Wanga, M. Liangb, J. Lia, W. Chenga and C. Lib, "Bearing fault diagnosis under unknown variable speed via gear noise cancellation and rotational order sideband identification,"

Mechanical Systems and Signal Processing, Elsevier, Vols. 62-63, pp. 30-53, 2016.

[123] P. McFadden and J. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *Journal of Sound and Vibration,* vol. 96, pp. 69-82, 1984.

[124] B. P. Bearings, "Bearing failure: causes and cures," 2017. [Online]. Available: http://www.schaeffler.com/remotemedien/media/_shared_media/08_media_library/01_ publications/barden/brochure_2/downloads_24/barden_bearing_failures_us_en.pdf. [Accessed 2 2017].

[125] S. Gowid, R. Dixon and S. Ghani, "Characterisation of major fault detection features and techniques for the condition monitoring of high-speed centrifugal equipment," *International Journal of Acoustics and Vibration,* vol. 21, no. 2, pp. 184-191, 2016.

[126] V. Oppenheim and R. Schafer, Discrete-Time Signal Processing, Englewood Cliffs, Prentice-Hall, 1989.

[127] J. KEPRT and P. BENES, "A comparison of AE sensor calibration methods," *Journal of Acoustic Emission,* vol. 26, no. 1, pp. 60-71, 2008.

[128] S. Vasilic, M. Kezunovic and D. Sobajic, "Optimizing performance of a transmission line relaying algorithm implemented using an adaptive, self-organized neural network," in *Paper 3, 14th PSCC Conference*, Servilla, 2002.

[129] S. Ross, Introductory Statistics, 3rd edition, Academic Press, 2005.

[130] C. Stergiou and D. Siganos , "Neural networks," *Surprice 96 Journal, Imperial College London,* vol. 4, no. 1, 2011.

[131] P. McCorduck, Machines Who Think, New York: W. H. Freeman & Co., 1979.

[132] Y. Cho and S. Kang, "Sensory evaluation using artificial intelligence," in *Emerging technologies for food quality and food safety evaluation*, CRC Press, 2011.

[133] "Multilayer Neural Network Architecture," Matlab help, MathWorks Company, 2015.

[134] J. Jesan, "The neural approach to pattern recognition," USA, ACM, 2004, pp. 2-2.

[135] S. Belciuga and F. Gorunescub, "Error-correction learning for artificial neural networks using the Bayesian paradigm. Application to automated medical diagnosis," *Journal of Biomedical Informatics,* vol. 52, pp. 329-337, 2014.

[136] M. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks,* vol. 6, no. 4, pp. 525-533, 1993.

[137] D. Meh, M. denislic, M. Lehtokangas and J. Saarinen, "Weight initialization with reference

patterns," *Neurocomputing,* vol. 20, no. 131, pp. 265-278, 1998.

[138] M. Fernández-Redondo and C. Hernández-Espinosa, "Weight initialization methods for multilayer feedforward," *ESANN'2001 proceedings - European Symposium on Artificial Neural Networks Bruges,* Vols. ISBN 2-930307-01-3, pp. 119-124, 2001.

[139] Z. Wang and A. Bovik, "Data mining concept and techniques, Mean squared error: Love it or leave it?," *IEEE Signal Processing Magazine,* vol. 98, pp. 98-112, 2009.

[140] M. Akhmet and E. Yılmaz, Neural networks with discontinuous/impact activations, Springer, 2013.

[141] R. Rojas, "The Backpropagation Algorithm," in *Newral Networks*, Berlin, Springer, 1996, p. Chapter 7.

[142] D. Ruan, Intelligent Hybrid Systems: Fuzzy Logic, Neural Networks, and Genetic Algorithms, Springer, ISBN: 978-0-7923-9999-5, 1997.

[143] P. Subbaraj and B. Kannapiran, "Artificial neural network approach for fault detection in pneumatic valve in cooler water spray system," *International Journal for Computer Applications,* vol. 9, no. 7, pp. 43-52, 2010.

[144] G. Dahl, A. McAvinney and T. Newhall, " Parallelizing neural networks training for cluster systems," in *PDCN '08 Proceedings of the IASTED International Conference on Parallel and Distributed Computing and Networks*, CA, USA, 2008.

[145] J. Chitode, Continuous TIme Signal and Systems, India: Technical Publications Pune, 2003.

[146] C. Sujatha, Vibration and Acoustics, New Delhi: tata McGraw Hill Education Private Limited, 2010.

APPENDIX (A)

MATLAB code for maintenance time interval optimisation

```matlab
Lo=9.15e-4; %per hour

Nj=2; %Number of components in subsystem

%Lo=2e-4*3*30; % for 3 month of operation

To=672 %Assuming that the time between preventive maintenances is 4 weeks (should

be less than the MTTF)


r=0.7;

M=1/Lo;

A=2.5;

Year=5*12*24;

T=200:1:1000;  %Tmax should be less than MMTF - MTTF =1/Lo

%T=Year; %maintenance interval= 1 year


Cpm=50000;

Ccm=100000; %per failure

Cs=(3.3e6/(12*30*24))*500; %cost of shut down/hr, knowing that, the price of LNG is

500 USD/Ton LNG = Approx. 190 000 USD/hr

Pr=3.3e6; %Production rate = 3.3 MTPA (Million Tons Per Annum)

Total_income=Cs*Pr; % Total income per year


num1=1-((0.1*A*T.^2)/M^2)+ (((0.09*A-0.2)*T)/M);

denum1=1-((0.1*A*To^2)/M^2)+ (((0.09*A-0.2)*To)/M);
```

```matlab
Le=r*Lo+(1-r)*Lo*((T./To).^(A-1)).*(num1/denum1);% effective failure rate per hour or
per Lo


Cst=Cs.*Le;

max1=max(Cst);

PM=Cpm./T;

CM=Ccm*Le;


plot(T,Cst,'-.b');

hold on

plot(T,PM,'--r');

hold on

plot(T,CM,'-.g');


Ctotal=Cst+CM*Nj+PM*Nj;


min1=min(Ctotal) % @T=1760 hr


hold on

plot(T,Ctotal,'-black');


xlabel('Time interval between periodic maintenance [h]')

ylabel('USD / h')

legend('Shutdown','PM cost','CM cost','Total cost')
```
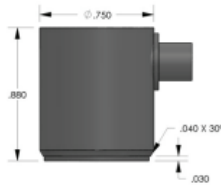
APPENDIX (B)

Data sheets

R6α Sensor

General Purpose, 60 kHz Resonant Frequency Sensor

DESCRIPTION AND FEATURES

The R6α is a narrow band resonant sensor with a high sensitivity. The sensor cavity is machined from a solid stainless steel rod, making the sensor extremely rugged and reliable. The ceramic face along with a 30 degree chamfer to cavity electrically isolates the sensor cavity from the structure under test assuring a low noise operation.

The compact size of the sensor makes it readily suitable for deploying in tight spaces for monitoring. The Alpha series family of sensors features an SMA connector versus the Microdot connectors found on MISTRAS' RXX series of sensors. The alpha series includes R3α, R6α, R15α, R30α, R50α, R80α and WSα sensors.

APPLICATIONS

This sensor can be used on metal and FRP structures such as pipelines or storage tanks in petroleum, refineries, chemical plants, and offshore platforms, due to its high sensitivity and low resonance frequency properties.
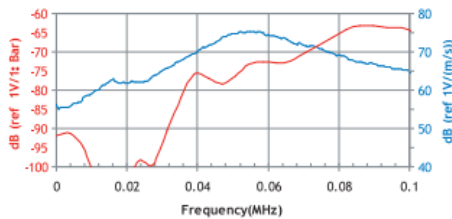
OPERATING SPECIFICATIONS

Dynamic

Peak Sensitivity, Ref V/(m/s)	75 dB
Peak Sensitivity, Ref V/µbar	-64 dB
Operating Frequency Range	35-100 kHz
Resonant Frequency, Ref V/(m/s)	55 kHz
Resonant Frequency, Ref V/µbar	90 kHz
Directionality	+/-1.5 dB

Environmental

Temperature Range	-65 to 175ºC
Shock Limit	500 g

Completely enclosed crystal for RFI/EMI immunity

Physical

Dimensions	0.75"OD X 0.88"H
	19 mm OD X 22.4 mm H
Weight	38 grams
Case Material	Stainless steel
Face Material	Ceramic
Connector	SMA
Connector Locations	Side
Seal	Epoxy

ORDERING INFORMATION AND ACCESSORIES

R6α	R6α or R6a
Magnetic Hold-Down	MHR15A
Sensor to Preamp Cable (1 or 2 meters)	1232-X-SMA
Amplifier subsystems	AE2A, AE5A or standard AE systems
Preamplifier	0/2/4, 2/4/6
Preamp to System Cable (specify length in 'm')	1234 - X

Sensors include

NIST Calibration Certificate & Warranty

Industrial air compressor from Paxton (model AT1200)

CHOOSE A SYSTEM FOR YOUR APPLICATION
This chart will help you find the AT-Series Blower and motor horsepower needed for your air pressure and flow specifications

If you need additional data, have questions, or would like assistance in sizing a blower or exhauster for your application or equipment, please don't hesitate to contact us. Our technical service staff and sales representatives are here to help, and we look forward to serving you.

APPENDIX (C)

MATLAB code for the proposed MFS2FI algorithm

```matlab
clear

clc


% loading the training datasets ( 4 sets for each machine condition)


data1=load('Healthy\1\1Impeller_ data matrix.mat');

data2=load('Healthy\3\1Impeller_ data matrix.mat');

data3=load('Healthy\8\1Impeller_ data matrix.mat');

data4=load('Healthy\10\1Impeller_ data matrix.mat');


data5=load('Leak\1\1Impeller_ data matrix.mat');

data6=load('Leak\3\1Impeller_ data matrix.mat');

data7=load('Leak\8\1Impeller_ data matrix.mat');

data8=load('Leak\10\1Impeller_ data matrix.mat');


data9=load('Impeller\1\1Impeller_ data matrix.mat');

data10=load('Impeller\3\1Impeller_ data matrix.mat');

data11=load('Impeller\8\1Impeller_ data matrix.mat');

data12=load('Impeller\10\1Impeller_ data matrix.mat');


data13=load('Belt\1\1Impeller_ data matrix.mat');

data14=load('Belt\3\1Impeller_ data matrix.mat');

data15=load('Belt\8\1Impeller_ data matrix.mat');

data16=load('Belt\10\1Impeller_ data matrix.mat');
```

```
data17=load('Both\1\1Impeller_ data matrix.mat');

data18=load('Both\3\1Impeller_ data matrix.mat');

data19=load('Both\8\1Impeller_ data matrix.mat');

data20=load('Both\10\1Impeller_ data matrix.mat');


X1=abs(fft(data1.data));

X2=abs(fft(data2.data));

X3=abs(fft(data3.data));

X4=abs(fft(data4.data));


X5=abs(fft(data5.data));

X6=abs(fft(data6.data));

X7=abs(fft(data7.data));

X8=abs(fft(data8.data));


X9=abs(fft(data9.data));

X10=abs(fft(data10.data));

X11=abs(fft(data11.data));

X12=abs(fft(data12.data));


X13=abs(fft(data13.data));

X14=abs(fft(data14.data));

X15=abs(fft(data15.data));

X16=abs(fft(data16.data));
```

```matlab
X17=abs(fft(data17.data));

X18=abs(fft(data18.data));

X19=abs(fft(data19.data));

X20=abs(fft(data20.data));


Solution=[1; 1; 1; 1; 1;1];

SofSolution=size(Solution);


% loading the second training dataset (a set will be only selected and used for the P
Matrix calculations)


data1t=load('Healthy\2\1Impeller_ data matrix.mat');

data2t=load('Healthy\5\1Impeller_ data matrix.mat');

data3t=load('Healthy\9\1Impeller_ data matrix.mat');


data4t=load('Leak\2\1Impeller_ data matrix.mat');

data5t=load('Leak\5\1Impeller_ data matrix.mat');

data6t=load('Leak\9\1Impeller_ data matrix.mat');


data7t=load('Impeller\2\1Impeller_ data matrix.mat');

data8t=load('Impeller\5\1Impeller_ data matrix.mat');

data9t=load('Impeller\9\1Impeller_ data matrix.mat');


data10t=load('Belt\2\1Impeller_ data matrix.mat');

data11t=load('Belt\5\1Impeller_ data matrix.mat');
```

```matlab
data12t=load('Belt\9\1Impeller_ data matrix.mat');


data13t=load('Both\2\1Impeller_ data matrix.mat');

data14t=load('Both\5\1Impeller_ data matrix.mat');

data15t=load('Both\9\1Impeller_ data matrix.mat');


X1t=abs(fft(data1t.data));

X2t=abs(fft(data2t.data));

X3t=abs(fft(data3t.data));


X4t=abs(fft(data4t.data));

X5t=abs(fft(data5t.data));

X6t=abs(fft(data6t.data));


X7t=abs(fft(data7t.data));

X8t=abs(fft(data8t.data));

X9t=abs(fft(data9t.data));


X10t=abs(fft(data10t.data));

X11t=abs(fft(data11t.data));

X12t=abs(fft(data12t.data));


X13t=abs(fft(data13t.data));

X14t=abs(fft(data14t.data));

X15t=abs(fft(data15t.data));
```

```matlab
%S=80000; % Frequency division

kk=1;


for SS=121000:-1000:1000


S=SS;


K=1;

NOC=5;  % Total number of machine conditions (Healthy, air leak, bearing with lack of

lubrication,...,etc)


for KK=2000:S:121000


  KKn=KK+S-1;


  if KKn>121000

    KKn=121000

  end


  X0_Freq(K)=KK;


  X1_max(K)=max(X1(KK:KKn));

  X2_max(K)=max(X2(KK:KKn));
```

```
X3_max(K)=max(X3(KK:KKn));

X4_max(K)=max(X4(KK:KKn));


X5_max(K)=max(X5(KK:KKn));

X6_max(K)=max(X6(KK:KKn));

X7_max(K)=max(X7(KK:KKn));

X8_max(K)=max(X8(KK:KKn));


X9_max(K)=max(X9(KK:KKn));

X10_max(K)=max(X10(KK:KKn));

X11_max(K)=max(X11(KK:KKn));

X12_max(K)=max(X12(KK:KKn));


X13_max(K)=max(X13(KK:KKn));

X14_max(K)=max(X14(KK:KKn));

X15_max(K)=max(X15(KK:KKn));

X16_max(K)=max(X16(KK:KKn));


X17_max(K)=max(X17(KK:KKn));

X18_max(K)=max(X18(KK:KKn));

X19_max(K)=max(X19(KK:KKn));

X20_max(K)=max(X20(KK:KKn));


X1t_max(K)=max(X1t(KK:KKn));

X2t_max(K)=max(X2t(KK:KKn));
```

```matlab
        X3t_max(K)=max(X3t(KK:KKn));


        X4t_max(K)=max(X4t(KK:KKn));

        X5t_max(K)=max(X5t(KK:KKn));

        X6t_max(K)=max(X6t(KK:KKn));

        X7t_max(K)=max(X7t(KK:KKn));


        X8t_max(K)=max(X8t(KK:KKn));

        X9t_max(K)=max(X9t(KK:KKn));

        X10t_max(K)=max(X10t(KK:KKn));

        X11t_max(K)=max(X11t(KK:KKn));


        X12t_max(K)=max(X12t(KK:KKn));

        X13t_max(K)=max(X13t(KK:KKn));

        X14t_max(K)=max(X14t(KK:KKn));

        X15t_max(K)=max(X15t(KK:KKn));


        XXX(K)=0;

        K=K+1;


    end


% Calculating the max and min values for max peaks (Benchmark Values)
```

```matlab
for KK=1:1:K-1


    Xmin(1,KK)=min([X1_max(KK) X2_max(KK) X3_max(KK) X4_max(KK) ]);

    Xmax(1,KK)=max([X1_max(KK) X2_max(KK) X3_max(KK) X4_max(KK) ]);


    Xmin(2,KK)=min([X5_max(KK) X6_max(KK) X7_max(KK) X8_max(KK) ]);

    Xmax(2,KK)=max([X5_max(KK) X6_max(KK) X7_max(KK) X8_max(KK) ]);


    Xmin(3,KK)=min([X9_max(KK) X10_max(KK) X11_max(KK) X12_max(KK) ]);

    Xmax(3,KK)=max([X9_max(KK) X10_max(KK) X11_max(KK) X12_max(KK) ]);


    Xmin(4,KK)=min([X13_max(KK) X14_max(KK) X15_max(KK) X16_max(KK) ]);

    Xmax(4,KK)=max([X13_max(KK) X14_max(KK) X15_max(KK) X16_max(KK) ]);


    Xmin(5,KK)=min([X17_max(KK) X18_max(KK) X19_max(KK) X20_max(KK) ]);

    Xmax(5,KK)=max([X17_max(KK) X18_max(KK) X19_max(KK) X20_max(KK) ]);


    Xtmin(1,KK)=min([X1t_max(KK) X2t_max(KK) X3t_max(KK) ]);

    Xtmax(1,KK)=max([X1t_max(KK) X2t_max(KK) X3t_max(KK) ]);


    Xtmin(2,KK)=min([X4t_max(KK) X5t_max(KK) X6t_max(KK) ]);

    Xtmax(2,KK)=max([X4t_max(KK) X5t_max(KK) X6t_max(KK) ]);


    Xtmin(3,KK)=min([X7t_max(KK) X8t_max(KK) X9t_max(KK) ]);

    Xtmax(3,KK)=max([X7t_max(KK) X8t_max(KK) X9t_max(KK) ]);
```

244

```matlab
Xtmin(4,KK)=min([X10t_max(KK) X11t_max(KK) X12t_max(KK) ]);

Xtmax(4,KK)=max([X10t_max(KK) X11t_max(KK) X12t_max(KK) ]);


Xtmin(5,KK)=min([X13t_max(KK) X14t_max(KK) X15t_max(KK) ]);

Xtmax(5,KK)=max([X13t_max(KK) X14t_max(KK) X15t_max(KK) ]);



end



%++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

% Testing

%++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++



% First file (healthy - Test)



% first row comparison (j = k-1)



X_test= X2t_max;  % Enter the sample number

 Selection= zeros(10, K-1); % comparing the measured signal with the benchmark values



AAA=zeros(5,5);




for KK= 1:1:K-1
```

```matlab
if X_test(1,KK)>= Xmin(1,KK) &&  X_test(1,KK) <= Xmax(1,KK)

  Selection(1,KK) = 11;


end


 if X_test(1,KK)>= Xmin(2,KK) &&  X_test(1,KK) <= Xmax(2,KK)

  Selection(2,KK) = 22;


 end


if X_test(1,KK)>= Xmin(3,KK) &&  X_test(1,KK) <= Xmax(3,KK)

  Selection(3,KK) = 33;


end


if X_test(1,KK)>= Xmin(4,KK) &&  X_test(1,KK) <= Xmax(4,KK)

  Selection(4,KK) = 44;


end


if X_test(1,KK)>= Xmin(5,KK) &&  X_test(1,KK) <= Xmax(5,KK)

  Selection(5,KK) = 55;


end


if X_test(1,KK)>= Xmin(5,KK) &&  X_test(1,KK) <= Xmax(5,KK)
```

```
end


for Kk=1:1:5,


  for j=1:1:K-1

    P(Kk, j)=abs(((Xmax(Kk,j)-X_test(1,j))/Xmax(Kk,j))*100);

  end


  for j=1:1:K-1

    P(Kk+5, j)=abs(((Xmin(Kk,j)-X_test(1,j))/Xmin(Kk,j))*100);

  end

end


for j=1:1:K-1,

  Pn=P(:,j);

  Pselected(1,j) = min(Pn);

  Index(j)= find(Pn==Pselected(1,j));

end


Index

P1=P

% Creating the Second half of the Selection Matrix
```

```matlab
for j=1:1:K-1

    if (Index(j)==1 || Index(j)==6)

        Selection(6,j)=1;

    end


    if (Index(j)==2|| Index(j)==7)

        Selection(7,j)=2;

    end


    if (Index(j)==3 || Index(j)==8)

        Selection(8,j)=3;

    end


    if (Index(j)==4 || Index(j)==9)

        Selection(9,j)=4;

     end


    if (Index(j)==5 || Index(j)==10)

        Selection(10,j)=5;
```

```matlab
    end

   end


Selection

K

S


% The decision making is based on the Selection Matrix


A1=find(Selection == 1);

A11=find(Selection == 11);

A2=find(Selection == 2);

A22=find(Selection == 22);

A3=find(Selection == 3);

A33=find(Selection == 33);

A4=find(Selection == 4);

A44=find(Selection == 44);

A5=find(Selection == 5);

A55=find(Selection == 55);


A1s=size(A1);

A11s=size(A11);

A1st=A1s+A11s;
```

```matlab
A2s=size(A2);

A22s=size(A22);

A2st=A2s+A22s;


A3s=size(A3);

A33s=size(A33);

A3st=A3s+A33s;


A4s=size(A4);

A44s=size(A44);

A4st=A4s+A44s;


A5s=size(A5);

A55s=size(A55);

A5st=A5s+A55s;


A(:,1)=[A1st(1,1);A2st(1,1);A3st(1,1);A4st(1,1);A5st(1,1)];

AS=A(:,1)

AA=find(AS==max(AS));

SS=size(AA);


if SS(1,1)>1 % There will be no solution if the algorithm detected more than one fault

    clear AA

    AA=[0];

else
```

```matlab
T= strcat(' The fault number is: ', num2str(AA) )

AAA(AA,1)=1;

end




% Second file (Leak - Test)



% first row comparison (j = k-1)



X_test= X5t_max;  % Testing



Selection= zeros(10, K-1); % comparing the testing datasets values with the benchmark
values



for KK= 1:1:K-1
   if X_test(1,KK)>= Xmin(1,KK) &&  X_test(1,KK) <= Xmax(1,KK)

     Selection(1,KK) = 11;


   end



   if X_test(1,KK)>= Xmin(2,KK) &&  X_test(1,KK) <= Xmax(2,KK)

     Selection(2,KK) = 22;


   end
```

```matlab
    if X_test(1,KK)>= Xmin(3,KK) &&  X_test(1,KK) <= Xmax(3,KK)

      Selection(3,KK) = 33;


    end



    if X_test(1,KK)>= Xmin(4,KK) &&  X_test(1,KK) <= Xmax(4,KK)

      Selection(4,KK) = 44;


    end



    if X_test(1,KK)>= Xmin(5,KK) &&  X_test(1,KK) <= Xmax(5,KK)

      Selection(5,KK) = 55;


    end


end


  for Kk=1:1:5,


    for j=1:1:K-1

    P(Kk, j)=abs(((Xmax(Kk,j)-X_test(1,j))/Xmax(Kk,j))*100);

    end



    for j=1:1:K-1
```

```matlab
        P(Kk+5, j)=abs(((Xmin(Kk,j)-X_test(1,j))/Xmin(Kk,j))*100);

    end

end


P2=P


for j=1:1:K-1,

Pn=P(:,j);

Pselected(1,j) = min(Pn);

Index(j)= find(Pn==Pselected(1,j));

end


Index


% Creating the Second half of  the Selection Matrix


for j=1:1:K-1


if (Index(j)==1 || Index(j)==6)

    Selection(6,j)=1;


end


if (Index(j)==2|| Index(j)==7)

    Selection(7,j)=2;
```

```
    end


if (Index(j)==3 || Index(j)==8)

    Selection(8,j)=3;


end


if (Index(j)==4 || Index(j)==9)

    Selection(9,j)=4;


end


if (Index(j)==5 || Index(j)==10)

    Selection(10,j)=5;


end



end


Selection

K

S
```

```matlab
A1=find(Selection == 1);

A11=find(Selection == 11);

A2=find(Selection == 2);

A22=find(Selection == 22);

A3=find(Selection == 3);

A33=find(Selection == 33);

A4=find(Selection == 4);

A44=find(Selection == 44);

A5=find(Selection == 5);

A55=find(Selection == 55);


A1s=size(A1);

A11s=size(A11);

A1st=A1s+A11s;


A2s=size(A2);

A22s=size(A22);

A2st=A2s+A22s;


A3s=size(A3);

A33s=size(A33);

A3st=A3s+A33s;
```

```matlab
A4s=size(A4);

A44s=size(A44);

A4st=A4s+A44s;


A5s=size(A5);

A55s=size(A55);

A5st=A5s+A55s;



A(:,2)=[A1st(1,1);A2st(1,1);A3st(1,1);A4st(1,1);A5st(1,1)];

AS=A(:,2)

AA=find(AS==max(AS));

SS=size(AA);


if SS(1,1)>1 % There is no solution if the algorithm detected more than one fault

   clear AA

   AA=[0];
else
T= strcat(' The fault number is: ', num2str(AA) )

AAA(AA,2)=1;
 end


% Third file (Impeller - Test)


% first row comparison (j = k-1)
```

```matlab
X_test= X8t_max;  % Testing


Selection= zeros (10, K-1); % comparing the measured signal with benchmark values


for KK= 1:1:K-1
  if X_test(1,KK)>= Xmin(1,KK) &&  X_test(1,KK) <= Xmax(1,KK)

    Selection(1,KK) = 11;


  end


   if X_test(1,KK)>= Xmin(2,KK) &&  X_test(1,KK) <= Xmax(2,KK)

    Selection(2,KK) = 22;


   end


  if X_test(1,KK)>= Xmin(3,KK) &&  X_test(1,KK) <= Xmax(3,KK)

    Selection(3,KK) = 33;


  end


  if X_test(1,KK)>= Xmin(4,KK) &&  X_test(1,KK) <= Xmax(4,KK)

    Selection(4,KK) = 44;


  end
```

257

```matlab
    if X_test(1,KK)>= Xmin(5,KK) &&  X_test(1,KK) <= Xmax(5,KK)

      Selection(5,KK) = 55;


    end


end


  for Kk=1:1:5,


    for j=1:1:K-1

    P(Kk, j)=abs(((Xmax(Kk,j)-X_test(1,j))/Xmax(Kk,j))*100);

    end


    for j=1:1:K-1

    P(Kk+5, j)=abs(((Xmin(Kk,j)-X_test(1,j))/Xmin(Kk,j))*100);

    end

  end



  for j=1:1:K-1,

  Pn=P(:,j);

  Pselected(1,j) = min(Pn);

  Index(j)= find(Pn==Pselected(1,j));

  end
```

Index

P3=P

```matlab
% creating the Second half of the Selection matrix

for j=1:1:K-1

    if (Index(j)==1 || Index(j)==6)

        Selection(6,j)=1;

    end

    if (Index(j)==2|| Index(j)==7)

        Selection(7,j)=2;

    end

    if (Index(j)==3 || Index(j)==8)

        Selection(8,j)=3;

    end

    if (Index(j)==4 || Index(j)==9)

        Selection(9,j)=4;
```

```matlab
            end


    if (Index(j)==5 || Index(j)==10)

        Selection(10,j)=5;


        end


    end


Selection

K

S


% The decision making is based on the Selection matrix


A1=find(Selection == 1);

A11=find(Selection == 11);

A2=find(Selection == 2);

A22=find(Selection == 22);

A3=find(Selection == 3);

A33=find(Selection == 33);

A4=find(Selection == 4);

A44=find(Selection == 44);

A5=find(Selection == 5);
```

```
A55=find(Selection == 55);


A1s=size(A1);

A11s=size(A11);

A1st=A1s+A11s;


A2s=size(A2);

A22s=size(A22);

A2st=A2s+A22s;


A3s=size(A3);

A33s=size(A33);

A3st=A3s+A33s;


A4s=size(A4);

A44s=size(A44);

A4st=A4s+A44s;


A5s=size(A5);

A55s=size(A55);

A5st=A5s+A55s;


A(:,3)=[A1st(1,1);A2st(1,1);A3st(1,1);A4st(1,1);A5st(1,1)];

AS=A(:,3)

AA=find(AS==max(AS));
```

```matlab
SS=size(AA);


if SS(1,1)>1 % There is no solution if the algorithm detected more than one fault

    clear AA

    AA=[0];

else

T= strcat(' The fault number is: ', num2str(AA) )

AAA(AA,3)=1;

end



% Fourth file (healthy - Test)

% first row comparison (j = k-1)

X_test= X11t_max; % Testing


Selection= zeros(10, K-1); % comparing the measured signal with the benchmark values


for KK= 1:1:K-1

    if X_test(1,KK)>= Xmin(1,KK) &&  X_test(1,KK) <= Xmax(1,KK)

        Selection(1,KK) = 11;


    end


    if X_test(1,KK)>= Xmin(2,KK) &&  X_test(1,KK) <= Xmax(2,KK)

        Selection(2,KK) = 22;
```

```matlab
    end


    if X_test(1,KK)>= Xmin(3,KK) &&  X_test(1,KK) <= Xmax(3,KK)

        Selection(3,KK) = 33;


    end


    if X_test(1,KK)>= Xmin(4,KK) &&  X_test(1,KK) <= Xmax(4,KK)

        Selection(4,KK) = 44;


    end


    if X_test(1,KK)>= Xmin(5,KK) &&  X_test(1,KK) <= Xmax(5,KK)

        Selection(5,KK) = 55;


    end


end


    for Kk=1:1:5,


        for j=1:1:K-1

        P(Kk, j)=abs(((Xmax(Kk,j)-X_test(1,j))/Xmax(Kk,j))*100);

        end
```

```matlab
    for j=1:1:K-1

    P(Kk+5, j)=abs(((Xmin(Kk,j)-X_test(1,j))/Xmin(Kk,j))*100);

    end

end




    for j=1:1:K-1,

    Pn=P(:,j);

    Pselected(1,j) = min(Pn);

    Index(j)= find(Pn==Pselected(1,j));

    end

    Index

    P4=P


% Calculating the Second half of the Selection matrix

for j=1:1:K-1


    if (Index(j)==1 || Index(j)==6)

        Selection(6,j)=1;


    end
```

```matlab
    if (Index(j)==2|| Index(j)==7)

        Selection(7,j)=2;


    end


    if (Index(j)==3 || Index(j)==8)

        Selection(8,j)=3;


    end


    if (Index(j)==4 || Index(j)==9)

        Selection(9,j)=4;


    end


    if (Index(j)==5 || Index(j)==10)

        Selection(10,j)=5;


    end



end
```

Selection

K

S

```matlab
% Decision making is based on the Selection matrix


A1=find(Selection == 1);

A11=find(Selection == 11);

A2=find(Selection == 2);

A22=find(Selection == 22);

A3=find(Selection == 3);

A33=find(Selection == 33);

A4=find(Selection == 4);

A44=find(Selection == 44);

A5=find(Selection == 5);

A55=find(Selection == 55);


A1s=size(A1);

A11s=size(A11);

A1st=A1s+A11s;

A2s=size(A2);

A22s=size(A22);

A2st=A2s+A22s;


A3s=size(A3);
```

```matlab
A33s=size(A33);

A3st=A3s+A33s;


A4s=size(A4);

A44s=size(A44);

A4st=A4s+A44s;


A5s=size(A5);

A55s=size(A55);

A5st=A5s+A55s;


A(:,4)=[A1st(1,1);A2st(1,1);A3st(1,1);A4st(1,1);A5st(1,1)];

AS=A(:,4)

AA=find(AS==max(AS));

SS=size(AA);


 if SS(1,1)>1 % There is no solution if the algorithm detected more than one fault

    clear AA

    AA=[0];

 else

T= strcat(' The fault number is: ', num2str(AA) )

AAA(AA,4)=1;

 end
```

```matlab
% Fifth file (Belt - Test)


% first row comparison (j = k-1)


X_test= X14t_max;  % Testing


Selection= zeros(10, K-1); % comparing measured signala with the benchmark values


for KK= 1:1:K-1
  if X_test(1,KK)>= Xmin(1,KK) &&  X_test(1,KK) <= Xmax(1,KK)

    Selection(1,KK) = 11;


  end


   if X_test(1,KK)>= Xmin(2,KK) &&  X_test(1,KK) <= Xmax(2,KK)

    Selection(2,KK) = 22;


   end


  if X_test(1,KK)>= Xmin(3,KK) &&  X_test(1,KK) <= Xmax(3,KK)

    Selection(3,KK) = 33;


  end


  if X_test(1,KK)>= Xmin(4,KK) &&  X_test(1,KK) <= Xmax(4,KK)
```

```matlab
      Selection(4,KK) = 44;

  end


  if X_test(1,KK)>= Xmin(5,KK) &&  X_test(1,KK) <= Xmax(5,KK)

    Selection(5,KK) = 55;

  end


end



  for Kk=1:1:5,


    for j=1:1:K-1
    P(Kk, j)=abs(((Xmax(Kk,j)-X_test(1,j))/Xmax(Kk,j))*100);
    end


    for j=1:1:K-1
    P(Kk+5, j)=abs(((Xmin(Kk,j)-X_test(1,j))/Xmin(Kk,j))*100);
    end
  end
```

```matlab
for j=1:1:K-1,

Pn=P(:,j);

Pselected(1,j) = min(Pn);

Index(j)= find(Pn==Pselected(1,j));

end



Index



P5=P



% Calculatin the Second half of Selection matrix



for j=1:1:K-1


if (Index(j)==1 || Index(j)==6)

    Selection(6,j)=1;


end


if (Index(j)==2|| Index(j)==7)

    Selection(7,j)=2;
```

```
        end


    if (Index(j)==3 || Index(j)==8)

        Selection(8,j)=3;


    end


    if (Index(j)==4 || Index(j)==9)

        Selection(9,j)=4;


    end


    if (Index(j)==5 || Index(j)==10)

        Selection(10,j)=5;


    end


end
```

Selection

K

S

```
A1=find(Selection == 1);

A11=find(Selection == 11);

A2=find(Selection == 2);

A22=find(Selection == 22);

A3=find(Selection == 3);

A33=find(Selection == 33);

A4=find(Selection == 4);

A44=find(Selection == 44);

A5=find(Selection == 5);

A55=find(Selection == 55);


A1s=size(A1);

A11s=size(A11);

A1st=A1s+A11s;


A2s=size(A2);

A22s=size(A22);

A2st=A2s+A22s;


A3s=size(A3);

A33s=size(A33);
```

```matlab
A3st=A3s+A33s;


A4s=size(A4);

A44s=size(A44);

A4st=A4s+A44s;


A5s=size(A5);

A55s=size(A55);

A5st=A5s+A55s;


A(:,5)=[A1st(1,1);A2st(1,1);A3st(1,1);A4st(1,1);A5st(1,1)];

AS=A(:,5)

AA=find(AS==max(AS));

SS=size(AA);


 if SS(1,1)>1 % if we have more than one maximum - equal potential of faults

    clear AA

    AA=[0];
 else
T= strcat(' The fault number is: ', num2str(AA) )

AAA(AA,5)=1;
 end
A

AAA
```

```matlab
%+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
% Determining the fault numbers that were detected based on the output
% of matrix AAA

% In each raw, one element only should be equal to one equals to 1 and others equal
should be equal to zero
% Per example, for fault no. 3, element AAA(3,3) should be equal to
% 1 while other should be equal to zero to avoid fault interference.

%+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
++++

F= zeros(5,1);

if AAA(1,1)>0 && AAA(1,2)<1 && AAA(1,3)<1 && AAA(1,4)<1 && AAA(1,5)<1
    F(1,1)=1;
    T=' Fault no. 1 can be detected ... !'
end

if AAA(2,1)<1 && AAA(2,2)>0 && AAA(2,3)<1 && AAA(2,4)<1 && AAA(2,5)<1
    F(2,1)=1;
    T=' Fault no. 2 can be detected ... !'
end

if AAA(3,1)<1 && AAA(3,2)<1 && AAA(3,3)>0 && AAA(3,4)<1 && AAA(3,5)<1
```

```matlab
    F(3,1)=1;

    T=' Fault no. 3 can be detected ... !'

end


if AAA(4,1)<1 && AAA(4,2)<1 && AAA(4,3)<1 && AAA(4,4)>0 && AAA(4,5)<1

    F(4,1)=1;

    T=' Fault no. 4 can be detected ... !'

end




if AAA(5,1)<1 && AAA(5,2)<1 && AAA(5,3)<1 && AAA(5,4)<1 && AAA(5,5)>0

    F(5,1)=1;

    T=' Fault no. 5 can be detected ... !'

end


FF(:,kk)=F

kk=kk+1

end


xlswrite('120FFTSegmentsLoop.xls',FF, 'FF Matrix-S=2-2.9Kto121-121.9K');

xlswrite('120FFTSegmentsLoop.xls',X0_Freq, 'Division Values (S)');
```

MATLAB code for the NN- based fault identification algorithm

```matlab
clear

clc

% Neural Network  training.


% loading the data files (4 runs for each condition)


data1=load('Healthy\1\1Impeller_ data matrix.mat');

data2=load('Healthy\2\1Impeller_ data matrix.mat');

data3=load('Healthy\3\1Impeller_ data matrix.mat');

data4=load('Healthy\4\1Impeller_ data matrix.mat');

data5=load('Healthy\5\1Impeller_ data matrix.mat');

data6=load('Healthy\6\1Impeller_ data matrix.mat');

data7=load('Healthy\7\1Impeller_ data matrix.mat');

data8=load('Healthy\8\1Impeller_ data matrix.mat');

data9=load('Healthy\9\1Impeller_ data matrix.mat');

data10=load('Healthy\10\1Impeller_ data matrix.mat');


data11=load('Leak\1\1Impeller_ data matrix.mat');

data12=load('Leak\2\1Impeller_ data matrix.mat');

data13=load('Leak\3\1Impeller_ data matrix.mat');

data14=load('Leak\4\1Impeller_ data matrix.mat');

data15=load('Leak\5\1Impeller_ data matrix.mat');

data16=load('Leak\6\1Impeller_ data matrix.mat');

data17=load('Leak\7\1Impeller_ data matrix.mat');

data18=load('Leak\8\1Impeller_ data matrix.mat');
```

```
data19=load('Leak\9\1Impeller_ data matrix.mat');

data20=load('Leak\10\1Impeller_ data matrix.mat');


data21=load('Impeller\1\1Impeller_ data matrix.mat');

data22=load('Impeller\2\1Impeller_ data matrix.mat');

data23=load('Impeller\3\1Impeller_ data matrix.mat');

data24=load('Impeller\4\1Impeller_ data matrix.mat');

data25=load('Impeller\5\1Impeller_ data matrix.mat');

data26=load('Impeller\6\1Impeller_ data matrix.mat');

data27=load('Impeller\7\1Impeller_ data matrix.mat');

data28=load('Impeller\8\1Impeller_ data matrix.mat');

data29=load('Impeller\9\1Impeller_ data matrix.mat');

data30=load('Impeller\10\1Impeller_ data matrix.mat');


data31=load('Belt\1\1Impeller_ data matrix.mat');

data32=load('Belt\2\1Impeller_ data matrix.mat');

data33=load('Belt\3\1Impeller_ data matrix.mat');

data34=load('Belt\4\1Impeller_ data matrix.mat');

data35=load('Belt\5\1Impeller_ data matrix.mat');

data36=load('Belt\6\1Impeller_ data matrix.mat');

data37=load('Belt\7\1Impeller_ data matrix.mat');

data38=load('Belt\8\1Impeller_ data matrix.mat');

data39=load('Belt\9\1Impeller_ data matrix.mat');

data40=load('Belt\10\1Impeller_ data matrix.mat');
```

```matlab
data41=load('Both\1\1Impeller_ data matrix.mat');

data42=load('Both\2\1Impeller_ data matrix.mat');

data43=load('Both\3\1Impeller_ data matrix.mat');

data44=load('Both\4\1Impeller_ data matrix.mat');

data45=load('Both\5\1Impeller_ data matrix.mat');

data46=load('Both\6\1Impeller_ data matrix.mat');

data47=load('Both\7\1Impeller_ data matrix.mat');

data48=load('Both\8\1Impeller_ data matrix.mat');

data49=load('Both\9\1Impeller_ data matrix.mat');

data50=load('Both\10\1Impeller_ data matrix.mat');


X1=abs(fft(data1.data(2000:122000)));

X2=abs(fft(data2.data(2000:122000)));

X3=abs(fft(data3.data(2000:122000)));

X4=abs(fft(data4.data(2000:122000)));

X5=abs(fft(data5.data(2000:122000)));

X6=abs(fft(data6.data(2000:122000)));

X7=abs(fft(data7.data(2000:122000)));

X8=abs(fft(data8.data(2000:122000)));

X9=abs(fft(data9.data(2000:122000)));

X10=abs(fft(data10.data(2000:122000)));


X11=abs(fft(data11.data(2000:122000)));

X12=abs(fft(data12.data(2000:122000)));

X13=abs(fft(data13.data(2000:122000)));
```

```
X14=abs(fft(data14.data(2000:122000)));

X15=abs(fft(data15.data(2000:122000)));

X16=abs(fft(data16.data(2000:122000)));

X17=abs(fft(data17.data(2000:122000)));

X18=abs(fft(data18.data(2000:122000)));

X19=abs(fft(data19.data(2000:122000)));

X20=abs(fft(data20.data(2000:122000)));


X21=abs(fft(data21.data(2000:122000)));

X22=abs(fft(data22.data(2000:122000)));

X23=abs(fft(data23.data(2000:122000)));

X24=abs(fft(data24.data(2000:122000)));

X25=abs(fft(data25.data(2000:122000)));

X26=abs(fft(data26.data(2000:122000)));

X27=abs(fft(data27.data(2000:122000)));

X28=abs(fft(data28.data(2000:122000)));

X29=abs(fft(data29.data(2000:122000)));

X30=abs(fft(data30.data(2000:122000)));


X31=abs(fft(data31.data(2000:122000)));

X32=abs(fft(data32.data(2000:122000)));

X33=abs(fft(data33.data(2000:122000)));

X34=abs(fft(data34.data(2000:122000)));

X35=abs(fft(data35.data(2000:122000)));

X36=abs(fft(data36.data(2000:122000)));
```

```
X37=abs(fft(data37.data(2000:122000)));

X38=abs(fft(data38.data(2000:122000)));

X39=abs(fft(data39.data(2000:122000)));

X40=abs(fft(data40.data(2000:122000)));


X41=abs(fft(data41.data(2000:122000)));

X42=abs(fft(data42.data(2000:122000)));

X43=abs(fft(data43.data(2000:122000)));

X44=abs(fft(data44.data(2000:122000)));

X45=abs(fft(data45.data(2000:122000)));

X46=abs(fft(data46.data(2000:122000)));

X47=abs(fft(data47.data(2000:122000)));

X48=abs(fft(data48.data(2000:122000)));

X49=abs(fft(data49.data(2000:122000)));

X50=abs(fft(data50.data(2000:122000)));


inputs=[X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14 X15 X16 X17 X18
X19 X20 X21 X22 X23 X24 X25 X26 X27 X28 X29 X30 X31 X32 X33 X34 X35 X36
X37 X38 X39 X40 X41 X42 X43 X44 X45 X46 X47 X48 X49 X50 ];
targets=[1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 ;

    0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0;

    0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0;
```

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 0 0 0 0 0

0 0 0 0 0;

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1

1 1 1 1 1];

```matlab
% Solve the Pattern Recognition Problem with Neural Network


% Create a Pattern Recognition Network
hiddenLayerSize = 50;

net = patternnet(hiddenLayerSize);


% Set up Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 40/100;

net.divideParam.valRatio = 10/100;

net.divideParam.testRatio = 50/100;


 % Train the Network
[net,tr] = train(net,inputs,targets);


% Test the Network
outputs = net(inputs);

errors = gsubtract(targets,outputs);

performance = perform(net,targets,outputs)
```

% View the Network

view(net)

% Plot

figure, plotconfusion(targets,outputs)

APPENDIX (E)

Axial and radial vibration signals comparison

In this section, the axial and radial vibration signals are compared. The selection of the best signal is based on the results of the comparison. The RMS values of radial and axial vibration signals are presented in Table 1. The results show that it is very difficult to identify the addressed machine faults based on the RMS values as the vast majority of RMS ranges overlap. However, the axial vibration signal could be better than the radial one as it can be used to identify MC 3.

Table 1: RMS values of radial and axial vibration signals of training datasets.

Radial vibration signals

	Min. RMS (Volt)	Max. RMS (Volt)
MC 1	1.628	1.664
MC 2	1.668	1.671
MC 3	1.627	1.675
MC 4	1.628	1.67
MC 5	1.641	1.701

Axial vibration signals

	Min. RMS (Volt)	Max. RMS (Volt)
MC 1	1.629	1.663
MC 2	1.660	1.662
MC 3	1.823	1.872
MC 4	1.626	1.668
MC 5	1.630	1.685

The Confidence Level which informs the minimum number of matching features between spectrum peaks of all machine conditions was also calculated to better select the signals with large differences. Based on the pre-identification matrices presented in Table 2, the difference between the confidence levels of radial and axial vibration spectra is very small. Due to the RMS results and the pre-identification matrices calculations, axial vibration signals will be employed in this characterisation study.

Table 2: Fault pre-identification matrices at a segment size of 10 Hz.

Radial Vibration signals

	MC 1	MC 2	MC 3	MC 4	MC 5
MC 1	**191**	127	129	145	92
MC 2	125	**199**	114	129	93
MC 3	136	114	**196**	115	76
MC 4	138	133	134	**186**	90
MC 5	104	98	115	115	**200**

Overall confidence Level = 41

Axial vibration signals

	MC 1	MC 2	MC 3	MC 4	MC 5
MC 1	**199**	116	130	113	79
MC 2	132	**204**	123	125	97
MC 3	125	111	**187**	134	107
MC 4	119	115	139	**196**	117
MC 5	108	111	102	113	**201**

Overall confidence Level = 48

APPENDIX (F)

Selected publications

List of publications

Six peer-reviewed papers were successfully written out of this study. The papers are as follows:

Journal publications (5 papers)

S. Gowid, R. Dixon and S. Ghani, "Optimization of reliability and maintenance of liquefaction system on FLNG terminals using Markov modelling," International Journal of Quality & Reliability Management, Emerald, vol. 31, no. 3, pp. 293-310, 2014

S. Gowid, R. Dixon and S. Ghani, "A Novel Robust Automated FFT-Based Segmentation and Feature selection Algorithm for Acoustic Condition monitoring Systems," Journal of Applied Acoustics, Elsevier , vol. 88, no. 1, pp. 66-74, 2015

S. Gowid, R. Dixon and S. Ghani, "Profitability, reliability and condition monitoring of LNG floating platforms: A review" Journal of Natural Gas Science & Engineering, Elsevier, vol 27, no. 3, pp. 1495-1511, 2015

S. Gowid, R. Dixon and S. Ghani, "Characterization of major fault detection features and techniques for the condition monitoring of high speed centrifugal equipment," Journal of Acoustics and Vibration, vol 27, no. 2, pp. 184-191, 2016

S. Gowid, R. Dixon and S. Ghani, "Performance Comparison between FFT Based Segmentation Algorithm and Neural Network for the Condition monitoring of Centrifugal Equipment" Journal of Dynamic Systems, Measurement and Control, ASME, vol 139, no. 6, doi: 10.1115/1.4035458, 2017

Conference publications (1 paper)

S. Gowid, R. Dixon and S. Ghani, "Assessment of liquefaction systems' process performance on LNG Floating export Terminals[*]," International conference on Mechanical, Automotive and Aerospace Engineering (ICMAAE 2013), Paper # 30111, 2-4 July 2013, Kula Lumpur, Malaysia

 *The paper was shortlisted by the conference committee for publication.