THE USE OF ADVANCED SOFT COMPUTING FOR MACHINERY CONDITION MONITORING

MAHMUD AHMED

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Huddersfield

December 2014

Abstract

The demand for cost effective, reliable and safe machinery operation requires accurate fault detection and classification. These issues are of paramount importance as potential failures of rotating and reciprocating machinery can be managed properly and avoided in some cases. Various methods have been applied to tackle these issues, but the accuracy of those methods is variable and leaves scope for improvement.

This research proposes appropriate methods for fault detection and diagnosis. The main consideration of this study is use Artificial Intelligence (AI) and related mathematics approaches to build a condition monitoring (CM) system that has incremental learning capabilities to select effective diagnostic features for the fault diagnosis of a reciprocating compressor (RC).

The investigation involved a series of experiments conducted on a two-stage RC at baseline condition and then with faults introduced into the intercooler, drive belt and 2nd stage discharge and suction valve respectively. In addition to this, three combined faults: discharge valve leakage combined with intercooler leakage, suction valve leakage combined with intercooler leakage and discharge valve leakage combined with suction valve leakage were created and simulated to test the model. The vibration data was collected from the experimental RC and processed through pre-processing stage, features extraction, features selection before the developed diagnosis and classification model were built.

A large number of potential features are calculated from the time domain, the frequency domain and the envelope spectrum. Applying Neural Networks (NNs), Support Vector Machines (SVMs), Relevance Vector Machines (RVMs) which integrate with Genetic Algorithms (GAs), and principle components analysis (PCA) which cooperates with principle components optimisation, to these features, has found that the features from envelope analysis have the most potential for differentiating various common faults in RCs.

The practical results for fault detection, diagnosis and classification show that the proposed methods perform very well and accurately and can be used as effective tools for diagnosing reciprocating machinery failures.

Dedication

To my wife, my sons and my daughter, for your patience and absolutely everything with love.

To my dad, my brothers and sisters, for their continuous support and encouragement while I have been away from home.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A	Ampere
AC	Alternating Current
ADC	Analog-Digital Converter
AI	Artificial Intelligence
AM	Amplitude Modulation
ANNs	Artificial Neural Networks
B&K	Bruel & Kjaer
BDC	Bottom Dead Center
BL	Loose Drive Belt
CBM	Condition Based Maintenance
CED	Cambridge Electronic Design
CM	Condition Monitoring
CP	Cylinder Pressure
CT	Current Transformer
DAG	Direct Acyclic Graph
DAQ	Data Acquisition
dB	Decibel
DC	Direct Current
DFT	Discrete Fourier Transform
DV	Discharge Valve
DVC	Discharge Valve Close
DVL	Discharge Valve Leakage
DVL+IL	Discharge Valve leakage with Intercooler combined fault
DVL+SVL	Discharge Valve leakage with Suction Valve leakage combined fault
DVO	Discharge Valve Open
DWT	Discrete Wavelet Transforms
ES	Expert system
FDD	Fault Detection and Diagnosis
FFT	Fast Fourier Transforms
FLS	Fuzzy logic system
GA	Genetic Algorithm
GUI	Graphical User Interface
H	Healthy
HAH	Half Against Half
HOS	Higher Order Spectra

hp	Horse power
HP	High pressure
Hz	Hertz
I/O	Input/Output
IAS	Instantaneous Angular Speed
Inr-L	Intercooler Leakage
IQR	Interquartile Range
IR	Infrared Thermography
LB	Lower Bound
LP	Low Pressure
MCSA	Motor Current Signature Analysis
MQC	Multivariate Quality Control
mRVM	Multiclass Relevance Vector Machine
Nnl	Normal negative log-likelihood
OAA	One Against All
OAO	One Against One
P.u	Per unit
PA	Atmospheric Pressure
PC	Personal Computer
PCA	Principle Components Analysis
PCB	Printed Circuit Board
PNN	Probabilistic Neural Network
Psi	Pounds per square inch
QPC	Quadratic Phase Coupling
Rad/s	Radians per second
RC	Reciprocating air compressor
RF	Rotational Frequency
RMS	Root Mean Square
RVM	Relevance Vector Machine
SPE	Squared prediction error
SV	Suction Valve
SVC	Suction Valve Close
SVL	Suction Valve leakage
SVL+IL	Suction Valve leakage with Intercooler combined fault
SVM	Support Vector Machine
SVO	Suction Valve Open
TDC	Top Dead Centre
UB	Upper Bound

V   Volte

WNNs   Wavelet Neural Networks

List of Notations

A_{fd} Maximum, flow area of the discharge valve $[mm^2]$

A_{fl} Maximum, flow area of the suction valve $[mm^2]$

A_L Leakage valve size $[mm^2]$

B_r Transmission ratio $[\%]$

C Damping constant of valve chamber $[N/ms^{-1}]$

C_i Speed of sound in the suction plenum $[ms^{-1}]$

cfs Force coefficient of suction valve

cfd Force coefficient of discharge valve

$C_{di}(x)$ Variable suction coefficient

$C_{dd}(x)$ Variable discharge coefficient

C_c SV Contact damping coefficients $[N.m^{-1}]$

C_{cd} DV Contact damping coefficients

C_{cs} SV Contact damping coefficients

C_d DV Contact damping coefficients $[N.m^{-1}]$

D Piston diameter $[mm]$

d Cylinder bore length $[mm]$

F Cylinder air pressure force $[N]$

fso Preset spring force of suction valve $[N]$

fdo Preset spring force of discharge valve $[N]$

fvd Discharge pressure force $[N]$

fvs Suction pressure force $[N]$

fgs Gravitation force for suction valve $[N]$

$f_{pL,H}$ Force produced by the air pressure $[N]$

fgd Gravitation force for discharge valve $[N]$

$f_{mL,H}$ Force produced by the inertial force of reciprocating mass $[N]$

g Earth's gravitational field strength $[9.8Nkg^{-1}]$

J Moment of inertia $[kgs^2]$

m_{rec} Reciprocating inertial mass $[g]$

$\dot{m}_{vi}$ Mass flow rate through the suction valve $[kgs^{-1}]$

$\dot{m}_{vd}$ Mass flow rate through the discharge valve $[kgs^{-1}]$

$\dot{m}_{ls}$ Mass flow rate of leakage suction valve $[kgs^{-1}]$

m Mass of air inside cylinder $[g]$

m_v Mass of the valve plate plus one-third of the spring mass $[g]$

ω IAS of the crank shaft $[rad/sec]$

γ Ratio of major specific heats of the process gas

ρ_c Density of the air in the cylinder $[kg/m^3]$

ρ_i Density of the air in the plenum $[kg/m^3]$

θ	Crank angle	[deg]
j	Moment of intertie	$[kg/m^2]$
k	Valve spring stiffness	$[N.m^{-1}]$
l	Connecting rod length	[mm]
m_p	Piston mass	[g]
m_{cr}	Connecting rod mass	[g]
m_{plate}	Mass of valve plate	[g]
m_{spring}	Mass of valve spring	[g]
T_m	Driving torque	[N.m]
$T_{pmL,H}$	Torque produced by cylinder air pressure and piston mass	[N.m]
$T_{fL.H}$	Friction torque	
$\dot{v}$	Variable cylinders volume	$[m^3]$
m_{co}	Cylinder clearance volume	$[m^3]$
v_c	Cylinder volume	$[m^3]$
ω	IAS of crank shaft	[rad/sec]
ω_v	Valve unit frequency	
ω_n	Natural frequency of the valve unit	[rad/sec]
x_{max}	Max valve plate displacement	[mm]
ω_s	The motor speed	[rad/sec]
XS_{max}	Suction valve max lift	[mm]
Xd_{max}	discharge valve max lift	[mm]
x_v	Valve plate displacement	[mm]
$\dot{x}_v$	Valve speed	$[ms^{-1}]$
$\ddot{x}_v$	Valve acceleration	$[ms^{-2}]$
x_p	Piston displacement	[mm]
$\dot{x}_p$	Vertical piston speed	$[ms^{-1}]$
$\ddot{x}_p$	Vertical piston acceleration	$[m.sec^{-2}]$
γ	Ratio of specific heats (1.4 for air)	
p_c	Density of the air in the cylinder	$[kg/m^3]$
p_i	Density of the air in the plenum	$[kg/m^3]$
T^2	Hotelling's statistic	
E	Residual error	

Declaration

No portion of the work presented in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

The Author

Mahmud Ahmed graduated from Tripoli University, with a B.Sc. in Computer Science in 1992 and worked in the Ministry of Higher Education from 1993 to 1999 as a programming engineer. In September 2000 he started his Master's Degree in Information Technology for Engineering at the Coventry University and graduated in September 2002.

He also worked at Musrata University as a lecturer assistant from 2004 to 2008.

In April 2009 he commences the research on CM that is presented in this thesis in the Centre for Efficiency and Performance Engineering (CEPE) at the University of Huddersfield.

Publications

1. **Ahmed, Mahmud**, Gu, Fengshou and Ball, Andrew (2010) Fault classification using an Artificial Neural Network based on Vibrations from a Reciprocating Compressor. In: Future Technologies in Computing and Engineering: Proceedings of Computing and Engineering Annual Researchers' Conference 2010: CEARC'10. University of Huddersfield, Huddersfield, pp. 92-97. ISBN 9781862180932.

2. **Ahmed, Mahmud**, Gu, Fengshou and Ball, Andrew (2011) Feature Selection and Fault Classification of Reciprocating Compressors using a Genetic Algorithm and a Probabilistic Neural Network. Journal of Physics: Conference Series, 305 (1). 012112. ISSN 1742-6596.

3. Abdusslam, S.A, **Ahmed, Mahmud**, Raharjo, Parno, Gu, Fengshou and Ball, Andrew (2011) Time Encoded Signal Processing and Recognition of Incipient Bearing Faults. IEEE In: Proceedings of the 17th International Conference on Automation & Computing. Chinese Automation and Computing Society, Huddersfield. ISBN 978-1-86218-098-7.

4. **Ahmed, Mahmud**, Abdusslam, S, Baqqar, Mabrouka, Gu, Fengshou and Ball, Andrew (2011) Fault Classification of Reciprocating Compressor Based on Neural Networks and Support Vector Machines. IEEE In: Proceedings of the 17th International Conference on Automation & Computing. Chinese Automation and Computing Society, Huddersfield. ISBN 978-1-86218-098-7.

5. Baqqar, Mabrouka, **Ahmed, Mahmud** and Gu, Fengshou (2011) Data Mining for Gearbox Condition Monitoring. IEEE In: Proceedings of the 17th International Conference on Automation & Computing. Chinese Automation and Computing Society, Huddersfield. ISBN 978-1-86218-098-7.

6. **Ahmed, M**, Gu, Fengshou and Ball, Andrew (2012) Fault Detection of Reciprocating Compressors using a Model from Principles Component Analysis of Vibrations. Journal of Physics: Conference Series, 364. 012133. ISSN 1742-6596.

7. **Ahmed, M**, Gu, Fengshou and Ball, Andrew (2012) Fault Detection and Diagnosis using Principal Component Analysis of Vibration Data from a Reciprocating Compressor. In: 18th International Conference On Automation And Computing (UKAC), 2012. IEEE, Cardiff, UK, pp. 461-466. ISBN 978-1-4673-1559-3.

8. Baqqar, Mabrouka, Wang, Tie, **Ahmed, Mahmud**, Gu, Fengshou, Lu, Joan and Ball, Andrew (2012) A General Regression Neural Network Model for Gearbox Fault Detection

using Motor Operating Parameters. In: 18th International Conference On Automation And Computing (UKAC), 2012.IEEE, Cardiff, UK,pp. 584-588. ISBN 978-1-4673-1559-3.

9. Alwodai, Ahmed, Shao, Y, Yuan, X, **Ahmed, M**, Gu, Fengshou and Ball, Andrew (2013) Inter-Turn Short Circuit Detection Based on Modulation Signal Bispectrum Analysis of Motor Current Signals. In: Proceedings of the 19th International Conference on Automation and Computing (ICAC) 2013: Future Energy and Automation. IEEE. ICAC 2013. Brunel University, London, UK. ISBN 978-1908549082.

10. **M. Ahmed, A**. Smith, F. Gu, A.D. Ball (2014) Fault Diagnosis of Reciprocating Compressors Using Revelance Vector Machines with A Genetic Algorithm Based on Vibration Data. In: Proceedings of the 20th International Conference on Automation & Computing, Cranfield University, Bedfordshire, UK, 12-13 September 2014. IEEE. ICAC 2014.

Chapter 1

Introduction to Condition Monitoring

This chapter reviews the importance of CM and fault diagnosis for the smooth running of industrial processes. The sections of the chapter sequentially present the motivation for this research project, outline the aims and objectives and present the methodology and structure of this thesis.

1.1   Background and Motivation

CM is carried out in various industries to ensure the reliability and state of machine health. Principally, CM allows maintenance personnel to monitor machine health by measuring certain key variables that are related to machine element performance. This method stands in contrast to shutting down the machine to physically observe any signs of wear and suggest maintenance accordingly. CM can reduce the need for machine shutdowns for observation and maintenance and allows maintenance teams to take timely action to resolve outstanding causes of machine wear and possible failure.(Liu, 2008).

The use of CM can be seen as a development from preventive maintenance, which itself developed from break down maintenance. Modern process requirements demand greater availability and reliability of machines which can only be provided through accurate monitoring of machine health. This allows maintenance personnel to determine the best possible course of action based on knowledge available from CM (Abd Kadir Mahamad, 2010).

CM has found greater importance in maintenance circles based on savings and system simplification it provides. Not only does CM allow the operator to make correct and on time maintenance decisions, it also allows a reduction in maintenance costs. The improvements offered in terms of greater system availability also provide direct financial benefit to processes that cannot afford to have prolonged maintenance delays. Overall a sizable reduction in maintenance costs and direct fiscal benefits offered by more reliable machines has pushed CM to the forefront of maintenance globally (Fuqing, 2011).

CM can be carried out in a number of different ways ranging from the manual tabulation of manually measured variables to more complex and intelligent systems that offer diagnosed causes for machine wear. Over the years, CM has evolved significantly given the need to diagnose faults in larger and more dynamic industrial systems. There has been an increase in the use of AI and a number of mathematical techniques, such as PCA, in order to isolate faults and offer diagnosis for industrial systems.

1.1.1 Need for AI Applications in CM

AI techniques have been applied to a number of different industrial systems including CM. In order to drive up the reliability of the fault detection mechanisms, AI and PCA are applied. More notably, the neural networks have found pervasive application in CM systems. The application of AI for CM is required in areas where analytical knowledge is difficult to come across. The use of AI allows creation of new knowledge from existing knowledge and input data from monitored variables (Shi, 2004). The use of AI and PCA techniques is required since vibration data sets contain a lot of data which results in the creation of a large set of features. Optimal feature selection is only achievable through the application of AI and PCA approaches. A comparison of AI and PCA application versus conventional methods such as time domain, frequency domain and envelope analysis reveals that the former results in greater efficiency and savings. The application of conventional methods requires human resources with the appropriate expertise as well as significant time, which is a financial cost to the maintenance establishment. In contrast, the application of AI and PCA techniques allows for much faster and more reliable fault detection without the hassle of added costs. However, the variables measured from CM require treatment at various levels to process them into discernible and actionable knowledge.

1.1.2 Soft Computing AI and PCA Based Diagnosis Methods

The focus of the current research is to utilise an AI framework that is further classified using PCA so as to detect faults for a RC. A RC rig has been utilised in order to study vibration characteristics through the introduction of faults. The same rig was also utilised to detect faults based on the AI learning and PCA processing of recorded data.

Vibration recorded from various locations on the RC rig was used for feature extraction. Time domain, frequency domain and envelope analysis techniques were applied for feature extraction. Tools utilised in the time domain included (but were not limited to) maximum, minimum, averages, root mean square, kurtosis and other statistical techniques. The various features identified through feature extraction of recorded vibrations were then utilised for feeding the AI model. The contention was to use processed vibration data in order to enrich the AI so that it could be used later for fault diagnosis. A number of different AI models were utilised in order to test their efficacy for the provided RC system. More primitive AI models included ANN derivatives. In addition, GAs were utilised to test their efficacy for fault diagnosis. Moreover, SVM methods were used for fault diagnosis purposes too. The major schema revolved around the use of neural networks with GA alone and the use of neural networks combined with GA and SVM simultaneously. All of the aforementioned AI models served the purpose of both detecting and then classifying various fault conditions.

Investigation of AI models mentioned above revealed that GA provided the most reliable and optimal fault detection and classification. The inclusion of a large feature set meant that ANN and SVM were not as effective as GA for fault detection and classification. Results produced

through GA based implementation of AI produced the most optimal feature selection and hence fault diagnosis.

It needs to be mentioned here that fault detection and diagnosis also depends on PCA. Features selected previously from time domain was used as baselines to compare features extracted from test runs. The comparisons between both sets of data were performed through PCA which allowed fault detection and diagnosis. Major techniques used for PCA included the utilisation of the Q statistic, the T^2 statistic and contribution plots. It must be mentioned here that fault detection and optimal feature selection were carried out through differing methods. The Q statistic and the T^2 statistic were used only in order to perform fault detection. In comparison, the contribution plots were utilised only in order to select the optimal features that had been derived from the processes mentioned above. The research process is detailed in the chapters that follow.

1.2   Research Topic

The research topic of this thesis is "The Use of Advanced Soft Computing for Machinery Condition Monitoring". The main part of this thesis is to improve the accuracy of the diagnosis system based on features or information that are fed into the AI techniques and PCA method, therefore the selection of the features is very important. This research explores the possibility of improving this matter by using PNN, SVMs, and RVMs which integrate with GAs.

1.3   Aim of the Research

The aim of this work is to develop a CM technique capable of detecting and diagnosing faults seeded into a RC and analysis using advanced computing techniques, and to practically assess the technique developed using a laboratory test rig.

1.4   Research Objectives

The main objectives of this research are stated as follows:

1. Development of effective features for AI model and PCA methods.

2. Development of diagnosis, classification methods based AI and PCA techniques for RC compressor failure.

In order to provide effective features, it is required to establish many features during feature extraction. But the most important thing is providing the useful features during features selection. In this research for features extraction, features from both time and envelope spectrum are considered. Meanwhile for feature selection, the distance evaluation technique is used due to its simplicity. On the other hand, in the development of diagnosis, classification of RC failure, the AI and PCA techniques are used. For these purpose eight different sets (classes) of data are used for diagnosis, classification of machine failure.

To fulfil the research purpose, the following objectives have been formulated:

Objective 1: To construct a comprehensive RC test rig to simulate faults and obtain experimental data on consequent system behaviour. This work will allow CM of the system using different monitoring strategies, with the same user interface, and moreover, to seed faults into the system.

Objective 2: To develop a mathematical model of the two stages RC to be used for CM. The experimental results will be used to verify the model predictions.

Objective 3: To introduce specific quantified faults into the RC in the different parts and both measure and predict the effects on compressor valves performance.

Objective 4: To review the baseline of vibration data under different operating condition using conventional signal processing methods, which will be referenced for comparison with more advanced methods.

Objective 5: To develop automatic data processing methods based on advanced soft computing such as NNs, GAs, SVMs and RVMs.

Objective 6: To develop of an effective feature selection approach for RC fault diagnosis by using PCA.

Objective 7: To provide some useful information to guide future research in this field.

1.5   Research Methodology

The research method is driven by problem identification and solution. Developing a powerful and innovative framework with an open and flexible architecture enables the user to integrate current and future research.

The research in this area of machinery CM was started by conducting a literature review. The overall research plan for this topic addresses the development of a machinery CM system. The aim of the methodology is to approach the problem in a systematic way and to propose and develop an appropriate solution.

The methodology consists of three steps:

1. The case study (experimental work) is based on CM data collected from a two-stage RC test rig.

2. Feature extraction was employed using the conventional methods including: time, frequency domain and envelope spectrum.

3. Feature selection and classification were investigated using two different approaches: AI and PCA. A hybrid approach that combines GA technique with other AI approaches was also developed and investigated.

Procedure: The procedure adopted to implement this methodology was to:

1. Define or examine the type of data collected from the test rig and the signal processing techniques implemented to extract features from the collected data. Two data types were collected to cover the healthy condition of RC and the faulty conditions as well. The faulty data were generated by seeding different physical faults with in the Test rig.

The collected data from different tests was not averaged because the compressor will be under variable discharge pressures which influence vibration responses significantly. The test rig was selected as it is one of the most important industrial machines and as a model of real world applications.

2. The capabilities of the time, frequency domain and envelope spectrum in feature extraction of RC were investigated through comparison between the healthy and faulty conditions. These techniques are simple and can be implemented easily to obtain an overview of the system behaviour.

3. The variety of extracted features was fed into the PNN, SVM and RVM to find the classification rate. Then, the same features were utilised in combined approaches GA-PNN, GA-SVM and GA-RVM. GA was used in order to facilitate the identification of the optimal parameters that implied during execution of AI approaches. These approaches are characterises by their capability for training and do not require any mathematical model for the system under consideration.

4. The PCA was also used for fault identification and classification for RC utilising the features extracted from the statistical time domain. PCA is an accurate, fast and powerful technique that depends on mathematical representation of the system.

1.6   Structure of the Thesis

This thesis is structured into eleven chapters as follows:

Chapter 1: Provides the background information on CM, this chapter describes various CM technologies. The issues of CM technology are described and AI techniques that can be used to address some of these problems are mentioned and the present research topic, aim and objectives.

Chapter 2: This chapter provides an introduction to the RCs and how it works and describes the common types of valves and reviews RC failures experienced in industry. Finally, it highlights the use of different CM techniques in fault diagnosis and gives a brief literature review of AI techniques and PCA method for machinery fault detection and classification.

Chapter 3: In this chapter, the test rig and its facilities are described, and then a brief description is given of the transducers used: accelerometers, shaft encoder and pressure sensors. It also describes data acquisition and data management. Finally, it explains how local faults of valve leakage, belt looseness and intercooler leak were seeded.

Chapter 4: This chapter describes the mathematical model of the RC, including quantifying physical parameters, valve opening and closing times. It then briefly explains how the simulation was implemented and comparison of model predictions and system behaviour made.

Chapter 5: This chapter introduces how the features extraction and presents the use of vibration techniques in compressor fault detection and the relative merits of analysing compressor valve signals using time domain, frequency domain and envelope spectrum techniques.

Chapter 6: This chapter begins by giving a background to AI, including its theoretical basis. It also explains the new novel PNN, SVM, RVM integrated with GA and how to apply them for RC fault detection and classification.

Chapter 7: This chapter implements the AI techniques to the problem of CM. PNN and GAs are used to build an automatic CM system to find high classification rate and optimal features.

Chapter 8: This chapter summarises the results performance of SVM for binary and multi-classes in classification and selection the optimal features via GA.

Chapter 9: This chapter develops RVM via OAO scheme with GA feature optimisation for applying this binary classifier to the compressors data, and examines the harmonics selected for classification to find the insight of the classifier in associated with the physical supports. Finally, the performance of the multiclass multi-kernel mRVM is also explored for obtaining more efficient fault classification.

Chapter 10: This chapter introduces the results of PCA method and its applications for fault classification using the features extracted from time domain for RC.

Chapter 11: Summarises the findings of the work and gives suggestions for future research.

Chapter 2

Reciprocating Compressors, Their Faults and Condition Monitoring Techniques

This chapter starts by giving an introduction on RCs and how they work. It then describes the common types of valves and reviews RC failures experienced in industry, and identifies valves as a main cause, the intercooler and the belt drive. Next, it discusses the more common types of conventional CM techniques used for reciprocal compressors. Then, the literature reviews of AI and PCA that are mainly used in fault detection and classification problems have been reviewed. Finally, the pre-processing data includes: feature extraction and feature selection are discussed.

2.1   Reciprocating Compressors

RCs are among the few and common machines used for compressing and processing of gas in the oil and gas industry. The designing and manufacturing principles of these compressors have been tested and proven a high level of reliability. Nevertheless, the reliability of these compressors can be improved through CM. CM slows down these machines enabling them to operate at their full potential. CM also helps in deciding if it would be beneficial to change the defective part of the compressor instantly or wait for a certain period of time. One of the foremost components of a RC is the valve; there are basically two types of valves in a RC - suction and discharge valve. Both of these valves account for almost forty percent of the cases where the valves are required to be shut down, and these valves account for almost more than fifty percent of the cases where they are the total cost of repair (B.-S. Yang, Hwang, Kim, & Chit Tan, 2005). Both of these facts serve as the driving reason for making improvements in the process of CM for predicting valve failure and the severity. It is used in oil refineries, gas transportation (pipelines), chemical plants, and refrigeration plants. It is also used in the manufacturing and blow moulding of glass, plastic (polyethylene terephthalate or PET) and for liquids.
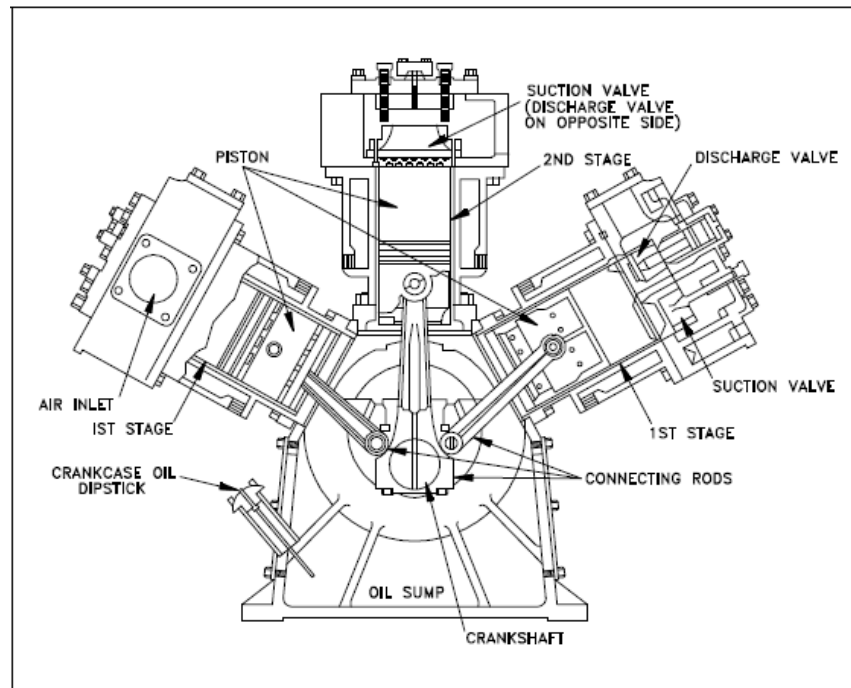
Figure 2.1 Cutaway view of the compressor (Bloch & Hoefner, 1996)

In most RCs, gaskets are used between adjacent parts to ensure leak proof conditions, because most of the surfaces are not machined finely enough to give a metal-tight body. Mainly gaskets are used between the cylinder head and the valve plate, between the valve plate and the compressor housing, between the compressor body and the bottom plate, if any, and also between the outer valves service and mounting bases. When the contact parts are tightly secured, shape and contour printed on the material, which is usually soft and resilient enough to take printing and thereby seal any oil or gas will possibly escape to the atmosphere, prevent penetration of air.

2.2   Working of a Reciprocating Compressor

A reciprocating piston uses the reciprocating action of a piston inside the cylinder for compressing the refrigerant. The process begins with the downward movement of the piston, which results in the creation of a vacuum inside the cylinder. During this process the pressure in the top intake and the pressure in the lower intake falls, due to this rise in pressure the intake valve is forced to open and once the valve of the intake opens up the refrigerant is then sucked inside the cylinder. Once the piston reaches its bottom point it then starts to move in an upward direction (Heinz & John, 1996). Afterwards, the intake valve then closes, leaving the refrigerant trapped inside the cylinder, and once the refrigerant is trapped inside the cylinder, the piston then continues to move in the upward direction to compress the refrigerant, which increases the amount of pressure inside the compressor. After a certain period of time there comes a point when the pressure exerted by the refrigerant is forced by the exhaust valve to open up and allows the compressed refrigerant to flow out of the cylinder.

This process is complete when the piston reaches its top most position, and afterwards it starts to move back in downward direction to its original position (Albert, Avery, Narin, & McAllister, 1991).
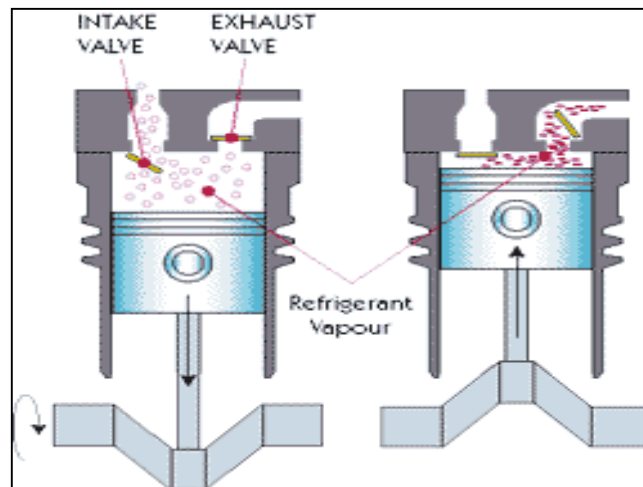
Figure 2.2 Air flow through a two-stage RC (Albert, et al., 1991)

RC valves can be opened and closed only by effect of cylinder pressure. This is achieved because the valves only open in one direction so that the intake valve opens in the direction towards the bottom dead centre of the cylinder and the discharge valve opens in the direction opposite to the intake valve. In the suction stage, the discharge valves are closed due to the vacuum effect created by the piston in its descending stroke (M Elhaj et al., 2008). Conversely, when the pressure has increased sufficiently, the discharge valve opens and the intake valve will be closed. The compressor duty cycle of the RC is divided into four stages that are defined below:

(A)   Top compression

In this phase of the cycle the cylinder is filled with gas.

(B)   Compression

In this phase the piston acts on the mass of its original volume reducing gas with a parallel increase in the pressure thereof. The cylinder valves are closed.

 (C)   Expulsion

Before the compression stroke is completed the exhaust valve opens. As a result of this the compressed gas is released from the cylinder, due to its own pressure through the valve. Before reaching the limit switch the discharge valve is closed, leaving the cylinder space filled with gas at discharge pressure.

 (D)   Expansion

In this phase of the cycle, both the exhaust valve and the inlet are closed. The piston starts moving from Reverse Run race and the gas contained within the cylinder undergoes a volume increase so that the pressure inside the system is reduced. Before reaching the final point the intake valve opens the cylinder.

(E)   Admission stage

In this phase of the cycle the piston goes back causing a depression in the cylinder which is compensated by the fresh gas inlet through the intake line. Just before reaching the lowest point of the race the intake valve closes, returning to state A) so that a new cycle begins.
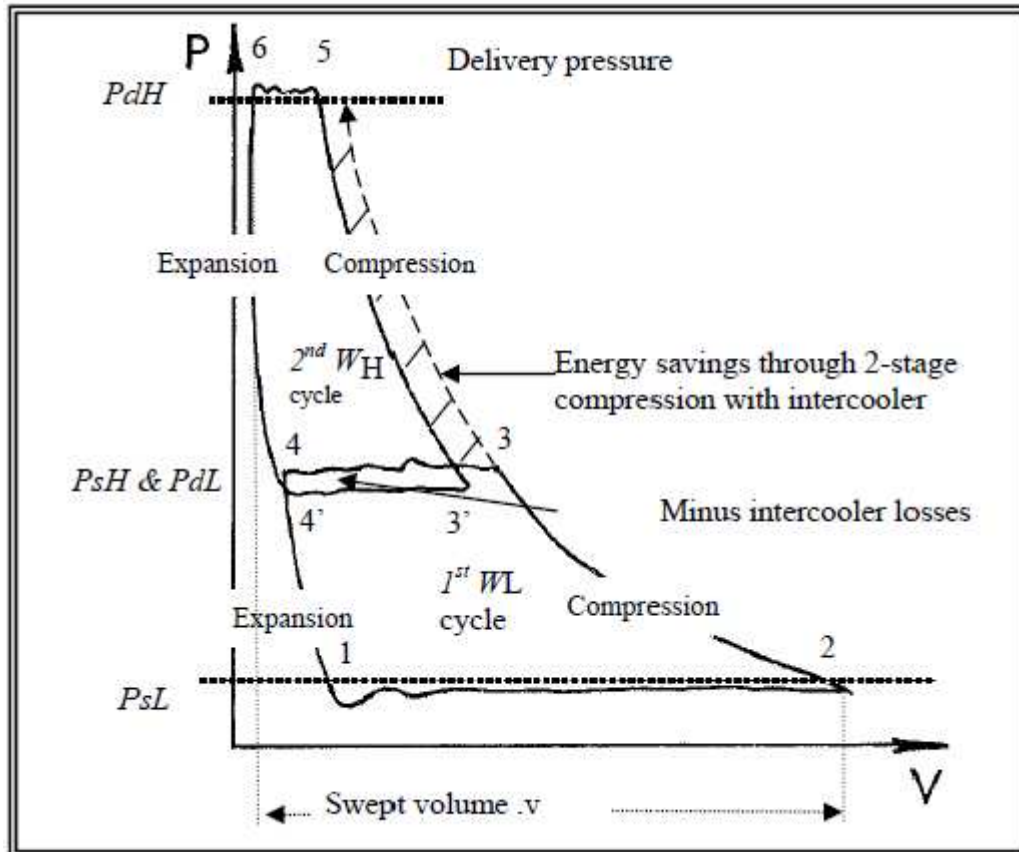
Figure 2.3 Different working stages of a RC (Zheng, 2005)

This type of compressor uses an automatic valve spring which opens only when sufficient pressure differential exists. Intake valves open when the pressure within the cylinder is slightly lower than the gas inlet pressure. The exhaust valve opens when the pressure in the cylinder is slightly higher than the pressure in the discharge line. Certain applications necessitate the use of high compression ratios (the ratio between the absolute intake pressure and gas absolute pressure the gas in the discharge) complicated by the need to verify in a single stage compression due to the high temperature reached by the gas in the discharge. Thus it is necessary to use serial compression that involves multiple compression stages. The gas is usually cooled between stages to reduce the temperature and volume before entering the next stage. It should be noted that each stage is constituted by a compressor in itself (Cho & Moon, 2005). This is designed and dimensioned to operate in series with one or more basic compressor elements, although they may all be fed from the same source but remain separate.

2.3   Types of Compressor Valves

Valves are automatic mechanisms placed on the suction and discharge of each of the cylinders, allowing gas flow in one direction, either into the cylinder (suction) or outwardly (drive). These mechanisms operate by pressure difference, although in certain conditions may be aided by springs. The following are different types of valves used in compressors:

2.3.1 Process Poppet Valves

The process poppet valve consists of a hole, usually round or oval, and a tapered plug, usually disc-shaped, placed at the end of a rod, also called a valve stem. There is a guide rod located around the valve that serves both to control the flow of a function element to all or nothing. The closure element rests on a seat ring and a circular section. As the closure is approaching the seat, the passage cross section is reduced and therefore increases the pressure loss by reducing the flow rate. In some applications, the pressure difference helps to close the valve, and other assistance to open it.
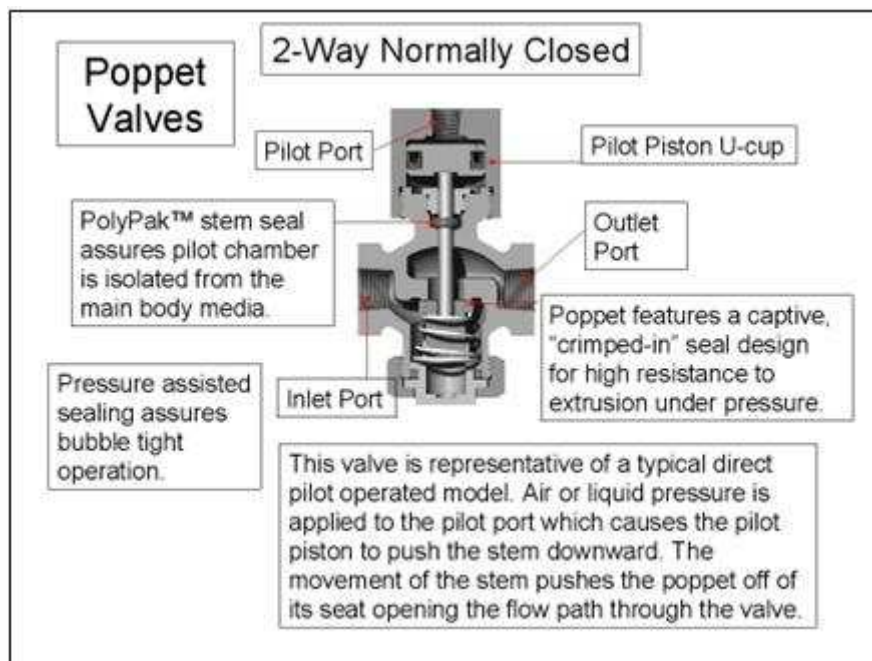
Figure 2.4 Process Poppet Valve (Dresser-Rand, 2005)

Poppet valves are very sturdy and resistant and commonly applied for industrial use in directional control valves. They are usually very tolerant to air contaminants (rust, dirt, etc) when used in compressed air service. They also allow high flows and high-speed operation. When the valve is actuated, it opens quickly and the area to pass the fluid is also large. The valve is operated by an actuator which in turn works on the valve stem, peeling off the disc from its seat to open (in the case of the normally closed valves) and allow passage of the fluid, or to support the disc in the seat and close (in the case of normally open valves), and to prevent the passage of fluid or gas. When the actuator stops applying force on the valve, a spring returns the valve to its original position (in the case of the normally closed valves),

causing the valve disc into contact with the seat; pressure fluid within the valve body helps keep it closed. In the case of normally open valves, fluid pressure assists the spring to separate the valve seat, opening the passage again.

2.3.2 HPS Valve

The HPS valves are one of the most versatile valves that are being used in compressors. The HPS valves consist of a number of circular plates that are moved with the displacement of the springs to which they are attached. These circular plates are then guided on the diameter, which enables the plates to hold their position over the slots in the seat. The buttons and rings of the HPS valves are made up of Hi-Temp© PEEK (Foreman, 2002). The HPS valves also contain a metallic ring that is capable of holding elements at very high temperature.

Figure 2.5 HPS valve (Foreman, 2002)

Throughout the lift, seat and exit areas , the HPS valves maintains a balanced flow, which results in the creation of a lower drop in pressure across the valve, limiting the loss of the valve and increasing the overall efficiency of the valve. The centre bolt of the valve is designed with a right hand thread on one half and a left hand thread on the other half. The internal components of the valves such as: springs, buttons, and rings are interchangeable between the discharge and the suction process. This minimises the maintenance and inventory cost.

2.3.3 PF Valves

PF valves are a ported plate type valve; these values are extremely versatile and can be used for an extensive amount of applications. The valves of the PF valves are joined with the help of radial ribs for the formation of a single plate. During the cycling process various coil springs make contact with the bottom of the plate, which results in the elimination of buttons on top of the springs. As a result of this process the compressor can perform operations at a very high speed, building the momentum and causing an efficient coil to coil contact in the spring.

Figure 2.6 PF valve (Foreman, 2002)

PF valves usually use a patented scalloped plate. Since there are very rare occasions in which these plates moves parallel to the seat, the plates located on the outside edge initially make contact with each other. This results in the creation of a higher amount of stress on the plate cracks and the outer edges. PF valves also consist of scallops located on the outside edges of the plate, just alongside the radial ribs which increase the cross sectional between the points of impact of the plates.

2.3.4 Reed Valves

Reed valves are commonly used in high performance versions of two-stroke engines, where it controls air-to-fuel mixture, its admission to the crankcase, and further up to the piston. Reed valves consist of a series of ports positioned adjacent to each other and each covered by a sheet with an extension which allows sufficient flow section even with small bending of the sheets.

Reed valves help to create an overpressure inside the compressor, since two-stroke engines do not have valves such as a four-stroke engine. Reed valves limit the flow of gas to one direction, opening and closing under changing pressure on each side. Modern versions usually consist of flexible metal or carbon fibre. The Reed valve sits between the carbonator and crank.
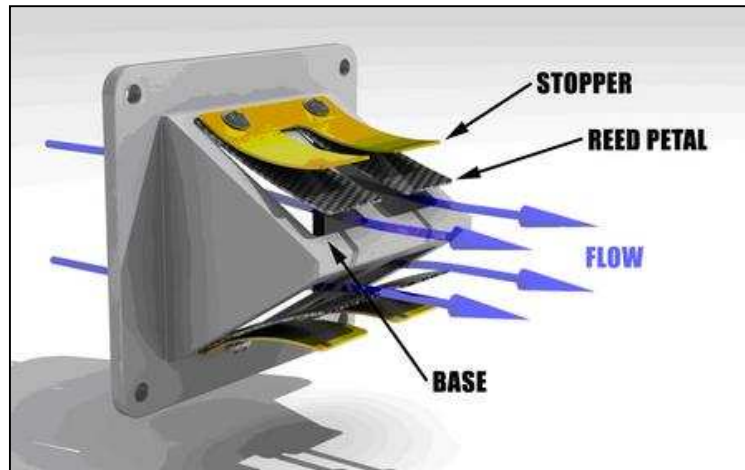
Figure 2.7 A diagram of the reed valve, showing the flexible reed petal, base, and stopper (Dhar, Tamma, Bhakta, & Krishna, 2014)

2.4   The Dynamics of Reciprocating Compressor Valves

The valve dynamics of a RC is determined by the distance between the seating and the plates (valve lift). The motion of the valve plate is simply determined by the help of the forces acting on these values. The calculation of the dynamics of the wall has been a normal routine process specifically at the design stage of the valves. This method is widely demonstrated and accepted for steel elements in industry (Quillen & Webster, 2001). To calculate the dynamics of a valve, the following three factors are considered along with the resulting force in Equation 2.1: the difference in pressure across the force area of the valve plate Av, the springing of the valve and lastly the contribution made by the viscous forces during the initial stages of the valve opening.

$$m_v \ddot{x}_v = (p_1 - p_2)A_v - k(x_v + l_1) - F_{adh}$$ Equation 2.1

Where:

m_v =Mass of the valve plates

k and l_1 = The stiffness of springing and initial deflection of the springs

p_1=Pressure in front of the valve

p_2=Pressure behind the valve

F_{adh}=Force

The force F_{adh} is obtained through the formula:

$$F_{adh} = f_1 \frac{\dot{x}_v}{{x_v}^3}$$ Equation 2.2

The factor f1 in Equation 2.2 depends completely on the geometric features of the properties and the value of the refrigerant. It also takes lubrication oil at the valve plate into account. These results are used for selecting the amount of valve lift required during the cycle, and the stiffness of a compressor. This process also determines the valve life, the pressure volume and the impact velocities on the seat and the guard. It also helps in the calculation of the total power loss and the total capacity caused by the valves during the cycle.

2.5   Main Failures in Reciprocating Compressors

2.5.1 Valve Failures

Valve failure in a RC can occur due to many reasons. Motriuk (Motriuk, 1996) indicates that the reason behind the improper installation of the valve during the compression cycles are a result of pressure pulsation inside the pipe, wrong selection of the valve meters and wide operating conditions. Valve failure can also occur due to corrosion as the valves of the compressor are frequently exposed to gases containing corrosive elements. During the compression cycle, these corrosive elements mix up with the gas and often cause the spring (or the moving elements inside the compressor) to halt permanently due to the fatigue caused by corrosion (Winandy, Saavedra O, & Lebrun, 2002).

Poor quality of gas can also result in valve failure, as the presence of debris and dirt in the gas causes accelerated wear. This results in the materialisation of poor sealing between the seat and the moving element, causing the guard and the flow holes in the seat to choke (not allowing the gas to flow through the valve). The liquids present in the gas steam also contribute to the valve failure, as large amount of liquids cause high impact slugging inside the compressor, resulting in reducing the effectiveness of the lubricant inside the compressor. In addition to this, it causes high amplitudes and pulsation that increases the frequency of the compressor, resulting in increased stress inside the compressor, and the potential for surge due to coil contacts on the valve springs. This effect causes a pressure drop that eventually causes the compressor unit to shut down.

Stiction is also among the many factors that could result in the failure of the valve. It occurs when a thin film of the oil causes the moving element to stick to the guard. This causes an additional amount of force for the spring to close the element once it is fully opened against the guard. When the spring is not strong enough, it will delay the close down, causing excessive impact velocity against the seat. The pressure force required for opening the moving elements will build up and result in excessive impact velocity and eventually halt the process when stiction is excessive.

2.5.2 Leak in Intercooler

In a RC the leak mostly occurs at the intercooler and during the process the clamps inside the compressor loosen and leak occurs. Leaks are mostly found at the bottom of the intercooler system and could be identified by pressurizing the compressor from 22 psi to 24 psi. These leaks can be detected through gentle wiggling of the pipes. There are certain circumstances where there are holes in the intercooler piping; there are other cases in which the leak occurs around the rubber pipes that are located inside the stock woven mesh pipes cracks (To, 1984).

2.5.3 Loose Drive Belt

In a RC there are certain situations in which the drive belt gets loose. This mostly happens when the belt is too tight around the body of the compressor. Under such circumstances the belt becomes overloaded, leading to wear and heat arising from friction as well as an increase in power consumption. In cases where the pressure increases significantly, the belt could break due to high tension. On the other hand, when the belt is too loose, the belt easily slips away from the compressor that leads to heat generation, unsteady rotation and abrasion on the belt that causes insufficient displacement of air inside the compressor (Iturriaga-Notario, 1978).

2.6   Overview of Practical CM Techniques

2.6.1 Vibration Monitoring

Vibration monitoring is one of the most commonly used techniques in CM whereby the condition of a machine is determined by analysing the vibration signals it generates. Every machine or process will produce vibrations in one form or another while in operation, and the vibration mechanisms of most machines are sufficiently understood enough to detect faulty operations from the vibration characteristics. The most commonly used transducer is the accelerometer which is particularly good for analysing rotating machinery due to its relative ease of use and low cost (Peters, 2002). Figure 2.8 shows trace of amplitude of the vibration detected by the accelerometer for the 2nd stage valves opening and closing under normal, healthy conditions and with 0.2%, and 0.8% discharge valve leakage. The 2nd stage discharge valve opens early at 185º for 0.2% valve leakage and 180º for 0.8% leakage, compared to a healthy valve which opens at about 192º. In this case the greater the leakage the greater the impact severity (M Elhaj, 2005).
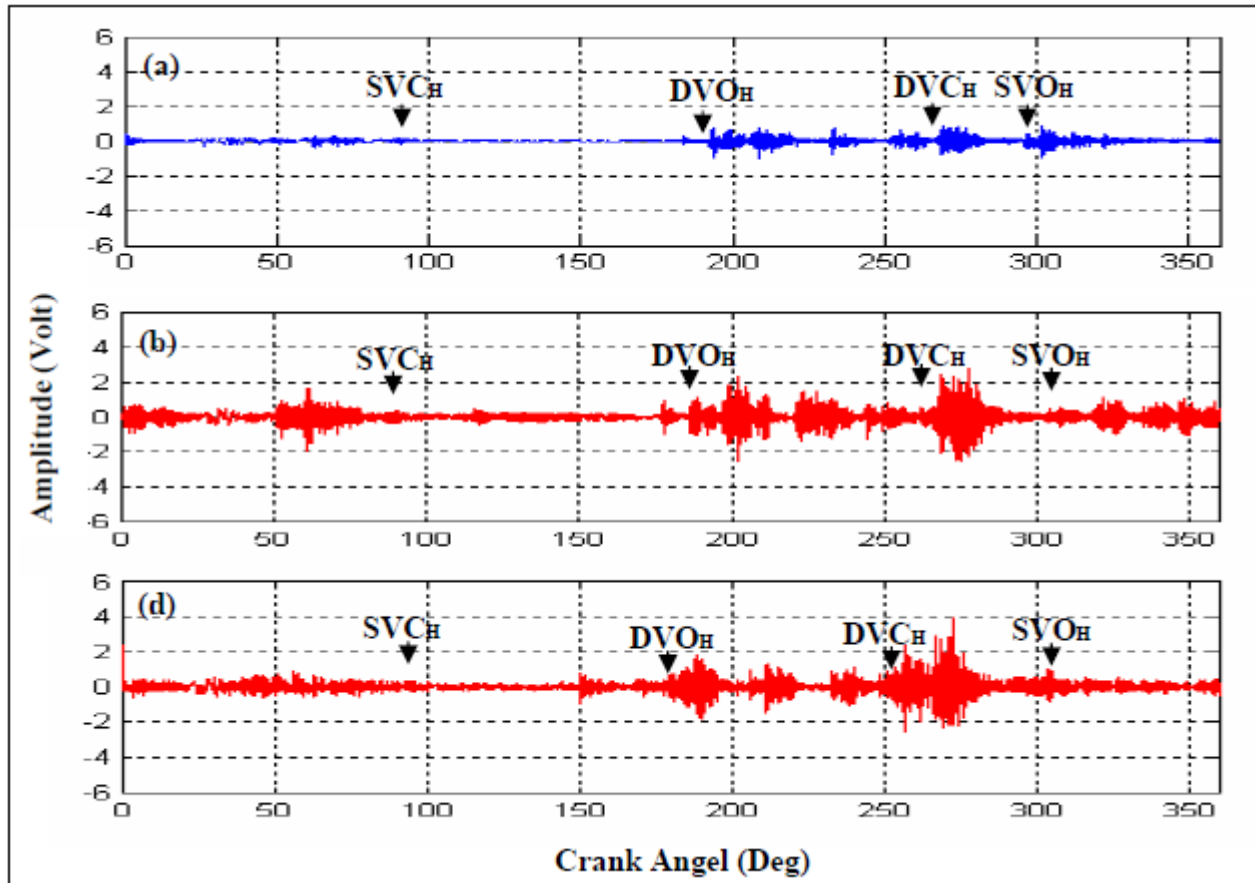
Figure 2.8 Time domain representation of vibration signal from accelerometer on the 2ⁿᵈ stage cylinder of a RC with valve leakage 2mm and without a leaky discharge (M Elhaj, 2005)

However, difficulties with vibration monitoring can occur due to the presence of multiple vibration sources within the machine for which the signals can combine in non-linear and possibly non-stationary ways.

2.6.2 In-cylinder Pressure Monitoring

In many machines such as diesel engines or RCs, the in-cylinder pressure signals usually have a lot of information on the condition and efficiency of the machine. The pressure transducer used for monitoring the cylinder pressure will, of course, depend on the machine. Figure 2.9 shows the dynamic pressure measured in the second stages of a RC with and without faults in the discharge valve.
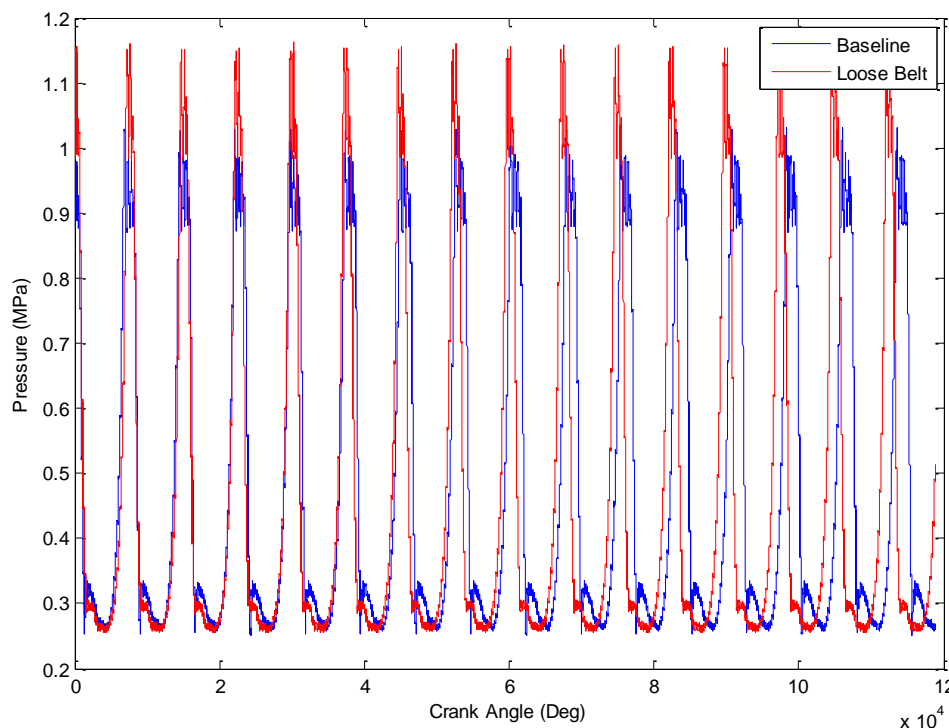
Figure 2.9 Measured cylinder pressures for a healthy second stage compressor and compressor with belt loosens (Naid, 2009)

Faults ranging from wear, sticking rings, to incorrect timing can be detected and diagnosed from cylinder pressure. For diesel engines an early indication of combustion problems can be obtained by measuring the in-cylinder pressure because of the strong relationship between cylinder pressure and power. In Figure 2.9 the typical cylinder pressure trace for a RC is shown in some detail. There is a rise in pressure as the piston moves "up" the cylinder. A short time before top dead centre the valve opens, causing a reduction in the rate of pressure rise in the cylinder. As the piston passes through TDC, the pressure in the cylinder begins to drop and there may by an oscillation in the cylinder pressure. The changes in pattern that occur with the presence of the given valve faults are clear.

2.6.3 Instantaneous Angular Speed Monitoring

The instantaneous angular speed (IAS) of reciprocating and rotating machinery will contain significant information about in-cylinder pressure, see Figure 2.10. This technique has been used to detect faults in fuel injection systems, in combustion processes and with valve leakage. This technique requires the use of speed sensors such as encoders, usually fitted onto the crankshaft of the flywheel. With a rotating compressor the IAS will vary over a working cycle; during the compression stroke the IAS will be less than during the expansion stroke. If the cylinder pressure is lower than it should be, the IAS will be faster than for the healthy case. The main disadvantage of this technique is that it is complicated to use. However, while it is a diagnostic technique that is useful for confirming the identification of faults detected by

other methods it has been used successfully to detect and identify faults such as leaky valves and deteriorating valve springs within a compressor, and rotor bar faults in an electric motor (Liang, 2000).

Considerable effort has been made on frequency domain analysis to obtain the IAS signal. These methods have been based on the Fast Fourier Transform (FFT) and rely largely on digital signal processing. In the near future, with the development of more advanced signal analysis techniques, it is expected that IAS would be more successful at fault detection and diagnosis (Yuhua Li et al., 2005).
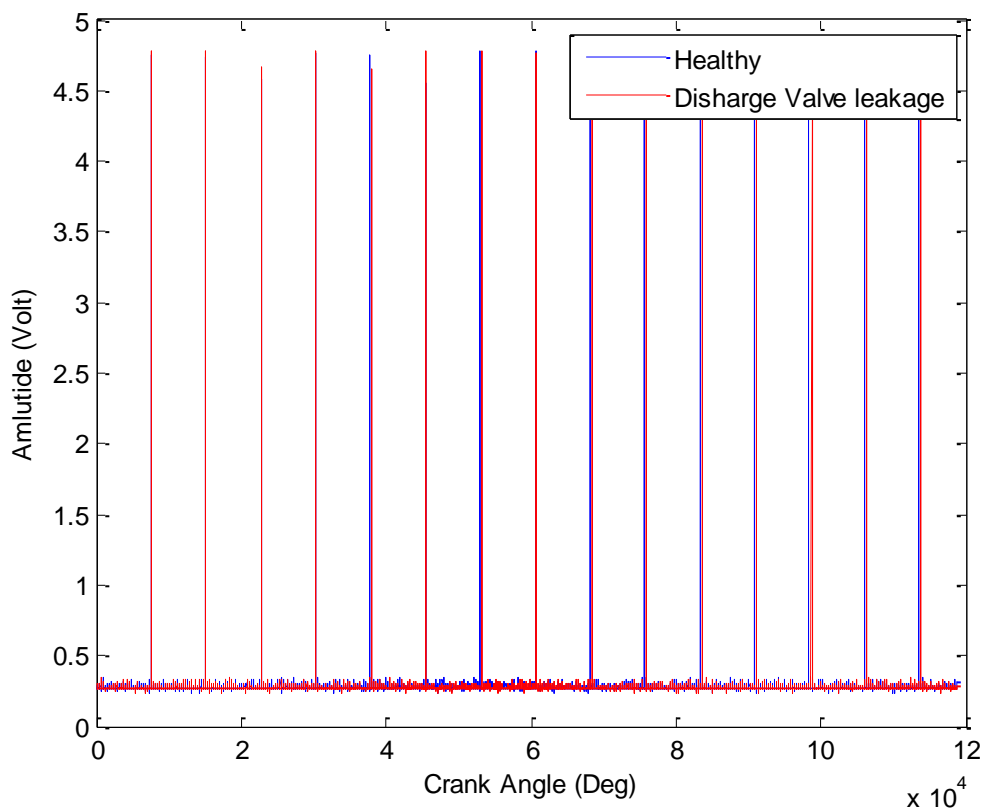
Figure 2.10 Measured cylinder pressures and Instantaneous Angular Speed for a healthy second stage compressor with small leak in discharge valve 2mm (Liang, 2000)

2.6.4 Acoustic Monitoring

Acoustic technology has been developed to include the field of CM. Acoustic analysis is now a recognised technique of non-destructive testing. This approach concentrates on the analysis of acoustic or noise waveforms produced by machinery processes (BenSasi, 2005). Microphones are often used to pick up acoustic signals to be compared against vibration-monitored waveforms. Microphones are sensitive, easy to mount and possess wide frequency response ranges that can give appropriate and comprehensive information. Therefore, microphones are utilised in many applications including gearbox, bearing, tool and engine CM (Erdelyi & Erie, 1956; Liang, 2000; Mayes, Steer, & Thomas, 1981) . On-line CM requires remote non-intrusive data collection, so the use of microphones is an attractive option. Moreover, acoustic

monitoring techniques offer a direct interpretation of vibration sources and generation mechanisms when the problem of ambient noise is to be addressed. However, because machines generate not only vibration but airborne noise as well, one of the major limitations of acoustic monitoring is the contamination of acoustic waveforms by background noise from similar adjacent machines (W. Li, 2000).

2.6.5 Infrared Thermography

Infrared thermography is now being targeted as a versatile tool for CM of equipment (Hung-Yi, Chao, & Tsai, 2005). IR imaging would enable fast and efficient approach in identifying the areas that are most in need of maintenance. The IR thermal imaging method utilise the radiant existence in the IR spectral band from measured objects to measure temperature. It is non-intrusive, applicable in remote areas and suitable for measurements of a large area. It can also serve to record data for subsequent storage and processing with a PC. Ay et al. (Ay, Jang, & Yeh, 2002) used an IR thermal imaging camera to observe the surface temperature of a plate finned-tube heat exchanger and calculated the local heat transfer coefficient. IR thermography on equipment shows that the surface thermal patterns are a consequence of internal conditions (Singh & Singh, 2011).
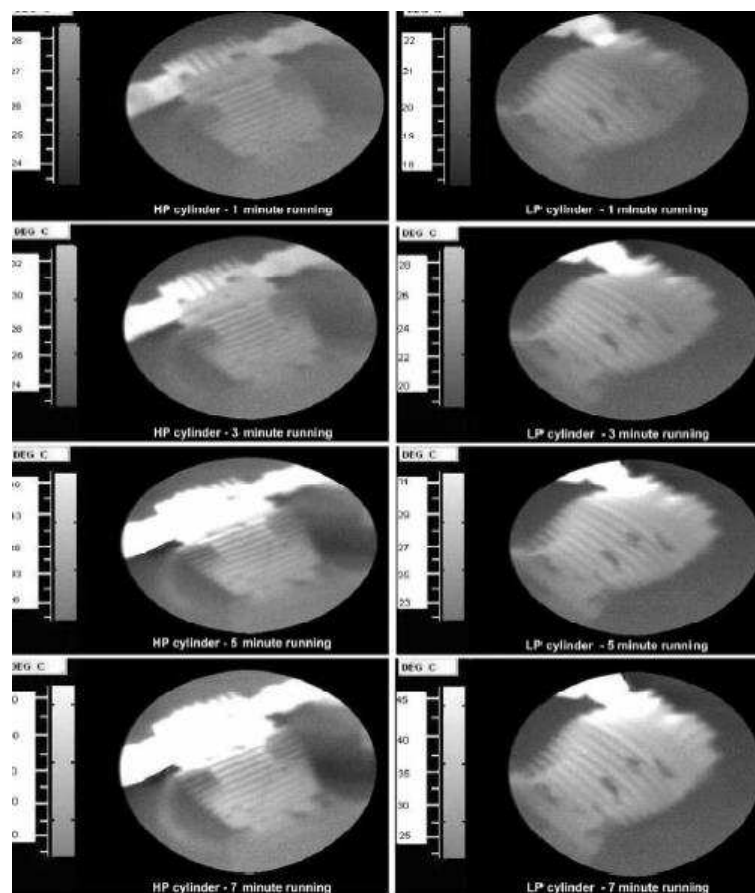
Figure 2.11 Thermographs of High Pressure and Low Pressure cylinders (Singh & Singh, 2011)

2.7 Literature review of AI Techniques and PCA Method for CM

Many AI techniques have been applied for fault diagnosis such ANN, SVM and RVM are briefly introduced with literature reviews. In addition, PCA method for fault detection and fault diagnosis are reviewed.

2.7.1 Neural Network

ANNs have potential application in the automated detection and diagnosis of machine condition. The use of the PNN has increased recently because it provides sound statistical confidence levels for its decisions. The goal of a PNN can be trained using test data to determine the features for the classifiers. These features have been calculated from the data obtained from the time and frequency domains and envelope analysis and were used for classifying the health of a RC. However, there remains the need to build an accurate and fast classification process using the optimal features which best characterize the system conditions and thus allow optimization of the parameters of the ANNs with a minimum number of features.

GAs have been used for automatic feature selection in machine CM (Samanta, Al-Balushi, & Al-Araimi, 2006). Moreover, GAs can be used to simultaneously find optimal structure of a spread value of PNN, in terms of concurrently determining the number of nodes in the hidden layers and connection matrices for evolving the ANN (Saxena & Saad, 2007).

Samanta and Al-Balushi (Samanta, et al., 2006) applied neural networks to the diagnosis of rolling element bearing faults using time-domain features from vibration signals. They also used GAs to optimize the features vector for gear fault detection using experimental vibration data from a gearbox (Samanta, Al-Balushi, & Al-Araimi, 2004). Orlowska-Kowalska (Kowalski & Orlowska-Kowalska, 2003) used a NN for induction motor fault diagnosis. Tiwari and Yadav (Tiwaria & Yadav, 2005) applied ANN in the CM of a defective RC using simulated data. Yang et al (B.-S. Yang, W.-W. Hwang, et al., 2005) presented classifiers including the self-organizing feature map (SOFM), learning vector quantization (LVQ), and SVM for fault features of a small RC. Yang et al., attempted to examine the performance of different multi-class SVM strategies on the diagnosis of faults in rotating machinery and compared the results with those obtained by ANNs (B.-S. Yang, Han, T., Hwang, W-W. , 2005). Jack et al examined the use of a GA to select the most significant input features from a large set of possible features in machine CM and shown that using a GA, a small subset of six allows more than 99% recognition accuracy, compared with an accuracy of 87.2% using an ANN without feature selection (L.B. Jack, Nandi., & McCormick., 2000). Worden et al. investigated GA in determining the parameters of Multi-Layer Perceptron (MLP) and selecting features for damage localisation in an aircraft wing (Worden K, January 2008), showing effectiveness of GA in performance improvement. Yang Genetic Programming (GP) used to combine optimally for engine valve fault detection (W.-X. Yang, 2006). Results obtained are better compared with Principal Component Analysis (PCA).

However, in these researches, more computational efficient NNs such as PNN have not been explored and a combination of ANN and GA has been received little attention for CM.

2.7.2 Support Vector Machines

SVMs have been introduced by Vapnik (V. N. Vapnik, 1995) and have become one of the most famous intelligent learning machines and a very good option to neural networks. SVMs are used for classification and regression analysis, data analysing, pattern recognition and nonlinear modelling. SVM is a popular and well-known method used for many applications such as signal processing , finding fault and control (Smola & Schölkopf, 2004).

(SVM) methods have been used along with grid search and other learning techniques. Similarly, SVM methods have been applied in tandem with GAs to obtain fault classification for fault prone software components. Comparably, SVM in tandem with GAs produced the best classification results, proving its superiority over other methods for fault classification (Martino, Ferrucci, Gravino, & Sarro, 2011). Mathematical modelling and testing of SVM methods with GA indicate their superiority over regular SVM methods in dealing with unbalanced classes to produce higher classification and faster learning (Amer, El-Garhy, Awadalla, Rashad, & Abdien, 2011). Hybrids of SVM methods such as combined SVM (CSVM) have been used extensively for process control such as in the Eastman process. Results indicate the superiority of SVM based methods over other methods of control (Tafazzoli & Saif, 2009).

SVM methods were employed in order to classify faults of reciprocating refrigeration compressors through the application of wavelet transform and statistical methods. Significant features were extracted from both raw noise signals and vibration signals. The selection of relevant RBF kernel parameters was carried out through iteration (B. S. Yang, Hwang, Kim, & Tan, 2005). In a similar application, SVM methods were applied to RCs butterfly valves to classify cavitation faults (B. S. Yang, Hwang, Ko, & Lee, 2005). Similar research was performed on RC valves to classify faults through vibration signals alone. Data for this purpose was gathered from the surface of the valve and the resulting vibration signals were decomposed by applying local wave methods (Ren, Ma, & Miao, 2005).

One of the larger problems posed by RC valves is the non-stationary and non-linear characteristics of the extracted vibration signals. In order to deal with the non-stationary and non-linear nature of such data, information entropy with good fault tolerance potential was utilised as the feature parameter fed to a SVM. This was utilised as being a comprehensive characteristic of the raw vibration signal. The resulting decision function was used to solve the limits of traditional fault classifications. The added strength of the SVM was its ability to be trained with only a few input samples to deal with multiple new faults (Zhigang Chen & Xiangjiao Lian, 2010).

The small linear pattern recognition performance and relatively small data sets extracted from RC valves present unique problems for fault classification. SVM has been utilised to deal with such limitations by employing information entropy since it is flexible as well as being liberal in

terms of non-linear behaviour. The SVM was trained by using a small vibration data set from RC valves characterised by information entropy. The fault classification results of SVM methods proved accurate enough for fault classification in valve failure of RCs (Cui, Zhang, Kang, & Lan, 2009).

Another limitation of vibration data from RCs is the small amount of fault data that can be extracted from regular runs compared to a large number of excitation sources. SVM was applied in tandem with Statistical Learning Theory (SLT) in order to overcome this challenge. Vibration signals were extracted from the rolling bearing in RC's crankcase through the utilisation of a test bed. A SLT scheme was developed to extract features that were then fed to the SVM for intelligent fault classification. Results showed that the application of these methods identified faults immediately with significant accuracy (Sheng, Jing, & Yabin, 2009).

2.7.3 Relevance Vector Machine

The RVM, introduced by M E. Tipping (M. E. Tipping, 2001), is a probabilistic sparse kernel model and is analogous to SVMs. It adopts a Bayesian approach to learning, by introducing a prior density over the weights, governed by a set of hyper parameters, whose most probable values are iteratively estimated from the data. Sparsity is achieved because in practice the posterior distributions of many of the weights are sharply peaked around zero. Furthermore, unlike the support vector classifier, the non-zero weights in the RVM are not associated with examples close to the decision boundary, but rather appear to represent prototypical examples of classes; the relevance vectors. The most compelling feature of the RVM is that, while capable of generalisation performance comparable to an equivalent SVM, it typically utilises dramatically fewer kernel functions. Furthermore, the RVM suffers from none of the other limitations of the SVM described above, as it is a probabilistic model.(Y. Hu & Luob, 2013).

RVM have been used in tandem with GA in order to classify faults in non-linear systems. The relative accuracy of RVM systems supported by GAs provide better classification rates for small learning data sets than for other comparable AI fault classification methods.

RVM has been utilised along with GA in order to optimally control nonlinear manufacturing processes. The technique relies upon discerning (approximately) the optimal control parameters of the manufacturing device. In turn, the non-linear behaviour of the manufacturing device has regression built in to filter out noise through the utilisation of a kernel based Bayesian structure. The GA tabulates the near optimal control parameters in order to maximise the required objective (Yuan, Wang, Yu, & Fang, 2007).

Rotating machinery fault diagnosis has been attempted using thermal imaging processed through RVM methods in combination with bi-dimensional empirical mode decomposition (BEMD) and generalised discriminant analysis (GDA). The BEMD enhanced thermal image is treated with GDA to reduce features, after which RVM is implemented for fault classification (Tran, Yang, Gu, & Ball, 2013).

RVM has been compared to SVM methods to demonstrate its robustness for gear fault detection. Compared to SVM, the RVM method required lesser kernel functions and learning time while demonstrating comparable performance (C. He, Li, Huang, Liu, & Fei, 2009).

RVM combined with GA has been utilised in state classification of roll bearings. The GA is applied to determine training parameters for RVM. Experimentation and analysis revealed that the application of GA in combination with RVM produced better results than back propagation neural networks and SVM (Yanhong Li & Liu, 2010).

A comparison of multi class RVM and SVM methods for low speed bearing fault detection revealed that RVM methods held great promise for accurate fault classification. Component analysis was carried out in order to classify features and to reduce the dimensions of the raw data set. Fault diagnosis was carried out with feature extraction and without it (A. Widodo et al., 2009).

Wavelet packet feature extraction was applied in tandem with RVM for detecting gear faults. Using the Fisher criterion, the discrimination power of the features is tabulated and two optimal features are selected in the time domain and wavelet domain. These are used as inputs to the RVM. Comparisons with SVM revealed that the RVM based method produced better results for online classification (N. Li, Liu, He, Li, & Zha, 2011).

RVM methods have been used on multi class discrimination problems in order to examine sparsity and recognition problems. RVM was used in tandem with multi class and multi kernel methods to test a number of different real world data sets. Results obtained from these methods were compared to results obtained from existing classification techniques. The application of multi kernel RVM methods demonstrated accuracy in producing multi class discrimination problems (Ioannis Psorakis, Theodoros Damoulas, & Mark A. Girolami, 2010a).

RVM methods were applied to analogue circuits for the diagnosis of faults modelled as multi class machine learning problems. Investigation was carried out on a first order Op-amp reluctance capacitance (RC) circuit in order to demonstrate the capabilities of RVM methods in resolving such problems. Results indicated that these methods could be utilized in order to diagnose faults in more intricate analogue circuits that involve a greater number of components (V. Jain, G. N. Pillai, & I. Gupta, 2011).

2.7.4 Principal Component Analysis

PCA is highly effective in reducing the overall dimensions of varied input data sets for more effective analysis. Over time, PCA has been adopted for use in different applications such as fault monitoring and diagnosis, signal processing, recognition of patterns, data compression and other similar tasks (Zhu, Bai, & Yang, 2009). PCA allows projecting large streams of input data onto a smaller dimensional space, so that the projected data is uncorrelated when compared to the original data. The various elements of such projected data are better known as the principal components (Mdlazi, Marwala, Stander, Scheffer, & Heyns, 2007). PCA has been employed with GA in order to reduce data dimensionality for use in fault diagnosis of

induction motors. PCA was employed to remove relative features, following which GA was employed to select the irrelative features and to optimise the ANN (B.-S. Yang, HAN, & YIN, 2006). Fault detection and diagnosis of plant subsystems has also been attempted using PCA. Normal plant operation decomposed through PCA was compared to faulty operation data through PCA decomposition to create thresholds for taking corrective actions. Real time monitoring of plant operation data was compared to both data sets with thresholds settled through Q statistics in order to detect faults (Villegas, Fuente, & Rodríguez, 2010).

Vibration monitoring of helicopter transmissions has been attempted using tri-axial accelerometers and PCA processing of the obtained data. The three different dimensions of acceleration data obtained using accelerometers were reduced to a single dimension using PCA for simpler processing. This approach is seen to provide a simpler and computationally robust technique for vibration monitoring in highly complex systems (Tumer & Huff, 2002). Independent PCA models suffer due to the control limits required for the Q and T^2 statistics. Also, the limits are produced assuming that the process data is Gaussian in character, which may lead to complications if the process data is not actually Gaussian in character. Probabilistic techniques have been used in conjunction with PCA (PPCA) in order to handle both Gaussian and non-Gaussian process data for fault detection and diagnosis in a process control environment. Outcomes signified improvement over simple PCA based control schemes, but certain areas still required improvement under the PPCA based control scheme (B. He, Yang, Chen, & Zhang, 2012).

PCA applications to process control are growing over time. Polyester film process monitoring has been attempted using Q and T^2 statistics through a PCA approach for multivariate quality control (MQC). When compared to other techniques, PCA provided a more robust model for fault detection although diagnosis was not highly reliable. It could be inferred that PCA standalone approaches are best suited to fault detection since fault diagnosis requires the application of other techniques for established reliability (Qin, 2003).

PCA has been used in CM applications to reduce the size of input space provided to the neural network for processing (Mdlazi, et al., 2007). When using the PCA method, the faulty signal is detected by either the T^2 or the Q statistic. The use of either statistic is not enough to provide the root cause of the fault, which must be determined using contribution plots. It can be seen that the contribution of the Q statistic is easier to determine compared to the T^2 statistic. Theoretically, if only a single principal component is recognised, which is responsible for a fault, the T^2 statistic could be used with ease. However, practically this is not possible since practical application requires the use of multiple scores that are connected to a fault (Kano, Hasebe, & Hashimoto, 2000).

The PCA can be used independently in order to classify faults if historical data is available for comparison. In case that there is no historical data for fault comparison and classification, contributions plots must be used to classify faults. The contribution of different variables, or their groups, to a monitored index can be used in order to recognise variables that are causing

the monitored index to reach out of bound values. There is a growing school of thought that tends to favour the use of contribution plots in comparison to PCA alone or PCA combined with other methods for fault identification. This method has been expressed as 'DISSIM' and is known to operate better than conventional PCA methods (Kano, Hasebe, Hashimoto, & Ohno, 2001).

Contribution plots can provide what variable is most strongly related to the observed fault. Typical batch processing operations utilise Q and T^2 statistics in order to identify the presence of faults. In turn, contribution plots are used to decipher what variables have the greatest role to play in the observed fault. It must be noted that this conventional method cannot apply if the identified variable(s) is the cause for or merely connected to the fault. Progressive PCA modelling has been used in batch operations to successively eliminate variables for progressive fault variable identification (Jie, Hong, & Zhang, 2010).

Contribution plots have been applied after PCA in order to detect sensor faults on a localised scale. PCA was used in order to detect faults while contribution plots were created in order to locate the fault on a local scale. The contribution plots were applied through a hierarchical scheme that allowed processing blocks and groups of variables for fault diagnosis (Benaicha, Guerfel, Bouguila, & Benothman, 2010).

2.8   Data Pre-processing

2.8.1 Feature Extraction

Achieving good performance of machine fault diagnosis is largely dependent on appropriate features extraction and features selection techniques. The selection of vital features from the targeted machine is the main contribution to increase the effectiveness of fault diagnosis process. Features extraction techniques can be categorized into three categories; time domain, frequency domain and envelope spectrum (A K Mahamad & Hiyama, 2008).

2.8.1.1   Time Domain Features

The features extracted from raw vibration signals are the statistical measures including root mean square (RMS), peak factor, lower bound, upper bound, entropy, variance, skewness, kurtosis, maximum value and range. Supposing that the signal is $x = x_1, x_2 \dots x_N$, these features are calculated by:

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[x_i - \bar{x}]^2} \qquad\qquad \text{Equation 2.3}$$

where x_i is a element of x; $\bar{x}$ is the mean value which is calculated by $\bar{x} = \frac{1}{N}\sum_{i=1}^{N}x_i$ and N is the number of sample points.

$$\text{Peak factor} = \frac{\max(x)}{\text{RMS}} \qquad\qquad \text{Equation 2.4}$$

$$\text{Lower bound} = \min(X) - \frac{1}{2}\frac{\max(X)-\min(X)}{\text{Nb}} \qquad\qquad \text{Equation 2.5}$$

Upper bound $= \max(x) + \frac{1}{2}\frac{\max(x)-\min(x)}{Nb}$

Equation 2.6

Entropy $= \sum_{i=1}^{N} p_i \log p_i$

Equation 2.7

Where p_i is the probability density of signal x

Variance $= \frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N-1}$

Equation 2.8

Skewness $= \frac{\frac{1}{N}\sum_{i=1}^{N} x_i{}^3}{\sigma^3}$

Equation 2.9

Where:

$\sigma = RMS$

Kurtosis $= \frac{\frac{1}{N}\sum_{i=1}^{N} x_i{}^4}{\sigma^4}$

Equation 2.10

Maximum $= abs(\max(x_i))$

Equation 2.11

Range $= \max(x_i) - \min(x_i)$

Equation 2.3

The interquartile range, normal negative log-likelihood value (Nnl) and Weibull negative log-likelihood value (WNL) have been calculated and used as the input features for PCA (Yadav & Wadhwani, 2011).

The formula interquartile range (IQR) is:

$IQR = Q_3 - Q_1$

Equation 2.13

Where:

$Q_3 = L + \left[\frac{\left(\frac{3}{4}n - CF\right)i}{f}\right]$

$Q_1 = L + \left[\frac{\left(\frac{1}{4}n - CF\right)i}{f}\right]$

L is lower limit of the class containing the quartile

n the number of data points

CF the cumulative frequency up to, but not including, the class containing the quartile

f the frequency in the class containing the quartile the quartile

i the class interval or class width

Weibull negative log-likelihood value and normal log-likelihood value were used recently for features extraction from vibration signals (Yadav & Wadhwani, 2011).

$-LogL = -Log \prod_{i=1}^{n} f\left(a, \frac{b}{x_i}\right) = -\sum_{i=1}^{n} \log f(a, b \backslash x_i)$

Equation 2.14

Where $f(x_i, a, b)$ is the probabilty density function. For Weibull negative log-likelihood function and normal negative log-likelihood function, the $pdfs$ are calculated as follows:

Weibull pdf $f(x_i \backslash a, b) = \frac{b}{a}\left(\frac{x_i}{a}\right)^{b-1} \exp^{-\left(\frac{x_i}{a}\right)^b}$

Equation 2.15

Norm pdf $f(x_i \backslash \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-(x_i - \mu/2\sigma^2)^2}$ Equation 2.16

Where μ and σ denote the mean and standard deviation respectively.

2.8.1.2   Frequency Domain Features

The frequency spectrum of time signals is computed using the Discrete Fourier Transform (DFT). The DFT of a N point time series $x(n)$ is given as:

$X(k) = \sum_{n=1}^{N} x(n)e^{-j2\pi(k-1)(n-1)/N}$ $K = 1,2, \dots \dots \dots N$ Equation 2.17

Normally, to compute the DFT, the Fast Fourier Transform algorithm (FFT) is employed(Vachtsevanos, Lewis, Roemer, Hess, & Wu).

2.8.1.3   Envelope Spectrum Features

Envelope analysis is suitable for diagnostics of machinery where faults have an amplitude-modulating (AM) effect on the characteristic frequencies of the machinery (Feng et al., 2013). Therefore in this situation, the envelope spectrum is required which can be used to localise the defects. Envelop spectrum features are proven to have a capability to reveal the fault sign from machine failure (Ahmed, Gu, & Ball, 2011).

$x = fft(x_{in})$ Equation 2.18

$$x_a(n) = \begin{cases} x(n), & n0, N/2 \\ 2 * X(n), & 1 < n < \frac{N}{2} - 1 \\ 0, & \frac{N}{2} + 1 < n < N - 1 \end{cases}$$ Equation 2.19

$x_a = ifft(x_h)$ Equation 2.20

$x_{env} = \sqrt{x_a * conj(x_a)}$ Equation 2.21

$x_{env} = |ifft(x_{env})|$

Where x_{in} is the vibration signal; X is the FFT of x_{in}; x_a is the FFT of analytic signal for x_{in}, x_a is the analytic signal for x_{in}; x_{env} is the analysed envelope signal and x_{env} is the envelope spectrum.

2.8.1.4   Data Normalization

The data sets were all normalized before training to put data on the same scale before further analysis of both the speed and success of training. In this work the scheme of normalization with zero mean and standard deviation of 1 for each feature set was attempted. The normalized value of e_i for E in the ith column is calculated as (Patki & V.Kelkar, 2013):

$Normalized(e_i) = \frac{e_i - \bar{E}}{std(E)}$ Equation 2.22

Where

$\bar{E} = \frac{1}{n}\sum_{i=1}^{n} e_i$ Equation 2.23

$std(E) = \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{n}(e_i - \bar{\bar{E}})^2}$ Equation 2.24

2.8.1.5   Distance Measures

Also are known as divergence or separability measure. For two classes, a feature X is preferred to another feature Y if induces a greater difference between the two classes, conditional probabilities than Y, if the difference is zero, then X and Y are indistinguishable. One of those methods to measure the distances is Euclidean distance method.

The Euclidean distances between examples of concepts is based on the idea that the greater the distance between the examples from different concepts the better the class separability(Guyon, Gunn, Nikravesh, & Zadeh, 2008). The Euclidean distances function is computed as follows:

$$D(X,Y) = \{\sum_{i=1}^{n}[(X_i - Y_i)^2 U(X_i - Y_i)]\}^{1/2}$$ Equation 2.25

$$U(X_i - Y_i) = \begin{cases} 0 & \text{if } X_i < Y_i \\ 1 & \text{if } X_i \geq Y_i \end{cases}$$ Equation 2.26

Where $D(X,Y)$ is the distance between two features X and Y size of n, and $U(X_i - Y_i)$ is the unit step function as defined by Eq 2.26.

2.8.2 Features Selection

The task of fault diagnosis of modern machinery has become difficult due to the growing complexity of machine design. But with the development of advance sensor technology and signal processing techniques, many features can be extracted from targeted machines for fault diagnosis purposes.

Normally, fault diagnosis has characteristic uncertainties. The uncertainties information can be reduced by using multiple sources of information. This means that the fault diagnosis capability can be improved by feeding inputs with multiple features. The problem may exist with the increase of features that can increase the difficulty of data analysis. It is unnecessary to employ all the features for fault diagnosis purposes. Some features can contribute significant information of faulty signs while some only contribute less information. Thus, it is necessary to have appropriate feature selection to increase the accuracy of the fault diagnosis process.

Various methods can be used for feature selection. Basically the features selection is done by evaluating the features based on other evaluation measurement such as distance, information, dependency and classification error rate obtained during measurement. Various methods can be used for feature selection such as modified distance discriminant technique (Xu, Xuan, Shi, Wu, & Hu, 2009b), distance evaluation technique (Q. Hu, He, Zhang, & Zi, 2007; Lei, He, Zi, & Hu, 2007; Xu, Xuan, Shi, Wu, & Hu, 2009a) , neural network (Matsuura, 2004). The selection of features selection is essential to increase the accuracy of fault diagnosis system.

2.9   Summary

This chapter introduced how the importance is the RCs with its operating method and the functions of parts of the compressor into which faults will be seeded, the valves, the

intercooler and the belt drive, and the consequences of such faults in an industrial context. Also the chapter has introduced to the reader the aspects of CM and various CM techniques.

Chapter 3

Test Rig and Experimental Arrangements

This chapter introduces the test rig with its faults seeding, instrumentation and system procedures. It begins by giving a brief description of the compressor characteristics and specifications; then the fault seeding. Thereafter, it discusses the operation of the transducers and its measured parameters, as well as a description of the data acquisition system. Finally, the calibrations of the pressure and vibration sensors are explained.

3.1   Test Rig Description

The Broom Wade TS9 RC was selected because such machines are widely used in industry and this particular compressor has previously been used by researchers in the Diagnostic Research Group and its elements are well known. The chosen compressor is a two-stage, single acting, V-formed with two cylinders designed to deliver compressed air between 0.55MPa and 0.8MPa to a horizontal air receiver tank with a maximum working pressure of about 1.38MPa. A diaphragm pressure switch on the tank could be set to switch off the electrical current to the motor or trigger data collection when the air pressure in the storage tank reached a prescribed value. As shown in Figure 3.1, the driving motor used was a three phase, squirrel cage, air cooled, type KX-C184 with a 2.5kW induction motor. It was mounted on the top of the receiver and transfers its power to the compressor through a pulley drive belt system. The transmission ratio was 3.2, which results in a crank shaft speed of 440rpm when the motor runs at its rated speed of 1420 rpm. An intercooler is situated between the cylinders.

Figure 3.1 Broom-Wade TS9 RC

Figure 3.1 also shows the compressor attached to a storage vessel(Wade, 1964). Basic technical specifications of the compressor are listed in Table 3.1 :

Compressor	Broom Wade TS9
Max working pressure	1.38 MPa
Number of cylinders	2 (90° opposed)
Piston stroke	76 mm
Speed	440 rpm
Motor power	2.5/3 HP
Voltage	380/420 V
Motor Speed	1420 rpm
Current	4.1/4.8 A
Piston Diameter [Low Pressure Cylinder]	93.6 mm
Piston Diameter [High Pressure Cylinder]	55.6 mm

Table 3.1 Compressor specifications

Intake air is filtered to ensure it is sufficiently clean and dry before entry into the compressor. The compressor initially compresses the air in the low-pressure cylinder and subsequently into the intercooler before the high-pressure cylinder, which has a smaller bore. The compressed air is ejected from the second cylinder into the storage cylinder which has a safety valve fitted to prevent excessive pressure build-up (M Elhaj, 2005). Splash lubrication is used to protect moving parts. The compressor drives a vane flywheel, the airflow from which passes over the coiled copper piping forming the intercooler, removing heat from it. Additional cooling of the intercooler is provided by the airflow in and out of a crankcase breather as the pistons move.

An isolator and a direct online starter connect the compressor to a three-phase current supply. A safety valve is fitted on the compressor to protect the motor from overload.

3.2   Fault Seeding

Seven common faults were separately introduced into a RC. These are a case of a leaky discharge valve in the high pressure cylinder, a leak in the intercooler and a loose drive belt. The experimental tests were carried out in the following sequence:

1. A healthy compressor with both stages operating normally.
2. A leaky discharge valve in the high pressure cylinder (second stage).
3. A leaky suction valve in the high pressure cylinder (second stage).
4. A leak in the intercooler.
5. A loose drive belt.
6. Discharge valve leakage combined with suction valve leakage.
7. Suction valve leakage combined with intercooler leakage.
8. Discharge valve leakage combined with intercooler leakage.

In order to determine the state of health of the compressor a reference or base-line signal (signature) is required. This reference was obtained when the healthy compressor was operating normally after a thorough inspection by qualified staff. The compressor was then

operated with each fault in turn, the signal from each transducer compared to that detected for normal operation, and deviations of the measured signal from the base line noted.

3.2.1 Valve Leakage Simulation

The valve leakage was introduced by drilling a small hole in the valve plate of the second stage discharge valve as shown in Figure 3.2. The hole was 2mm diameter and it is 2% of the flow cross-sectional area.
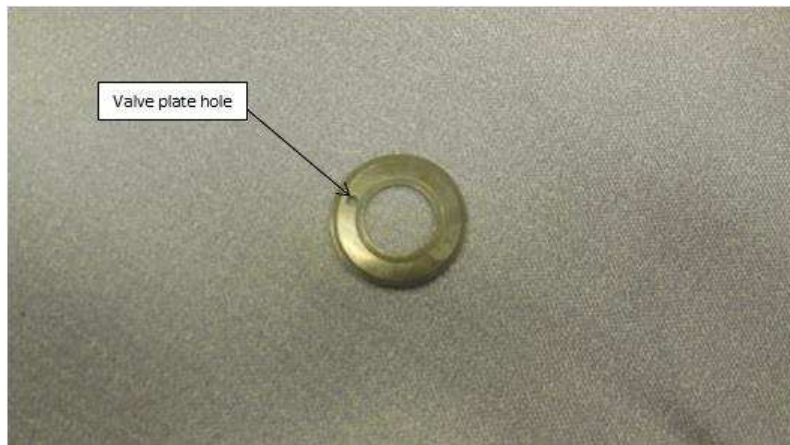
Figure 3.2 Leak in 2nd stage valve plate

3.2.2 Leak in Intercooler

Leakages are common in joints in the pipe work carrying the process gas from the first stage to the second. Here a loose intercooler joint is seeded into a compression joint close to the second cylinder. The pipeline screw nut, shown on the Figure 3.3, was loosened to create the leak. A small leakage was achieved by turning the nut through one turn. Unfortunately, whist this represented a realistic leakage; it was not possible to quantify the leak as a proportion of the area of the flow cross-section.

Figure 3.3 Leak in the intercooler

3.2.3 Loose Drive Belt

To model a loose belt arising from belt wear due to friction the separation of the centres of the two pulleys was reduced from 169mm to 167mm, equivalent to a 0.5% increase in belt length.
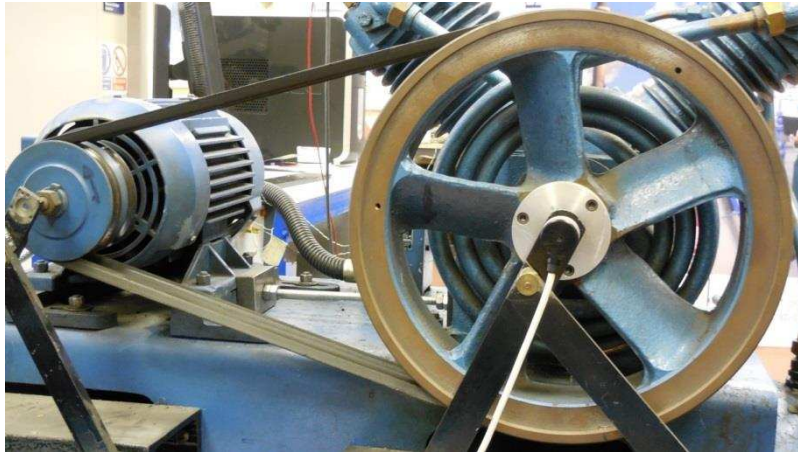
Figure 3.4 A belt fault

3.3   Measurement Transducers

The test rig has been used for several previous CM research projects and so a number of different transducers were already fitted. This includes accelerometers, instantaneous angular speed (IAS) encoder, static and dynamic pressure sensors and thermocouples. Each of these transducers is connected to a data acquisition system (DAQ) by coaxial BNC cables to reduce signal noise.

3.3.1 Accelerometers

Two accelerometers, type YD-5-2 with frequency range 0-15kHz, sensitivity $45mv/ms^{-2}$, capable of withstanding temperatures of up to 150°C and acceleration of up to $2000ms^{-2}$, were positioned at the inlet and outlet valves of the compressor to measure vibration (Kjaer, 1997). Figure 3.5 shows how each accelerometer is connected via a screw-threaded brass stud bonded to the casing with ceramic cement, which helps to avoid over-heating.

Figure 3.5 Location of an accelerometer at top of cylinder head

3.3.2 Dynamic Pressure Sensor

For dynamic pressure measurements, an analogue pressure transducer was placed on the head of each cylinder. The pressure sensors were fitted via a small hole drilled into the head of each cylinder, as shown in Figure 3.6. These sensors were GEMS type 2200 strain gauge pressure transducers with an output of 100mV for full range pressure when used with a 10Vd.c. power supply. These sensors have a range of up to about 4MPa (600psi) and an upper frequency limit of about 4 kHz.

Figure 3.6 Dynamic pressure sensor on cylinder head

3.3.3 Static Pressure Sensor

A Gem type PS20000 static pressure sensor is installed on the air storage tank, see Figure 3.7. Its operating range is from 0 to 1.35MPa (200Psi), with a maximum output 100mV when supply voltage is 15V. Operating temperature range is −20°C to +105°C.

Figure 3.7 Static pressure sensor on air storage tank

This sensor is used to monitor the tank pressure and can be used to trigger data collection automatically when the pressure reaches the prescribed set points. Knowing the pressure delivered to the storage tank allows the efficiency of the compressor to be calculated.

3.3.4 Temperature Measurement

Figure 3.8 shows the K type thermocouples with a linear response at temperatures between -20ºC to 220ºC. Thermocouples were inserted alongside the pressure sensor to measure the temperature of the air in each compressor cylinder. This ensured the safe operation not only of the pressure sensors, but also of the compressor itself.
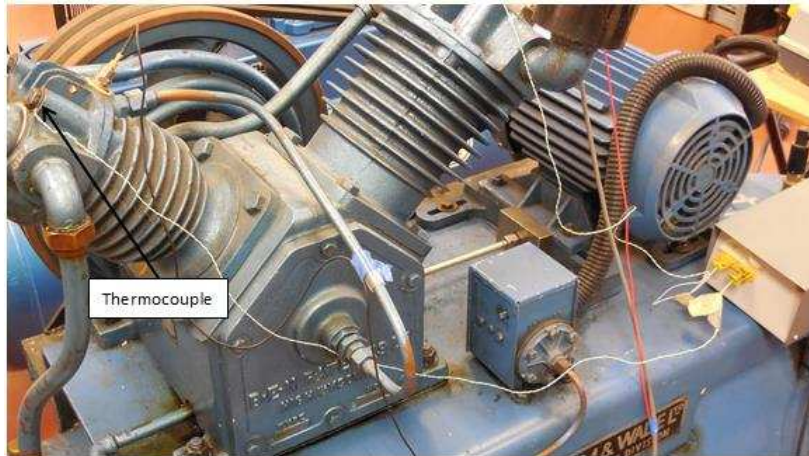
Figure 3.8 Thermocouple

3.3.5 Shaft Encoder

Figure 3.9 shows the Hengstler incremental optical encoder which was attached to the drive shaft using a spindle adapter and used to measure the IAS of the drive shaft accurate to 1º. This allowed small changes in shaft speed to be measured and recorded. The encoder may be connected directly to the PC via the DAQ.

Figure 3.9 Hengstler encoder sensor

3.3.6 Current Transducer

The Hall Effect current transducer, RS number 286-327, used to measure the stator current is mounted on a printed circuit board (PCB). This transducer, see Figure 3.10, is designed for measuring circuit current without actually connecting into the circuit. Technical specifications are given in Table 3.2.
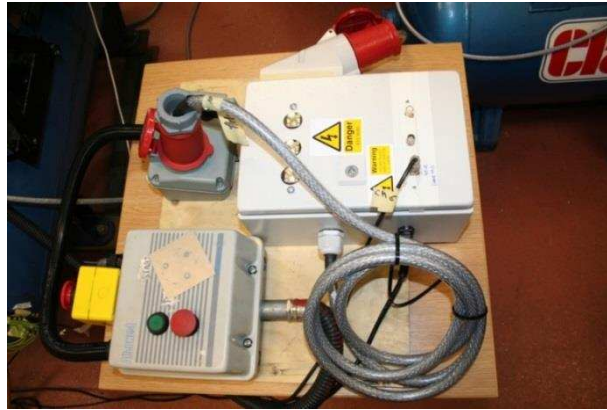
Figure 3.10 Three-phase current measuring unit

Hall Effect current transducer	(RS 286-327)
Response time (Inst.)	<1 μ sec.
Supply voltage	±15V dc, (±5%)
Operating temperature	0 °C to +70°C
Bandwidth	DC to 100kHz
Analogue output voltage	5V

Table 3.2 Technical specifications for Hall Effect current transducer (RS 286-327)

3.3.7 Microphone

Airborne sound was measured using Bruel & Kjaer precision microphone type 4130 as shown in Figure 3.11. This has a linear response up to about 16 kHz which is well above the upper frequency limit of the sound of interest (about 10kHz) and a dynamic range to 130dB which is well above the maximum noise level expected. The microphone was 10cm from the cylinder band hence it was in the acoustic near field for all frequencies below about 1 kHz.

Figure 3.11 High-bandwidth microphone

Parameter	Value
Model	MK224
Type	4130
Sensitivity	41.7mV/Pa
Frequency response	16kHz
Maximum range	130dB

Table 3.3 Specification for Bruel & Kjaer precision microphone

3.4   Data Acquisition System

3.4.1 Hardware: 1401 CED

In the present work, the ADC used is the Power 1401 CED as shown in Figure 3.12 and Figure 3.13, from Cambridge Electronics Ltd. Analogue to digital converter is an electronic device, often an integrated circuit that converts an analogue voltage to a digital value. Two main parameters of interest in ADCs are the rate at which the converter can sample analogue values, and the resolution at which it can resolve the values. Sampling rate is given in samples and the resolution is given as a percentage of the maximum voltage that the converter can resolve, or the number of bits that this corresponds to. It can record waveform data, event (digital) data and marker information at 400 kHz with 16bit resolution. It can also simultaneously generate waveforms and digital outputs for real-time, multi-tasking experimental systems(Ltd, 1991).We only use seven channels of sixteen channels. Channel '1' collects the data from low pressure transducer. Channel '2' collects the data from the high pressure transducer. Channel '3' collects the data from the low vibration accelerometer. Channel '4' collects the data from high vibration accelerometers. Channel '5' collects the data from shaft encoder to measure the angle index signal. Channel '6' collects the data to measure angular position mark for position TDC piston and channel '7' collects the data from current transducer.

Figure 3.12 The front view of the Power 1401 CED

Figure 3.13 The rear view of the Power 1401

3.4.2 Software: Lab Windows TM/CVI Version 5.5

The DAQ software was written in the programming language C and developed using National Instruments Lab Windows TM/CVI Version 5.5, which includes a large set of run-time libraries for instrument control, data acquisition and analysis (National Instrument Company, 2000). A graphical user interface (GUI) editor is included. The Lab Windows environment includes such features as automatic code generation which makes measurement much easier than traditional C or C++ environments.

The DAQ software enables multiple channels of dynamic data (e.g. IAS, vibration, dynamic pressure, motor current, sound and temperature) to be acquired simultaneously at different rates and data lengths. The software package has an acquisition set-up panel with the control commands and status indicators listed on the screen. This allows the user to modify, for example, the sampling frequency and sample data length and ensure an optimal data set is collected for subsequent off-line analysis.

Figure 3.14 shows the set-up screen: sampling frequency is 62.5 kHz; the data length is 30,642 samples. Thus the time between data consecutive points was 0.4903 sec.

Figure 3.15 shows the panel displaying the data collected for compressor monitoring. Each channel shows a different parameter: dynamic pressures from the low and high pressure cylinders, the acceleration levels on the low and high pressure cylinders, the current signal and the encoder signal. A trigger signal initiates data collection every time the piston passes TDC, so that each pulse is always measured at exactly the same crank position. This ensures accurate time-domain averaging. Running averages of six consecutive data segments were used to obtain a cleaner signal.
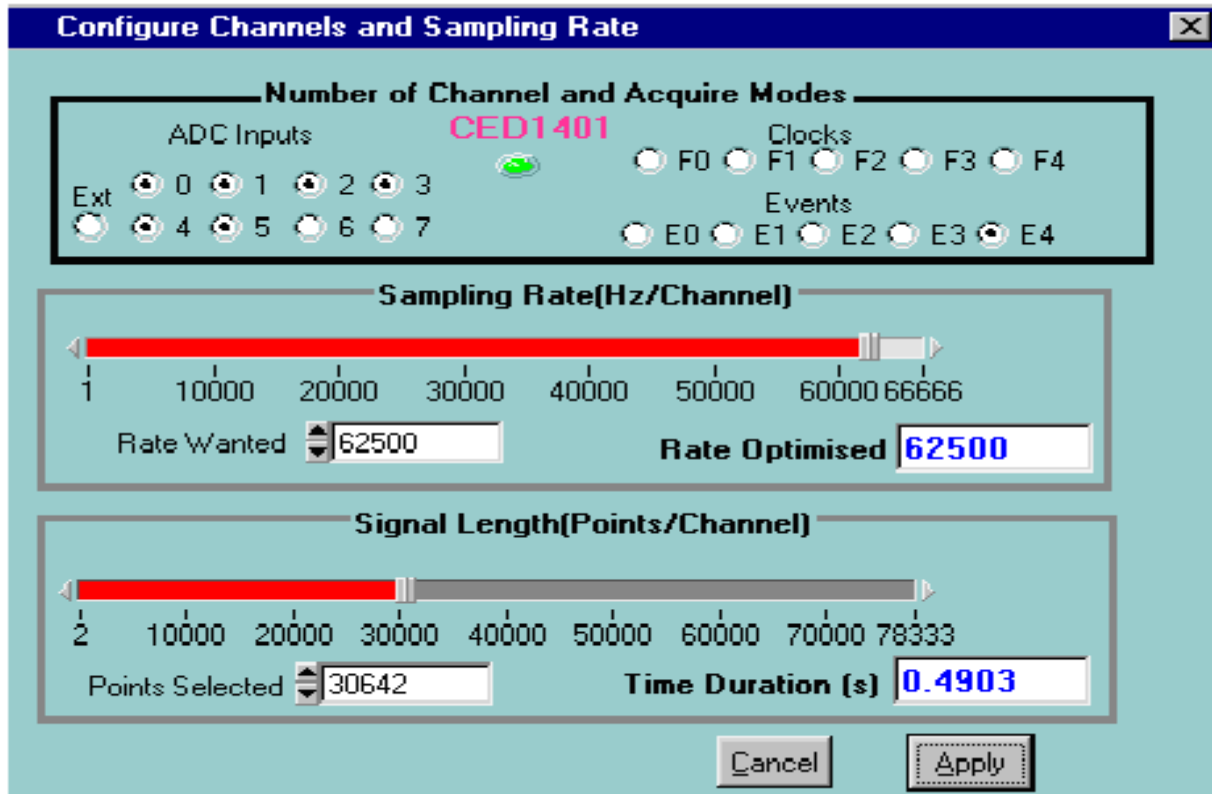
Figure 3.14 Setup screen for data acquisition
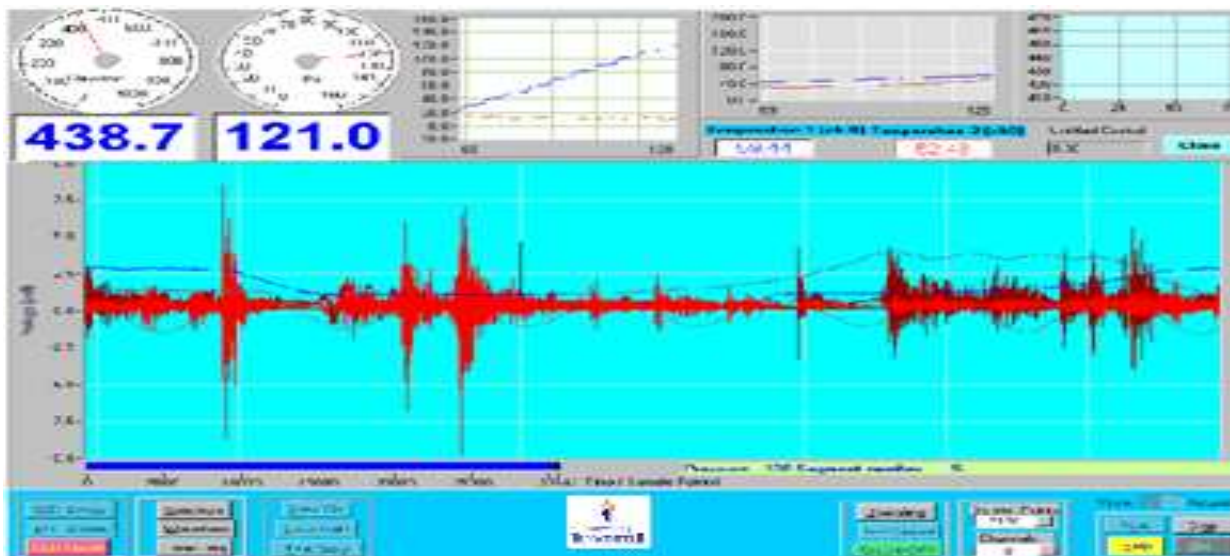
Figure 3.15 Data acquisition in progress

3.4.3 Data Management and Measurement Practice

Figure 3.16 is a schematic diagram of the test system and shows the sensor positions. To help ensure good practice, a standard test procedure was developed. For each set of data, 4 files were collected. The same transducers were always connected to the same input ports on the 1401PLus. Care was taken to ensure that information regarding data collection was always placed in the directory created for that purpose. The specific test, compressor load and fault

conditions, and related files were identified by a filename structure established for that purpose.
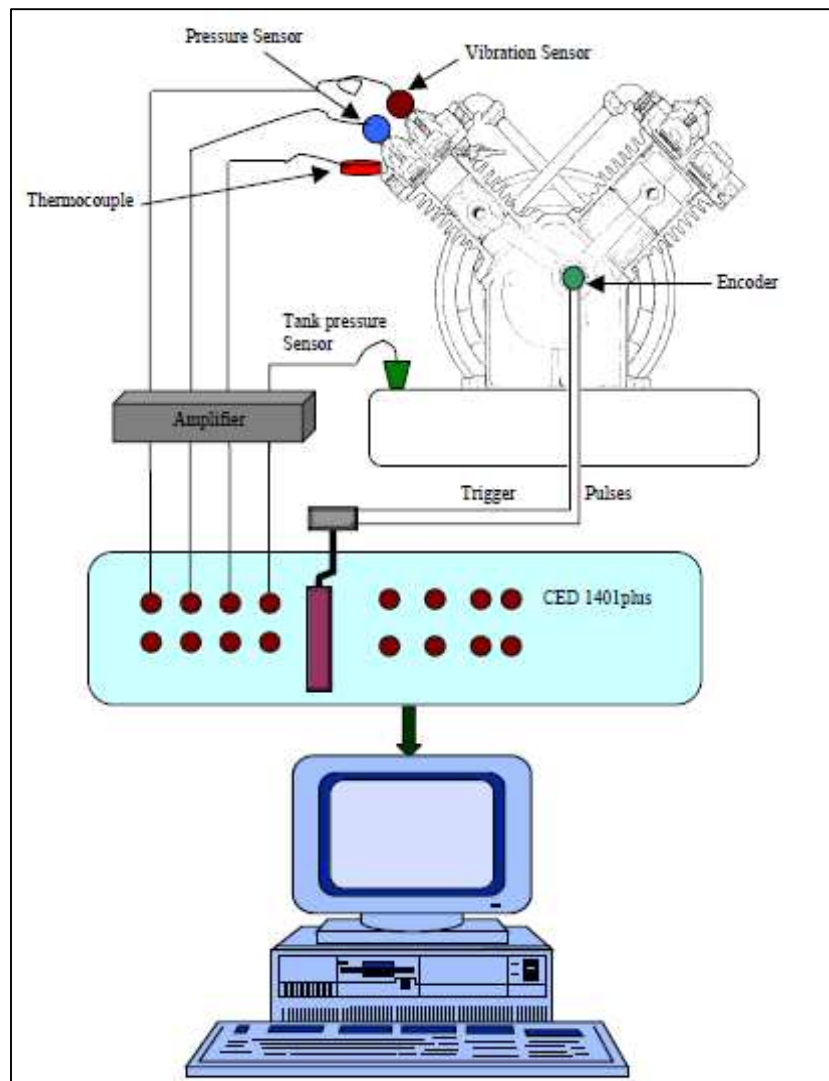
Figure 3.16 Schematic diagram of compressor test system

3.5   Calibration of Instruments

3.5.1 Pressure Transducer Calibration

The various sensors and instruments were calibrated through the data acquisition system before the test was conducted. To achieve this, readings were taken by the data acquisition as the input signals were swept through their operating ranges, in order to calculate the gain and offset values for each individual signal. The GEMS-2200 in-cylinder pressure transducer was calibrated by applying a known pressure to the sensor and measuring the voltage from the charge amplifier. The dead weight system used had an accuracy of 0.1%. These figures were then used to calculate the gain and offset values for the transducer.

3.5.2 Vibration Transducer Calibration

Generally the manufacturers follow the International code ISO/FDIS16063-21 (ISO/16063-21, 2003) which gives the guidelines for the calibration of vibration by comparison to a reference transducer. The calibration procedure adopted uses a 1g sinusoidal vibration generator (Electro-dynamic shaker (JZK-51-SHAKER)) with varying frequency. A well-calibrated comparison accelerometer is used to characterise the measured responses of the accelerometers to be calibrated.

The Electro-dynamic shaker, model JZK-51-SHAKER, has a wide frequency range from 10 Hz - 2 kHz. The accelerometers were calibrated using the data acquisition system before the tests started. To achieve this, readings were taken by the data acquisition as the input signals were swept through their operating ranges, in order to calculate the gain and offset values for each individual signal. The difference between accelerometers readings value were found. These differences are due to their sensitivity difference (11.070 and 11.799 mV/ms-2 for vibration first stage and the second stage respectively). The accuracy of the both sensors were determined and included in the features extraction algorithms.

3.6 Summary

This chapter provides the reader with a detailed description of the test facility. Facilities and relevant instrumentation that would be used to study the vibration, current, pressure, temperature are included. All the data during experiments were collected using the system discussed later and then analysed the experimental data in the PC on MATLAB code In general, the test facilities and instrumentation have been tested and proved to be sufficient for common RC condition monitoring and fault detection.

Chapter 4

Vibration Characterisation Based on Mathematical Models and Numerical Studies of A reciprocating Compressor

In this chapter a mathematical model is developed for the numerical simulation of the behaviour of the two-stage RC used in the test rig and subsequently used for characterising fault signatures. The model consists of a crank shaft, fluid flow and valve impact dynamics to detail the valve motion, the intercooler flow, in-cylinder pressure variation and belt transmission behaviour. Moreover, it has developed fault models to understand the changes of signatures under intercooler leakages, abnormal valve events and belt wears. The studies allow gaining in-depth understanding of the vibration responses and pave basis for develop effective methods for signal and data analysis in subsequent works.

4.1   Introduction

A number of techniques have been developed in last two decades for CM of RCs. As reviewed in the work (M Elhaj, 2005) and (Mahmud, Ann, Feng, & Andrew, 2014) , most of these techniques are developed based on data-driven approaches. They are good at developing detection systems as the method can be trained with small data samples without any knowledge of the physics of systems. However, they generally have the problems with less accuracy in diagnosis and poor generalisation capability.   .

To overcome these shortages, this study will takes into account the dynamics of the system with maximum degrees in developing the data driven based AI approaches. To this end, this chapter with examine the dynamics of the RC by a numerical simulation. It consists of developing the mathematic models of a two-stage compressor, simulating the dynamic responses in MATLAB under different fault cases and then characterising the general relationships between vibration sources and various fault cases. In addition, the model reliability, simulation implementation and key results have been also verified by experiments under representative operating conditions.

4.2   Implementation

The core part in converting the model into MATLAB codes is to decide on an appropriate equation solver from among a number of MATLAB options. The fourth-order Runge–Kutta integration algorithm (ode45) was selected for this simulation. This is the algorithm recommended by MATLAB and has been proved to be stable and efficient in solving non-stiffness differential equations. However, there are some degrees of stiffness in the valve

equations because of the seat impact transients, that is a single step size for the computations would lead either to inaccurate solutions (too large a step size for the transient period) or take far too long to compute (too small a step size for the more linear stage). To take into account the influence of possible stiffness in the solving process, a small step-size was selected so that the solutions of the valve motions were stable, then for comparison, other solvers such as ode23s and ode15s were tested. The latter is recommended by MATLAB for solving stiff equations, and no significant differences were found between either of them and the results predicted by ode45. The numerical solution process is continued up to a time equivalent to the duration of 5 compressor cycles, when all of the solutions have reached steady operation. Each process begins at the TDC of the piston in the 1st cylinder (crank angle $\theta=0^o$), when both suction and discharge valves are closed and the cylinder pressure is relatively low, and finishes after 5 revolutions of the crankshaft, again at TDC with $\theta = 5 \times 360^o = 1800o$. The logical flow chart of the simulation model is shown in Figure 4.1 below.
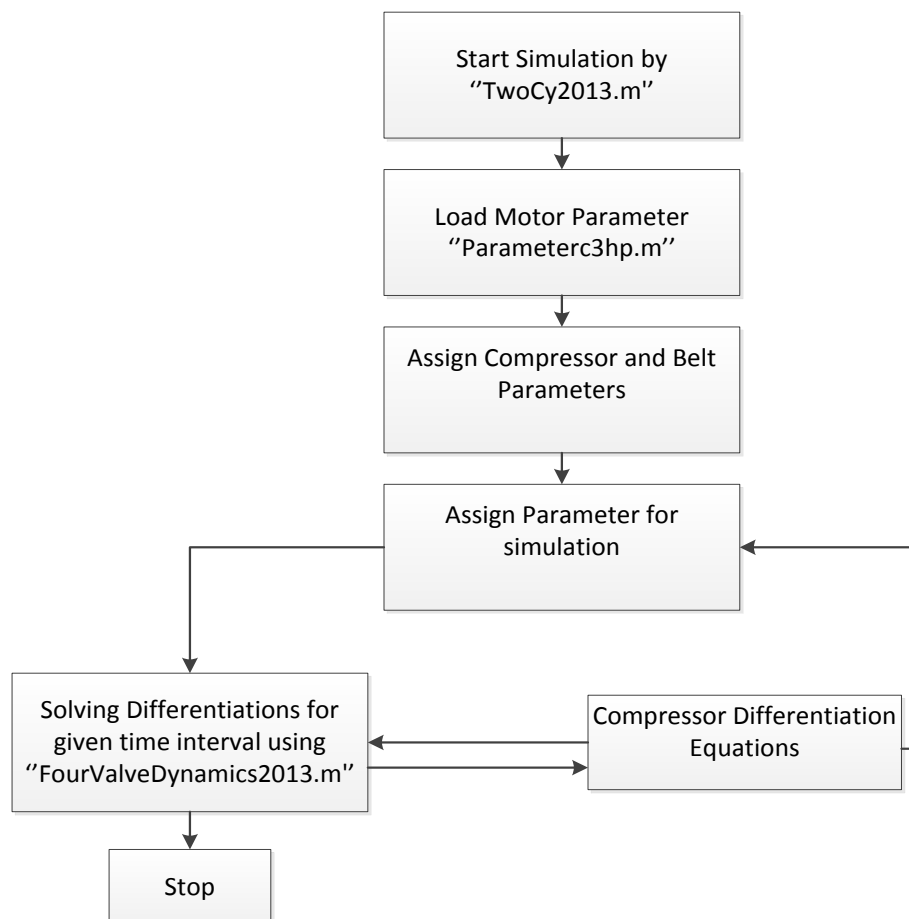
Figure 4.1 Flowchart of compressor simulation model

4.3   Dynamics of the Piston Mechanism

4.3.1 A typical Piston Cylinder System

Figure 4.2 and Figure 4.3 show the typical piston cylinder system of RC, the piston with length of con-rod l crank and radius r. The operation of cycle begin at TDC where $\theta = 0°$ and ending with $\theta = 360°$ each revolution of crankshaft, which is assumed to rotate at constant angular speed. The crankshaft rotates driven by a prime mover and causes the con-rod to move the piston in the cylinder, compressing the gas which exerts a force F directed against the motion of the piston (Ball, 2000).
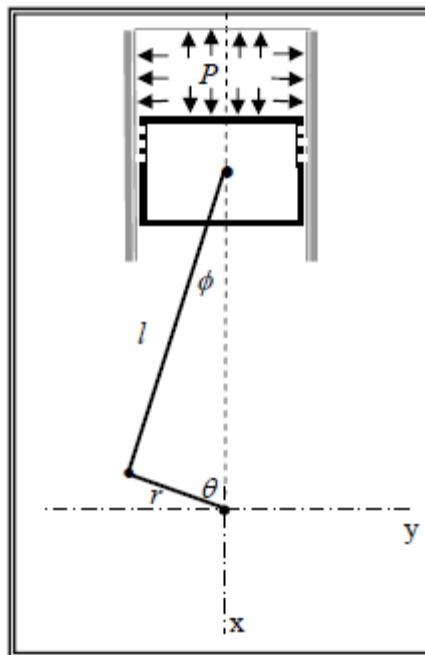
Figure 4.2 Forces due to gas pressure (M Elhaj, 2005)

Figure 4.3 depicts the reaction to the force F which can be resolved into components; $\text{F}/\cos\emptyset$ acting along the con-rod and $\text{Ftan}\emptyset$ acting in the ZY axes. The force $\text{F}/\cos\emptyset$ induces a torque M_t which acts in the plane to rotate the crankshaft in a counter clockwise direction Z.

$$M_t = \left[\frac{F}{cos\emptyset}\right] rsin(\theta + \emptyset) \qquad\qquad \text{Equation 4.1}$$
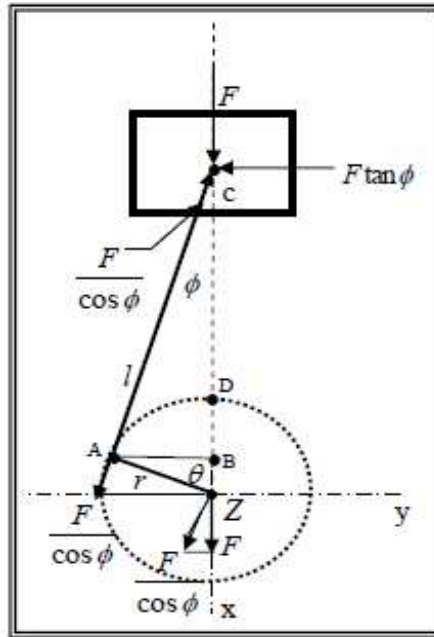
Figure 4.3 Action of the gas pressure forces (M Elhaj, 2005)

Since there is no translational movement of the crankshaft, the bearing of the crankshaft must exert forces F in the ZX vertical direction and $Ftan\emptyset$ in the ZY horizontal direction. The forces transmitted to the stationary parts of the compressor are:

- F upwards on the cylinder head
- $Ftan\emptyset$ to the right on the piston guide
- F downwards on the crankshaft bearings at Z
- $Ftan\emptyset$ to the left on the crankshaft bearings at Z

Figure 4.4 presents the simple piston and the crank arrangement, where x_p is the downward displacement of the piston from the TDC and θ is the crank angle from the TDC.

If the crankshaft is rotating at constant angular speed ω, then the position of piston can be calculated in terms of the crank angle θ.

$$x_p = r\left(1 - rcos(\theta)\right) + lcos(\theta) \qquad \text{Equation 4.2}$$

However, considering the geometry of the crankshaft and con-rod;

$$lsin(\emptyset) = rsin\,\theta = rsin\,\theta \qquad \text{Equation 4.3}$$

Consequently

$$cos\emptyset = \sqrt{1 - \frac{r^2}{l^2}sin^2\theta} \qquad \text{Equation 4.4}$$

By substituting Eq.(4.4) into Eq.(4.2):

$$x_p = r\left(1 - rcos(\theta)\right) + l\left[1 - \sqrt{\frac{r^2}{l^2}sin^2(\theta)}\right] \qquad \text{Equation 4.5}$$

The Eq. (4.5) can be differentiated to obtain expressions for the velocity and the acceleration of the piston:

$$\ddot{x}_p = \omega\, r\, \sin\theta\,(1 + \cos\theta)\Big/\sqrt{1 - \frac{r^2}{l^2}\sin^2\theta} \qquad\qquad \text{Equation 4.6}$$

If the assumption is made that $\left(\frac{r^2}{l^2}\right)\sin^2\theta \ll 1$ and this is reasonable since $\frac{r}{l} = \frac{1}{4}$ and $\sin^2\theta < 1$, (unless $\theta = (2n+1)$ then the expression for the acceleration may be written as:

$$\ddot{x}_p = \omega^2 r \left(\cos\theta + \frac{r}{l}\cos 2\theta\right) \qquad\qquad \text{Equation 4.7}$$
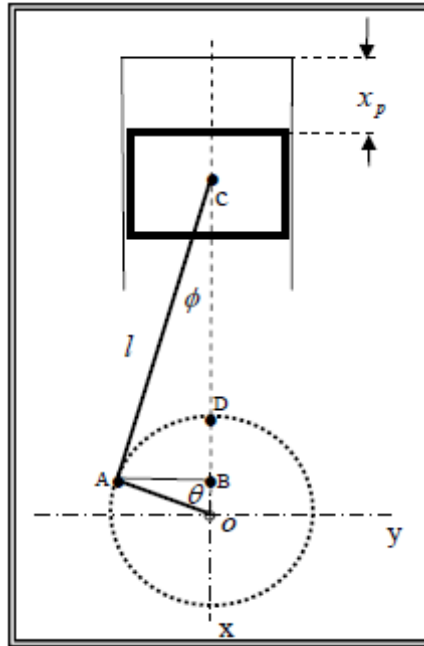
Figure 4.4 Diagrammatic representation of piston displacement (M Elhaj, 2005)

4.3.2 Two Stage Piston Crank Arrangement

The position of the piston in each cylinder of the two stage compressor is considered positive when it moves downward, as shown in Figure 4.5. For a Broom Wade TS-9 compressor, the displacement of the piston in the 1st stage lags the 2nd stage by a phase of п/2. If the initial, lower pressure stage is taken as the reference stage, the displacements of the pistons in the two stages are:

$$x_{pL} = r(1 - r\cos(\theta)) + l\left[1 - \sqrt{1 - \frac{r^2}{l^2}\sin^2(\theta)}\right] \qquad\qquad \text{Equation 4.8}$$

$$x_{pH} = r\left(1 - r\cos\left(\theta + \frac{\pi}{2}\right)\right) + l\left[1 - \sqrt{1 - \frac{r^2}{l^2}\sin^2(\theta)}\right] \qquad\qquad \text{Equation 4.9}$$
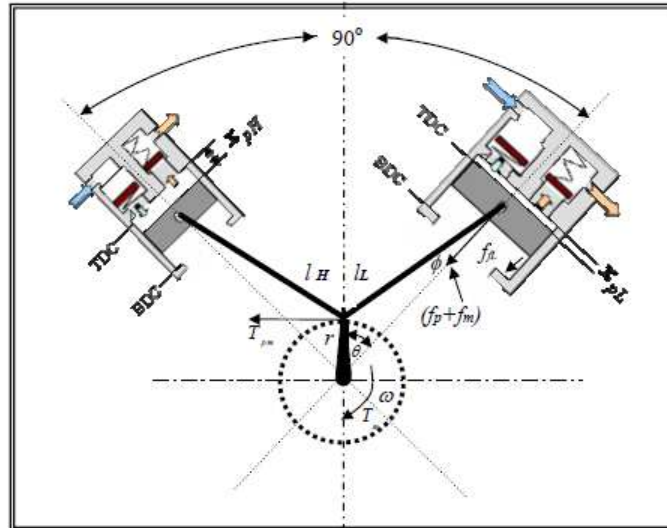
Figure 4.5 Simplified model of two stage air compressor (M Elhaj, 2005)

The velocity $\dot{x}_{pL,H}$ and acceleration $\ddot{x}_{pL,H}$ of the pistons of the low pressure and high pressure cylinders follow by differentiation:

$$\dot{x}_{pL} = \omega\, r\, \sin\theta(1 + \tfrac{r}{l}\cos\theta)/\sqrt{1 - \tfrac{r^2}{l^2}\sin^2\theta} \qquad \text{Equation 4.10}$$

$$\dot{x}_{pH} = \omega\, r\, \sin(\theta + \tfrac{\pi}{2})(1 + \tfrac{r}{l}\cos(\theta + \tfrac{\pi}{2}))/\sqrt{1 - \tfrac{r^2}{l^2}\sin^2(\theta + \tfrac{\pi}{2})} \qquad \text{Equation 4.11}$$

$$\ddot{x}_{pL} = \omega^2 r(\cos\theta + \tfrac{r}{l}\cos 2\theta) \qquad \text{Equation 4.12}$$

$$\ddot{x}_{pH} = \omega^2 r(\cos(\theta + \tfrac{r}{l}) + \tfrac{r}{l}\cos 2(\theta + \tfrac{\pi}{2})) \qquad \text{Equation 4.13}$$

4.3.3 Crank Angle Model

According to the simplified dynamic model, the equation of crank motion can be derived according to Newton's second law:

$$J\frac{d^2\omega r}{dt^2} = T_{em}(t) - T_{pmL,H}(t) - T_{fL,H}(t) \qquad \text{Equation 4.14}$$

Where crankshaft angular speed ω is a function of time t, j is the equivalent inertial moment of the system (consisting of the reciprocating and rotating parts of the compressor, an electric motor and power transmission shaft or a belt connected the motor and compressor).

$T_{em}(t)$ is the driving torque from the electric motor, $T_{pmL,H}(t)$, is the resultant torque to the air pressure inside the cylinder and the unbalanced inertial force and the con-rod of the 1st and 2nd stages and $T_{fL,H}(t)$ is the friction torque of the low pressure and high pressure cylinders.

4.3.4 Calculation of Torques

This section indicates the mathematical expressions for the driving load torque of the power unit, the torque induced by the gas pressure in both cylinders, and the vertical unbalanced inertial force in the two-stage compressor. Yang et al (J. Yang, Pu, Wang, Zhou, & Yan, 2001) show the driving torque from the motor to be:

$$T_{em} = \frac{P_w}{\omega_s} B_r$$ Equation 4.15

Where

P_w is the motor power in Watts,

B_r = is the belt transmission ratio

ω_s = is the motor speed in rad/sec

Based on the eq $M_t = \left(\frac{F}{\cos\emptyset}\right).r\sin(\theta + \emptyset)$, it can be shown that the resultant torque due to the air pressure inside the cylinder and the reciprocating inertial force of the piston and the con-rod of the low pressure and high pressure cylinders is:

$$T_{pmL} = (f_{pL} + f_{mL})R_{eL}$$ Equation 4.16

$$T_{pmH} = (f_{pH} + f_{mH})R_{eH}$$ Equation 4.17

Where:

$$R_{eL} = r\sin\theta + \frac{r}{2l}\sin 2\theta / \sqrt{1 - \frac{r^2}{l^2}\sin^2(\theta)}$$ Equation 4.18

$$R_{eH} = r\sin\left(\theta + \frac{\pi}{2}\right) + \frac{r}{2l}\sin^2\left(\theta + \frac{\pi}{2}\right) / \sqrt{1 - \frac{r^2}{l^2}\sin^2\left(\theta + \frac{\pi}{2}\right)}$$ Equation 4.19

Where:

f_{pL} and f_{pH} are the forces produced by the gas pressure in the lower and higher pressure cylinder respectively, f_{mL} and f_{mH} are the inertial forces produced by the reciprocating mass of the low and high pressure cylinder respectively, r is the crank radius, R_{eL} and R_{eH} are the effective radius of the crankshaft for the low and high pressure cylinders respectively, and l is the length of the con-rod length (J. Yang, et al., 2001).

The force produced by the air pressure in both cylinders becomes:

$$f_{mL} = m_{recL}\ddot{X}_{pL}$$ Equation 4.20

$$f_{mH} = m_{recH}\ddot{X}_{pH}$$ Equation 4.21

Where:

$\ddot{X}_{pL}$ and $\ddot{X}_{pH}$ are vertical piston acceleration [m.sec^{-2}]

The RC inertial mass of both stages m_{recL} and m_{recH} are calculated from the following equations:

$$m_{recL} = m_{pL} + 0.5m_{crL}$$ Equation 4.22

$$m_{recH} = m_{pH} + 0.5m_{crH}$$ Equation 4.23

Where m_{crL} and m_{crH} are the reciprocating inertial mass of both stages, $m_{pL.H}$ is the piston mass of both cylinders and $m_{crL,H}$ is the con-rod mass.

4.4   Cylinder Pressure Model

The airflow thorough a two-stage RC is shown schematically in Figure 4.6. The airflow enters the intake and it is then cleaned by the air filter and passes through the 1st stage suction port into the 1st cylinder. After being compressed, the air passes through the 1st stage discharge chamber, to the intercooler, then through the second stage suction port, and into the 2nd stage

high pressure cylinder. The air is then compressed and passes out through the 2nd stage discharge chamber, and then into the storage tank.
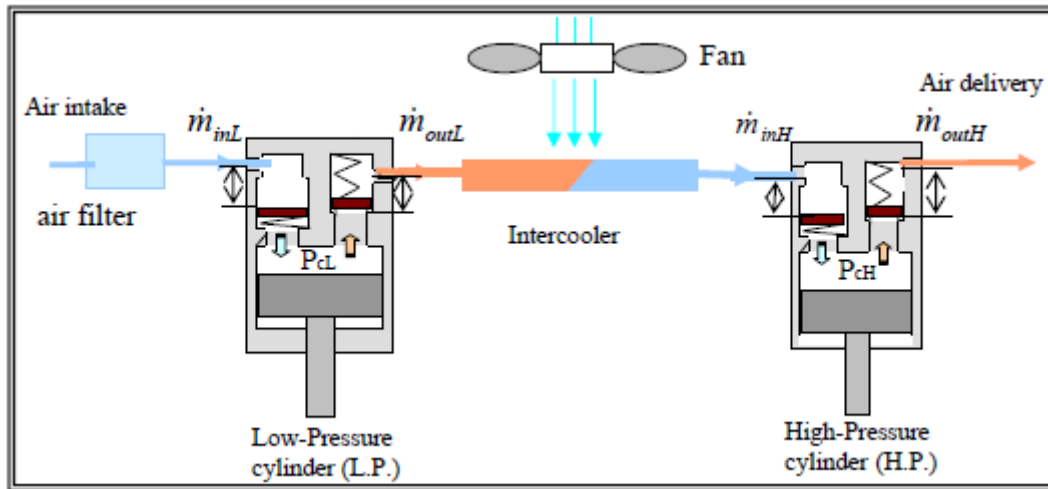
Figure 4.6 Air flow through a two-stage RC

4.4.1 Cylinder Pressure

The equation for the instantaneous cylinder pressure in each cylinder can be derived from the first law of thermodynamics (Stronach, Johnston, & Cudworth, 1984) as:

$$\frac{dP_{CL}}{dt} = \frac{1}{V_{CL}}\left[C_{iL}^2 \frac{dm_{viL}}{dt} - C_{CL}^2 \frac{dm_{vdL}}{dt} \gamma P_{CL} \frac{dV_{CL}}{dt}\right]$$
<div align="right">Equation 4.24</div>

The cylinder pressure of the 2nd stage is:

$$\frac{dP_{CH}}{dt} = \frac{1}{V_{CH}}\left[C_{iH}^2 \frac{dm_{viH}}{dt} - C_{CH}^2 \frac{dm_{vdH}}{dt} \gamma P_{CH} \frac{dV_{CH}}{dt}\right]$$
<div align="right">Equation 4.25</div>

Where

V_C=Cylinder volume

$\frac{dm}{dt}$, Mass flow rates through the suction & discharge valves

C =Local speed of sound=$\sqrt{(\gamma RT_{CL,H}}$

R =The gas constant, $287m^2s^{-2}K^{-1}$ for air

γ = Ratio of major specific heats for the process gas =1.4 for air

The absolute temperature of air in the cylinder may be calculated from:

$$T_{CL,H} = T_{iL,H}\left(\frac{P_{CL,H}}{P_{iL,H}}\right)^{\frac{\gamma-1}{\gamma}}$$
<div align="right">Equation 4.26</div>

Where

$T_{iL,H}$=Average absolute temperature of the suction air in low and high pressure cylinders respectively,

$P_{iL,H}$=Suction pressure to low and high pressure cylinders respectively, and

$P_{CL,H}$=The internal cylinder pressure, low and high pressure cylinders respectively.

4.4.2 Cylinder Volume

According to (McCarthy, 1997) the volume of the cylinder in both stages $V_{cL,H}$ can be determined using:

$$V_{cL,H} = V_{coL,H} + S_{PL,H}X_{PL,H}$$ <div style="text-align:right">Equation 4.27</div>

Where:

$V_{coL,H}$=Clearance volume for the 1st and 2nd stages respectively

$S_{PL,H}$=Cross-sectional area of the piston,$=\frac{\pi}{4}d_{L,H}^2$

$X_{PL,H}$ = Piston displacement

d =Piston Diameter

The rate of change of cylinder volume $\frac{dV_{CL,H}}{dt}$ can be determined using:

$$\frac{dV_{CL}}{dt} = S_{PL}.\frac{dX_{PL}}{dt}$$ <div style="text-align:right">Equation 4.28</div>

$$\frac{dV_{CH}}{dt} = S_{PH}.\frac{dX_{PH}}{dt}$$ <div style="text-align:right">Equation 4.29</div>

Where:

$\frac{dX_{PL,H}}{dt}$ = The piston speed for the 1st or 2nd stage

4.5 Mass Flow Models

The rate of change of the cylinder pressure as given by equations (4.25) and (4.26) depends on the mass flow rate through the suction and discharge valves. Figure 4.7 shows the mass flow model.
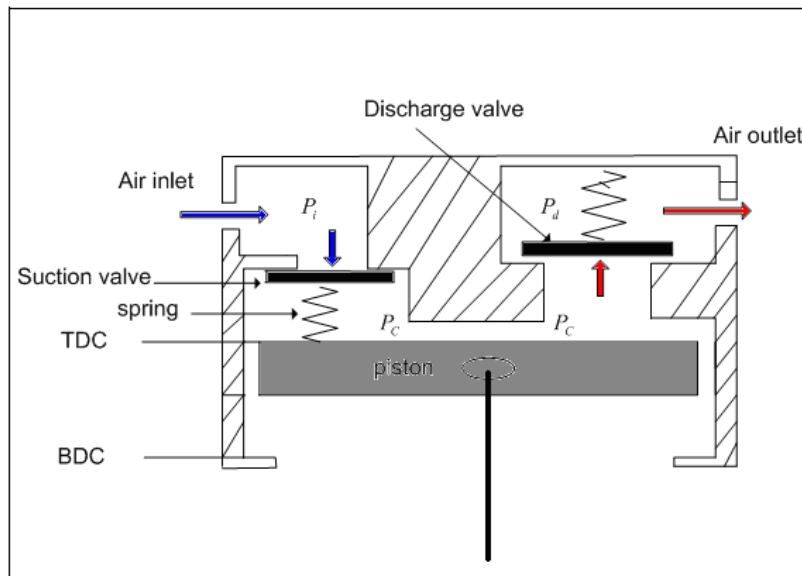
Figure 4.7 Mass flow models (Chaykosky, 2002)

Flow rates $\frac{dm_{viL,H}}{dt}$ and $\frac{dm_{vdL,H}}{dt}$ through the suction and discharge valves, respectively, of both cylinders. The expressions of mass flow rates can be derived by (Daniel 1994):

$$\frac{dm_{vd}}{dt} = \beta_d C_{dd}.A_{fd}\sqrt{2pc|P_c - P_d|}$$ <div style="text-align:right">Equation 4.30</div>

$$A_{fd} = 2\pi.r_{vd}.\max(x_{vd})$$ <div style="text-align:right">Equation 4.31</div>

Where:

A_{fd}=The maximum flow area of the discharge valve

r_{vd} =The discharge valve plate radius

$\max(x_{vd})$ =The maximum discharge valve plate displacement

C_{dd} =The discharge coefficient

P_d =The pressure in the discharge plenum

P_c =The density of the air in the cylinder

pc=The pressure in the cylinder

$\beta_d = \text{sign}(P_c - P_d)$, which is +1 for normal flow and −1 for backflow

Likewise, the mass flow rate through the suction valve is modelled as:

$$\frac{dm_{vi}}{dt} = \beta_d C_{di}. A_{fi}\sqrt{2pc|P_i - P_d|}$$
<div align="right">Equation 4.32</div>

$$A_{fi} = 2\pi. r_{vi}. \max(x_{vi})$$
<div align="right">Equation 4.33</div>

A_{fi} =The maximum flow area indicated

r_{vi} =The suction valve plate radius

$\max(x_{vi})$ =The maximum suction valve plate displacement

C_{di} =The suction coefficient

P_i =The pressure in the suction plenum

P_c =The density of the air in the cylinder

pc=The pressure in the cylinder

$\beta_d = \text{sign}(P_i - P_c)$

4.5.1 Suction Mass Flow Rate

During suction the mass flow rate ($m_{viL,H}$) of the air through the valve can be expressed as:

$$m_{viL,H} = \beta_{iL,H} C_{diL,H}(x). A_{fiL,H}\sqrt{2P_{cL,H}|P^e_{iL,H} - P_{cL,H}|}$$
<div align="right">Equation 4.34</div>

Considering the terms in equation (4.35) in order:

$\beta_{iL,H} = \text{sign}(P^e_{iL,H} - P_{cL,H})$, which are+1 for normal flow and -1 for backflow,

$C_{diL,H}(x)$ is a variable suction coefficient

$$C_{diL,H}(x) = 0.42\frac{x_{L,H}}{x_{\max L,H}}$$
<div align="right">Equation 4.35</div>

$A_{fiL,H} = 2\pi. r_{LH}. d_{iff}$ is the flow area around the valve plate (where d_{iff} is the distance between the outer edge of the valve plate and the inner wall of the valve chamber), $P_{cL,H}$ is the density of the air in the cylinders, $P^e_{iL,H}$ is the pressure in the suction plenum, $P_{cL,H}$ is the cylinder pressure. Figure 4.8 shows the linear relationship between the normalised valve displacement and the flow coefficient.
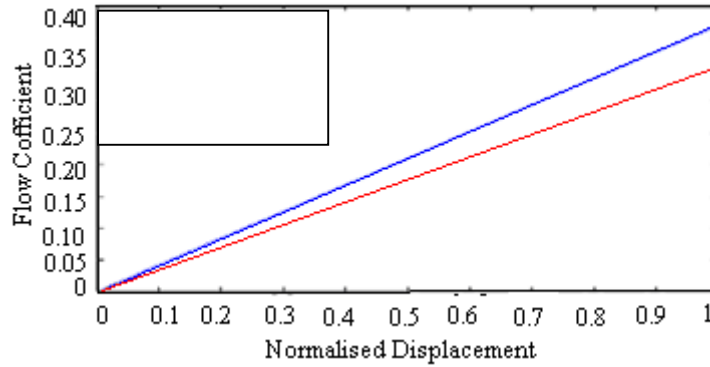
Figure 4.8 Variable flow coefficients for the suction and discharge valves (M Elhaj, 2005)

$$P_{iH} = P_{dL}(\frac{P_{iH}}{P_{dL}})^{1/\gamma}$$
<div align="right">Equation 4.36</div>

Where $P_{iH,L}$ is density of the air at intake of each cylinder, P_{dL} is density of the air at 1st stage discharge, and the air density is taken as $P_i = 1.77 \text{kg/m}^3$

4.5.2   Discharge Mass Flow Rate

The discharge mass flow rate is given by:

$$\dot{m}_{vdL,H} = \beta_{dL,H} C_{ddL,H}(x).A_{fdL,H}\sqrt{2P_{cL,H}|P_{cL,H} - P^e_{dL,H}|}$$
<div align="right">Equation 4.37</div>

Considering the terms in equation (4.37) in order:

$C_{ddL,H}(x) =$ is a variable discharge coefficient which accounts for the reduced flow area resulting from the separated flow.

$$C_{ddL,H}(x) = 0.35\frac{x_{L,H}}{x_{\max L,H}}$$
<div align="right">Equation 4.38</div>

$A_{fdL,H} =$ is the maximum flow area of discharge valve.

The flow coefficients are based on those of Price (M. Elhaj, Gu, & Ball, 2004), who took measurements on similar valves. To allow for the possibility of backflow, the absolute value of the pressure drop across the valve is taken, where $P^e_{dL,H}$ is the pressure in the discharge plenum,

$\beta_{dL,H} = \text{sign}(P_{cL,H} - P^e_{dL,H})$, which is +1 for normal flow and -1 for backflow.

$$P_{cL,H} = P_{iL,H}\left[\frac{P_{cL,H}}{P_{iL,H}}\right]^{\frac{1}{\gamma}}$$
<div align="right">Equation 4.39</div>

Where, $P_{cL,H}$ is the density of the air in the cylinders, $x_{\max L,H}$ is the maximum valve plate displacement, $x_{L,H}$ is the valve displacement. The two discharge valves have the same maximum displacement; the two suction valves have the same displacement and $P_{iL,H}$ is the density of the air in the plenum.

$$P_{dL,H} = P_{iL,H}\left[\frac{P_{dL,H}}{P_{iL,H}}\right]^{\frac{1}{\gamma}}$$
<div align="right">Equation 4.40</div>

Where $P_{dL,H}$, is the density of the air in discharge valves, and $P_{iL,H}$ is the density of the air in the suction valves.

4.6   Model of Valve Motion

The discharge and suction valves in a RC are similar in design with each constructed with a spring, a valve plate, and a chamber. RC valves normally open and close once per one revolution of the crankshaft. The movement of the valve plate may be described by a series of processes: the initial opening of the valve and movement of the valve until it impacts on the valve seat (M. Elhaj, et al., 2004; Kryter & Haynes; Naid, 2009). The valve plate motion prior to impact is assumed linear and can be approximated as a single-degree-of-freedom system, as shown in Figure 4.9.
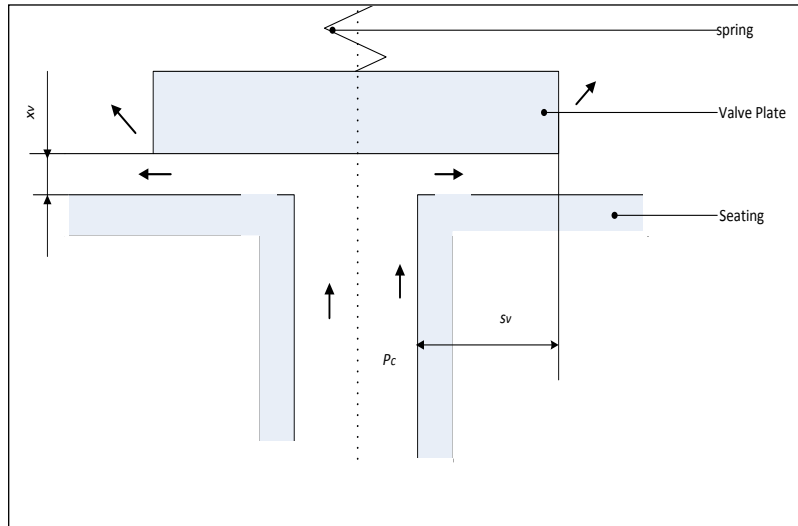
Figure 4.9 Schema of a plate valve

4.6.1 Discharge Valve Motion

The equation of motion of the suction valve and gas force on the valve during the rapid opening, described by Newton's second law in the following equation:

$$m_{vd}\frac{d^2x_{vd}}{dt^2} + C_d\frac{dx_{vd}}{dt} + k_dx_{vd} = f_d \hspace{3cm} \text{Equation 4.41}$$

Where:

m_{vd} =Total mass of the discharge valve unit, = mass of valve plate plus one-third of the spring mass.

f_d =Force on the discharge valve plate due to the difference in pressures between the cylinder and discharge pressures (Pc and Pd, respectively)

k_d = Discharge valve spring stiffness,

c_d =Damping constant of the discharge valve chamber,

$\frac{dx_{vd}}{dt}$ =Discharge valve velocity

$\frac{d^2x_{vd}}{dt^2}$ =Discharge valve acceleration

x_{vd} =Vertical displacement of discharge valve plate.

4.6.2 Suction Valve Motion

The equation of motion of the suction valve can be represented as:

$$m_{vs}\frac{d^2x_{vs}}{dt^2} + C_d\frac{dx_{vs}}{dt} + k_sx_{vs} = f_s$$
<div align="right">Equation 4.42</div>

Where:

m_{vs} =Total mass of the suction valve unit = mass of valve plate plus one-third of the spring mass.

f_s = Force on the discharge valve plate due to the difference in pressures between the cylinder and discharge pressures (Pc and Pd, respectively).

k_s =The suction valve spring stiffness

C_s =The damping constant of the suction valve chamber

$\frac{dx_{vs}}{dt}$ = The suction valve velocity

$\frac{d^2x_{vs}}{dt^2}$ =The suction valve acceleration

x_{vs} =The vertical suction valve plate displacement.

The equation of motion of the discharge valve before it impacts on the seat is as follows;

$$m_{vdL,H}\ddot{x}_{vdL,H} + C_{dL,H}\dot{x}_{vdL,H} + K_{vdL,H}x_{vdL,H} = \sum f_{vdL,H}$$
<div align="right">Equation 4.43</div>

When the valve plate hits the valve seat (valve totally open or closed) the equation of motion becomes:

$$\dot{m}_{vdL,H}\ddot{x}_{vdL,H} + C_{cL,H}\dot{x}_{vdL,H} + K_{cdL,H}x_{vdL,H} = \sum f_{vdL,H}$$
<div align="right">Equation 4.44</div>

$m_{vdL,H}$= is the mass of the valve plate plus one-third of the spring mass,

$C_{dL,H}$=is the damping coefficient,

$K_{vdL,H}$=is the non-linear springs stiffness during the period when the valves are not in contact with the valve seats,

$K_{cdL,H}$=is the contact stiffness between seat and valve plates,

$C_{cL,H}$=is the contact damping coefficient,

$\ddot{x}_{vdL,H}$ =is the valve acceleration,

$\dot{x}_{vdL,H}$ =is the valve velocity,

$x_{vdL,H}$ =is the valve displacement, and

$\sum f_{vdL,H}$=is sum of the total forces acting on the valve plate.

The displacement of the discharge valve is governed by equations (4.43, 4.44) where the force term on the right-hand side is given by:

$$\sum f_{vdL,H} = fvd_{vdL,H} - fgd_{L,H} - fdo_{L,H}$$
<div align="right">Equation 4.45</div>

Where

$fgd_{L,H} = -mg_{L,H}$, is the weight of the discharge valves

$fvd_{L,H}$=is the preset spring force (Daniel 1994; Fleming, 1989).

The last term describes the pressure force, $fvd_{L,H} = cfd_{L,H}Sv_{L,H}(Pc_{L,H} - fd_{L,H})$ for the discharge valves, where: $cfd_{L,H}$, is the force coefficient, $Sv_{L,H}$ is the slot area for a single channel, $Pc_{L,H}$ is the cylinder pressure, and $fd_{L,H}$ is the pressure in the discharge plenum.

4.7   Fault Simulation

The faults examined are leakage through the 2nd stage discharge and suction valves, intercooler and a loose drive belt. Simulation is used in each case to predict the signatures for these faults which will then be tested experimentally. The results of the simulation should also provide helpful guidance in examining the experimental data.

4.7.1 Simulation of Discharge Valve Leakage

In the simulation, a leak is modelled as additional flow through an orifice in parallel to the normal valve flow. It is assumed that gas will flow through the leak when a pressure difference exists across the leak. This will occur for the discharge and the following equation was used to determine the mass flow rate for gas flow through the discharge valve orifice.

$$\dot{P}_{cH} = \frac{1}{V_{cL,H}}[C^2{}_{iL}\dot{m}_{vil} - C^2{}_{cH}\dot{m}_{vdH} - \gamma P_{cH}v_{cH} - C^2{}_{iH}\dot{m}_{iH-}C^2{}_{cH}\dot{m}_{dH}]$$

Equation 4.46

$$\dot{m}_{dH} = \beta_{dH}C_{dH}A_{dH}\sqrt{2P_{cH}|P_{cH} - P_{dL}|}$$

Equation 4.47

Where A_{dL}=the leakage flow area of the discharge valve

$\beta_{dl,H} = sign(P_{dL,H} - P_{cL,H})$= the flow direction parameter: $+1$ for the leakage flow from the cylinder into discharge passage, and -1 for leakage flow from the discharge passage back into the cylinder.

4.7.2 Simulation of Suction Valve Leakage

Similarly for the suction valve the calculation is:

$$\dot{m}_{sH} = \beta_{sH}C_{sH}A_{sH}\sqrt{2P_{cH}|P_{cH} - P_{sL}|}$$

Equation 4.48

Where

A_{sL}=the leakage flow area of the suction valve.

4.7.3 Simulation of Belt Transmission Dynamics
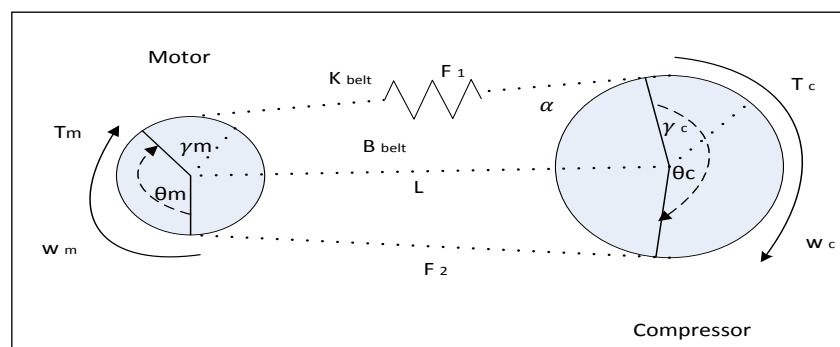
Figure 4.10 Belt transmission systems

$$v_s = v_0(1 - s)$$

Equation 4.49

$$v_0 = \frac{1}{B_r}$$

Equation 4.50

Where

v_0 =is the inverse of the transmission ratio B_r

The transmission ratio is $B_r = \frac{r_c}{r_m}$ and $s = 0.001$ is a suitable value for the slip coefficient (Kwon, Tae-Suk, Lee, Dong-Hoon, & Sul, 2005).

$$J_m \frac{d\omega_m}{dt} = T_{em} - C_b \Delta \dot{\theta} - K_b \Delta \theta \qquad \text{Equation 4.51}$$

$$J_c \frac{d\omega_c}{dt} = (C_b \Delta \dot{\theta} + K_b \Delta \theta)/v_0 T_c \qquad \text{Equation 4.52}$$

Where

$$\Delta \theta = \theta_m v_s - \theta_c \qquad \text{Equation 4.53}$$

$$\Delta \dot{\theta} = \dot{\theta}_m v_s - \dot{\theta}_c \qquad \text{Equation 4.54}$$

K_b = is the spring constant of the belt (Nm/rad)

C_b = is damping constant of the belt and pulleys [Nms/rad]

4.7.4 Simulation of Intercooler Leakage

Assuming there is a small leakage in the intercooler of the two stage compressor then

$$\dot{P}_{in} = \frac{1}{V_{in}} [C^2{}_{cL} m_{vdl} - C^2{}_{iH} m_{viH} - C^2{}_{in} m_{in}] \qquad \text{Equation 4.55}$$

$$m_{in} = \beta_{in} C_{in} A_{in} \sqrt{2 P_{in} |P_{in} - P_0|} \qquad \text{Equation 4.56}$$

Where

$A_{in} = 2\pi r_{L,H} x$ is leakage flow area allowing gas to escape from the intercooler

$$\beta_{in} = sign(P_{in} - P_0) \qquad \text{Equation 4.57}$$

4.8   Model Validation

The model was validated by comparing the results of the simulation with experimental data. The measurement procedures are described subsequently. The valves motion for the high pressure stage of the RC is described only for different values of discharge pressure: 80, 100 and 120psi.

4.8.1 Numerical solution procedure

The 2nd stage piston leads that of the 1st stage by π/2 and so, initially, the piston will be approximately in the middle of its suction stroke, and the pressure in the 2nd stage cylinder will be higher than in the first. When either piston is at TDC, the volume enclosed within the cylinder will be equal to the relevant clearance volume, and the pressure and temperature of the gas within the cylinder will be assumed to be equal to the discharge gas pressure and temperature. By simultaneously solving the pressure equations 4.24 and 4.25 and mass flow rate equations for the discharge and suction valves equations 4.30 and 4.34 and using the first law of thermo-dynamics, the instantaneous cylinder pressure and mass of the gas inside the cylinder can be computed. The rate of change of cylinder volume is calculated using equations 4.28 and 4.29. By using the valve dynamics equations 4.37, 4.42 and 4.41, the valve lift for the suction and discharge valves can be determined. The temperature of the gas within the cylinder is calculated using the equation of state, 4.26. This process is repeated for consecutive time intervals up to the moment the suction valve opens. The determination of the suction process begins with an assumed pressure within the cylinder. The valve lift is then

calculated for this pressure. By using the calculated valve lift, and solving the pressure equation 4.24 and continuity equations 4.30 and 4.31 simultaneously, the pressure and mass of air inside the cylinder may be calculated. This procedure is repeated until the calculated and assumed cylinder pressures are nearly the same. Using the equation of state, the cylinder temperature is calculated. The process of suction (re-expansion) continues up to the time when the suction valve opens. With the shutting of the suction valve, compression begins and the values of cylinder pressure and temperature may be determined by steps parallel to those used in the expansion procedure. As soon as the discharge valve opens, the discharge process starts. During the discharge procedure, the cylinder pressure and temperature values are determined using steps similar to those used in the suction process. At the end of the discharge cycle (TDC), the calculated cylinder pressure and temperature are compared with the values assumed at the start of the calculations. If there is any significant difference, the calculations are repeated (with the calculated end values now being the start values for the next cycle of calculations) until there is sufficiently good agreement between the assumed start and the calculated end values. The prediction of the compressor performance in terms of volumetric efficiency, volume flow rate and the performance ratio is achieved by computing the mass flow rate through suction and discharge valves together with other parameters. A theoretical simulation for a healthy two-stage, two cylinder, RC was carried out assuming the 1st stage suction valve was open to atmosphere, (atmospheric pressure assumed to be 101kPa (14.7psi)). The 1st stage discharge pressure was 0.27MPa (40psi) (M Elhaj, 2005). The model assumed that during compressor operation the discharge pipe of the 1st stage was connected to the suction pipe of 2nd stage through the intercooler. The 2nd stage suction pressure was taken as 0.21MPa (32psi). The maximum discharge pressure was taken to be 0.8MPa (120psi), and the exhaust pipe of 2nd stage was assumed connected directly to the storage tank. To evaluate the effects of different storage pressures on the system, the storage tank release valve was set at different discharge pressures.

4.8.2 Physical Parameters and Constants

The important numerical parameters needed as inputs to simulation are listed in Table 4.1 and Table 4.2.

	Low Pressure Cylinder	High Pressure Cylinder
Piston Mass [kg]	1.78	0.89
Piston Head Diameter [mm]	93.6	55.6
Cylinder Bore [mm]	101.6	63.5
Suction Pressure [kPa/psi]	100/14.7	220/32.2
Discharge Pressure [kPa/psi]	270/39.7	816/120
Suction temperature [^{o}C]	21	41
Discharge temperature [^{o}C]	50	80

Table 4.1 First and second stage piston and cylinder parameters

A. MOTOR PARAMETERS	
Motor Speed [rpm]	1420
Motor Power [Kw]	2.2
B. COMPRESSOR	
Number of Cylinders	2 (90º Opposed)
Compressor Speed [rpm]	425
Flywheel Ratio	3:1
Diameters of the motor pulley [cm]	12
Diameters of the compressor pulley [cm]	36
Tank Capacity [litres]	272
Piston Stroke [mm]	76.2
Connection Rod Length [mm]	171.6
Crank Radius [mm]	38.1
C. VALVE SYSTEM	
Maximum Suction Valve lift [mm]	2.2
Maximum discharge Valve lift [mm]	2.6
Mass of Valve Plate, Low Pressure Cylinder [g]	2.3
Mass of Valve Plate, High Pressure Cylinder [g]	2.1
Mass of Valve Spring , Low Pressure Cylinder [g]	1.0
Mass of Valve Spring, High Pressure Cylinder [g]	2.0
Outer Radius Valve Plate, Low Pressure Cylinder [mm]	21.0
Outer Radius Valve Plate, High Pressure Cylinder [mm]	14.0
Inner Radius Valve Plate, Low Pressure Cylinder [mm]	12.5
Inner Radius Valve Plate, High Pressure Cylinder [mm]	10.5

Table 4.2 Electrical Motor, Broom Wade TS 9-16 Compressor and Valve System Details

4.9   Simulation

Dynamic models play an important role in the development of the pressure and the vibration based machinery diagnosis. Creating the model can be used to understand and predict fault signatures. The model can also be used as guidance and the same faults can then be introduced into the machine with the results from the model indicating which ones are detectable and how the recovered forces should vary.

4.9.1 Second Stage Valve Operations

For the piston in the 2nd stage, the higher pressure cylinder leads the first by $\pi/2$, thus when the 1st stage piston is at TDC, the 2nd stage suction valve will be open and the discharge valve closed. Figure 4.11 shows the estimated crank angles at which the discharge valve opens and closes. However, the time at which the discharge valve opens and closes is expected to depend significantly on the discharge pressure. The higher the discharge pressure, the later the valve will open because a higher pressure in the discharge tank will require a correspondingly higher pressure in the cylinder before the 2nd stage discharge valve spring opens.
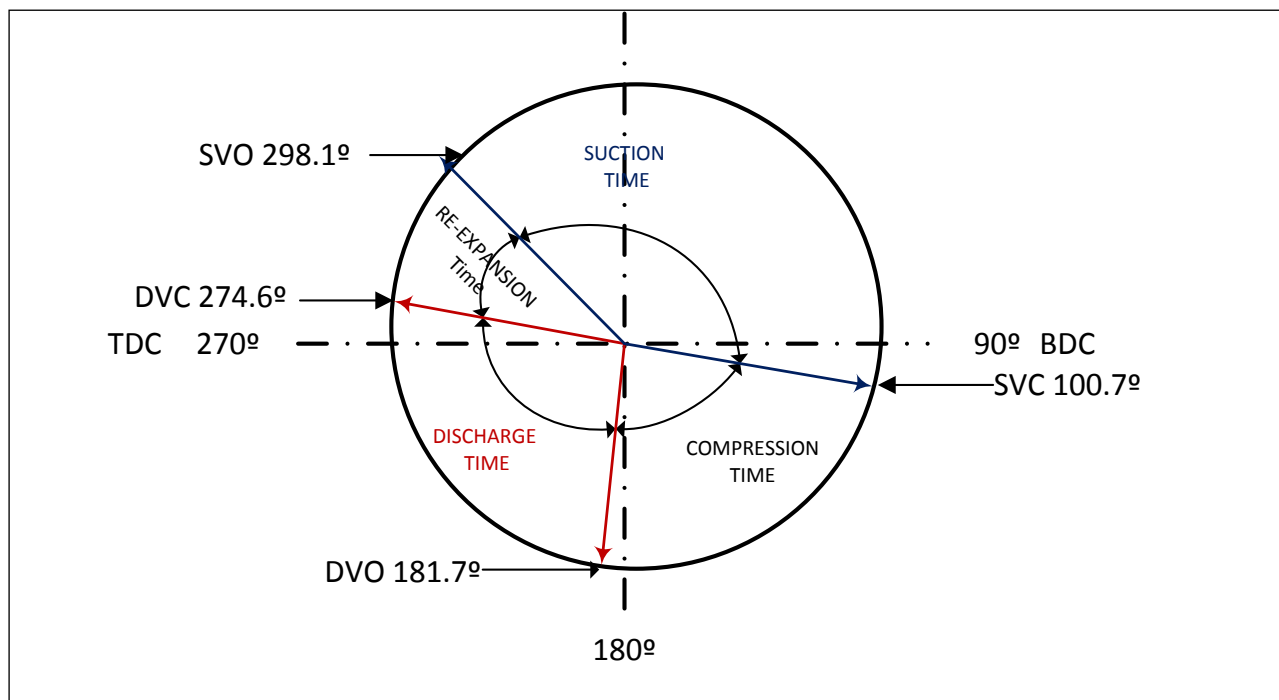
Figure 4.11 Predicted crankshaft angles at which the valves open and close for 2nd stage of a healthy compressor with discharge pressure 100psi

Figure 4.12 is compiled from the predicted displacement of the suction and discharge valves in the 2nd stage of the compressor, as shown in Figure 4.12. The crank angles at which the valves open and close are obvious and it can be clearly seen that the suction valve is predicted to open at a crankshaft angle 298.1º and the discharge valve is predicted to open at an angle of 181.7º. The closing of the suction and discharge valves are predicted as 100.7º and 274.6º, at discharge pressure 100psi respectively.
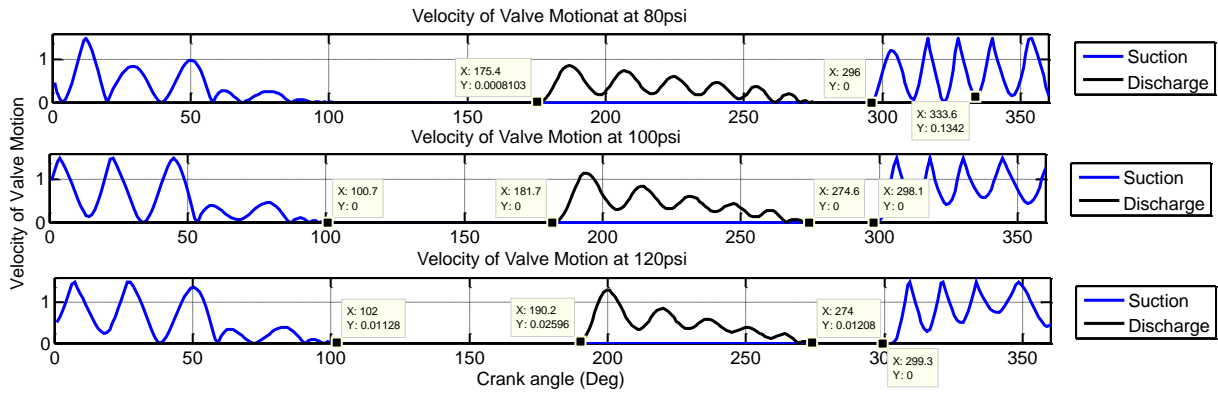
Figure 4.12 Predicted motion of the discharge valve and suction valve for the 2nd stage of a healthy compressor with different discharge pressure

4.9.2 Comparison of Measured and Simulated Signals

Figure 4.13 shows the pressure in the 2nd stage cylinder, as measured experimentally and show the same pressure as predicted by the mathematical model. The measured and predicted cylinder pressure shows a rapid drop during re-expansion, as the piston moves down the cylinder. When the cylinder pressure drops sufficiently below that of the inlet manifold, the suction valve opens and air enters the cylinder. The pressure shows some slight initial oscillation, particularly in the 2nd cylinder, due to suction valve flutter. The pressure fluctuation rapidly dies away and the gas pressure attains almost constant value during the remainder of the suction time, as the piston moved down the cylinder. After the suction valve closes and the piston moves up the cylinder, the air above it is rapidly and uniformly compressed causing a corresponding increase in the cylinder pressure.
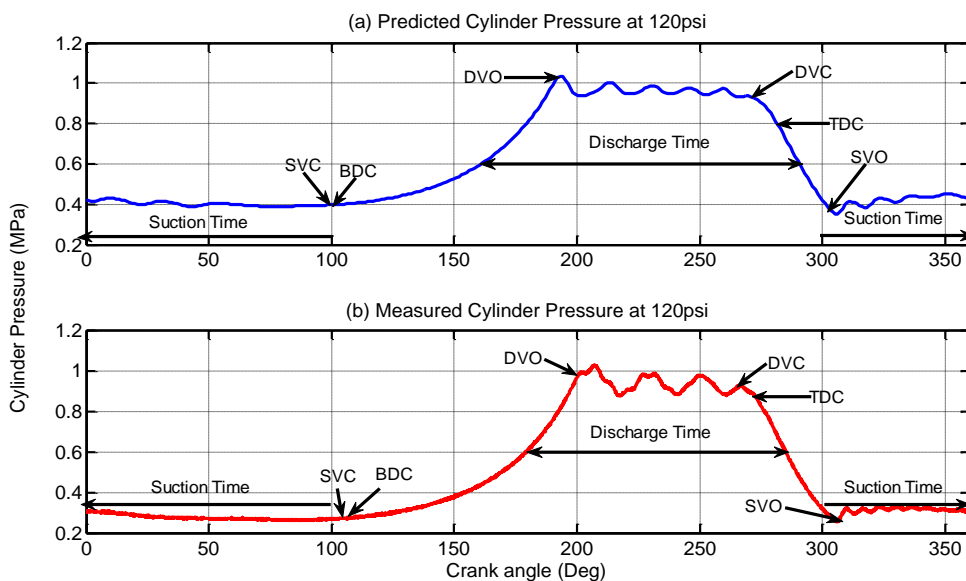
Figure 4.13 Predicted pressure and Experimental results in the 2nd stage cylinder of a healthy compressor during cycle pressure 120psi

Figure 4.14 shows the measured and predicted pressures for different values of the discharge pressure; 80, 100 and 120 psi. The maximum pressure in both cylinders increases with the discharge pressure. The predictions of the model are confirmed both quantitatively and qualitatively. Although the operational range of the compressor is specified as 80–120 psi, it was considered useful to use discharge pressures as low as 80 psi to test the model.
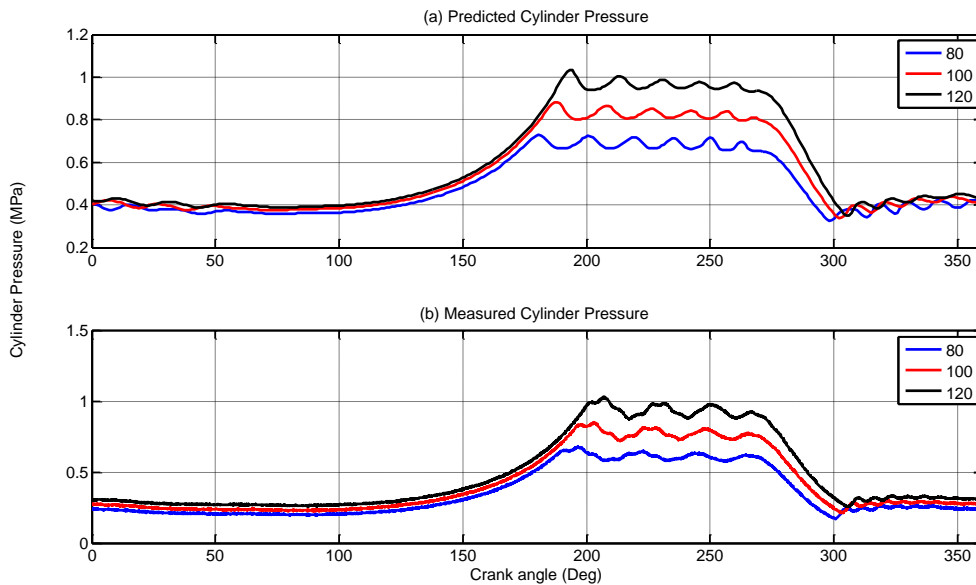
Figure 4.14 Predicted cylinder pressure and measured cylinder pressure for different discharge pressure for base line compressor

A more direct comparison of the pressures between the measured and predicted is shown in Figure 4.15. The difference that occurs during the compression and discharge processes of the 2nd stage is considered most likely due to small error in the values assigned to the flow coefficients parameters used under different pressures in the equations.
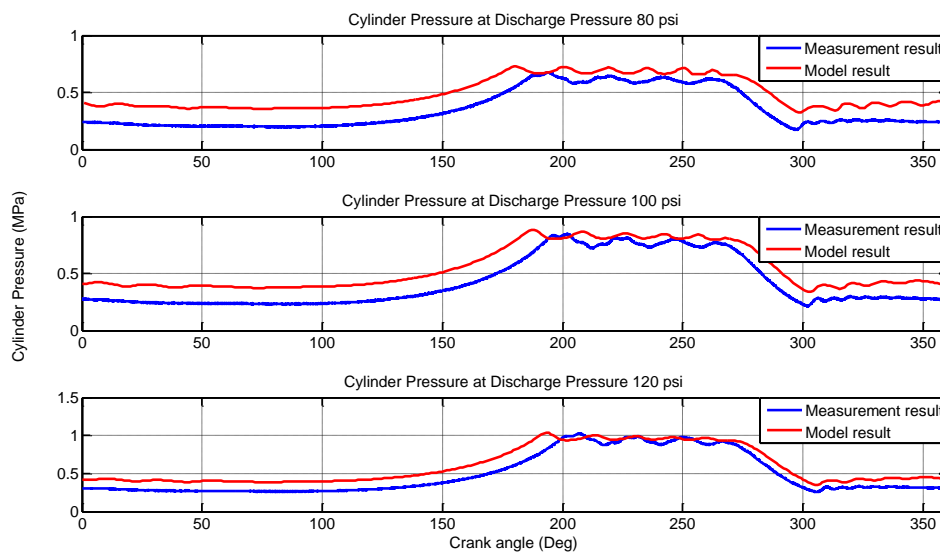
Figure 4.15 Direct comparison of experimental and simulated dynamic pressure signal in 2nd cylinder with different discharge pressure
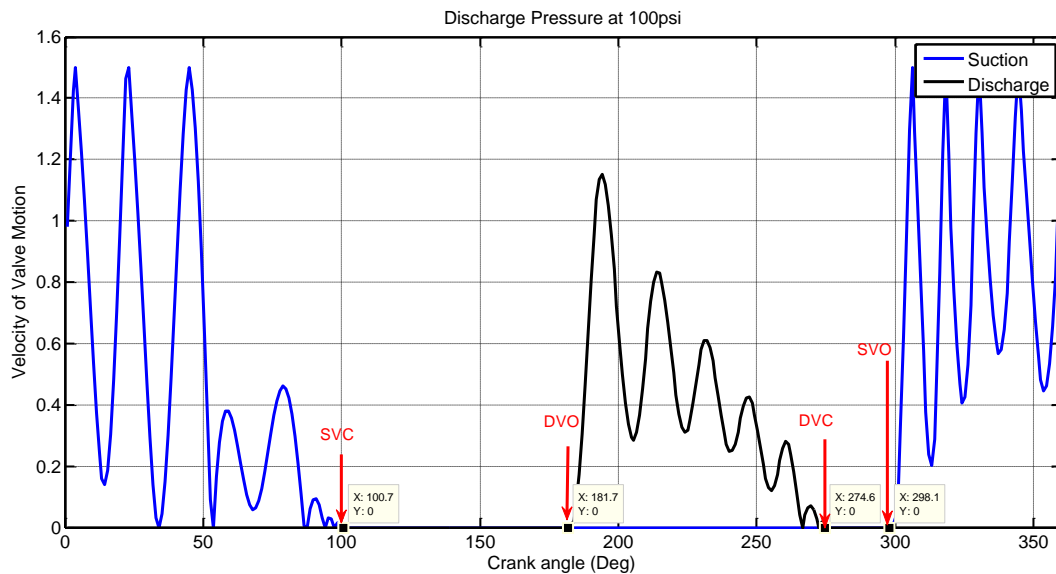
Figure 4.16 Predicted motion of the discharge valve and suction valve for the 2nd stage of a healthy compressor at discharge pressure 100 psi

Figure 4.17 shows the measured pressure and vibration signal from the head of the 2nd stage compressor cylinder for one complete period of the compressor cycle and four significant transient vibration responses can be seen, each consistent with a corresponding valve impact event: suction valve opening, suction valve closing, discharge valve opening and discharge valve closing and the distance between the two events is identified as valve plate bounces. When the TDC of the piston in the 1st stage cylinder is taken as the reference position, the sequence of the four events for the 2nd stage will be identified as show in Figure 4.17. It can be seen clearly that the opening impact of the discharge valve coincides with the largest transient vibration. The closing of the discharge valve as can be seen by the motion of the medium channel in the valve. The next transient vibration correlates with the opening of the suction valve as presented by the motion of the smallest channel in the valve, and the next one after that coincides with the suction valve closing.
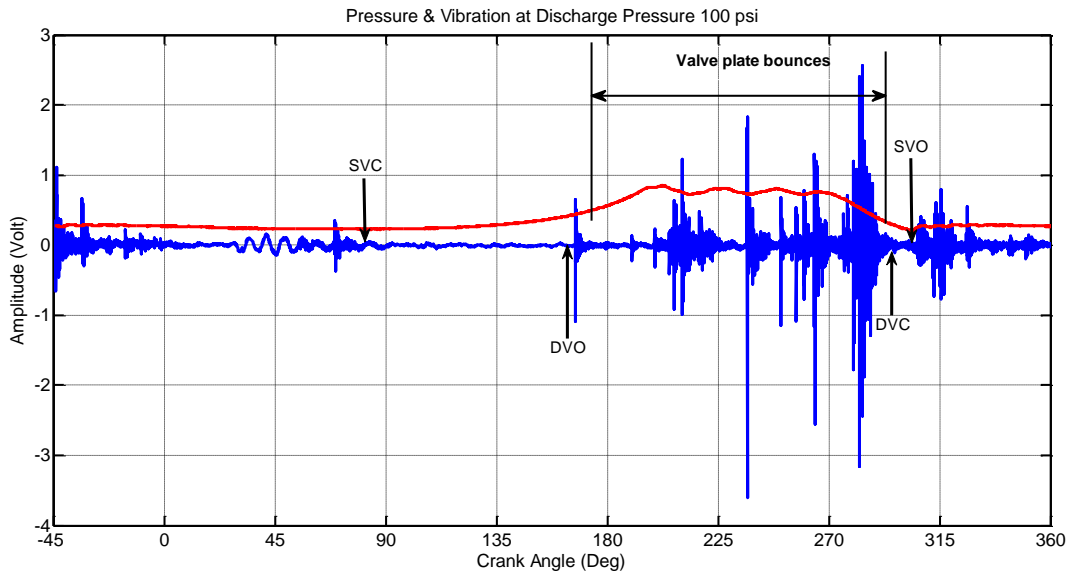
Figure 4.17 Pressure and Vibration signals from 2nd stage cylinder head as a function of crank angle for a healthy compressor with discharge pressure (100psi)

Figure 4.18 clearly shows how change in the discharge tank pressure affects the opening and closing times of the 2nd stage valves which are shown for three discharge tank pressures: 80psi, 100psi and 120psi. Particularly striking is the change in the crankshaft angle at which the 2nd stage discharge valve opens with increase in discharge pressure; as expected the higher the tank pressure the later the valve opens. If the discharge valve closes earlier, the greater the discharge pressure in the storage tank. The greater the discharge pressure, the pressure in the cylinder must be greater to open the discharge valve, and this will require greater movement of the piston up the cylinder and hence a greater crank angle. See Table 4.3.

Discharge Pressure	DVO	DVC	SVO	SVC
80psi	162.4	290.3	303.0	85.44
100psi	169.6	292.4	306.0	86.63
120psi	173.3	293.1	307.7	92.17

Table 4.3 Showing suction and discharge valve opening and closing angles as a function of discharge pressure for a healthy compressor at 2nd stage
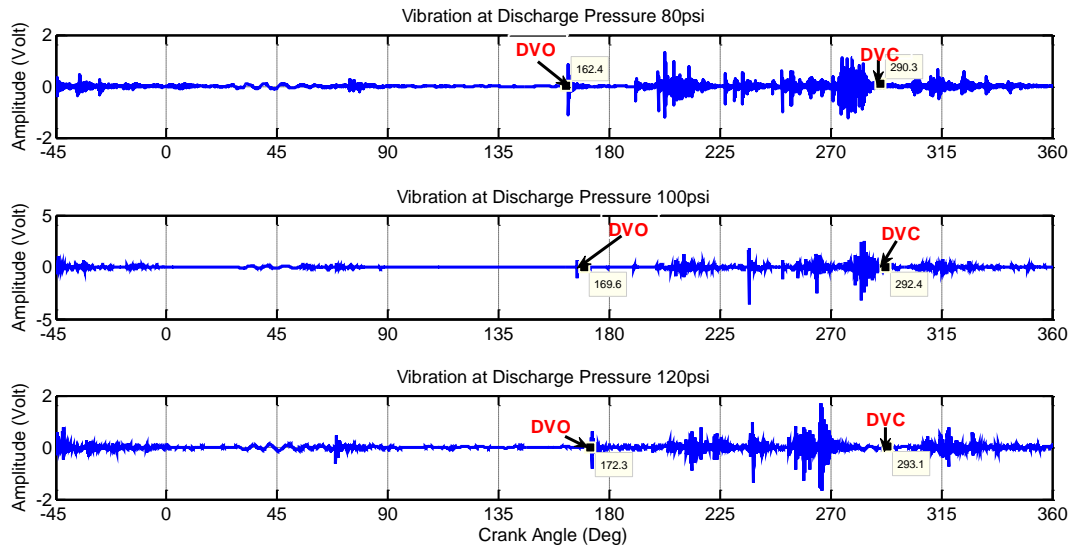
Figure 4.18 Measured vibration signals from 2nd stage cylinder head showing suction and discharge valve opening and closing angles as a function of discharge pressure, for a healthy compressor

4.9.2.1   *Weakened Suction Valve Spring*

A worn suction valve spring is introduced into the compressor by reducing the spring stiffness. The predicted results on the effects of valve displacement due to weakening the suction valve spring are shown in Figure 4.19. As the figure shows, the suction valve motion is affected by this fault; the weakening spring causes the suction valve to open slightly earlier, because a low pressure drop is required across the valve for opening. The most noticeable effect is the discharge valve opens slightly later, with higher impact velocity as the springs weaken. As seen in the figure, the later closing occurs because of the spring force needed to close the valve has been reduced. The impacts also increase in strength if the piston has reversed direction prior to valve closure, forcing the still open valve plate harshly into the seat.
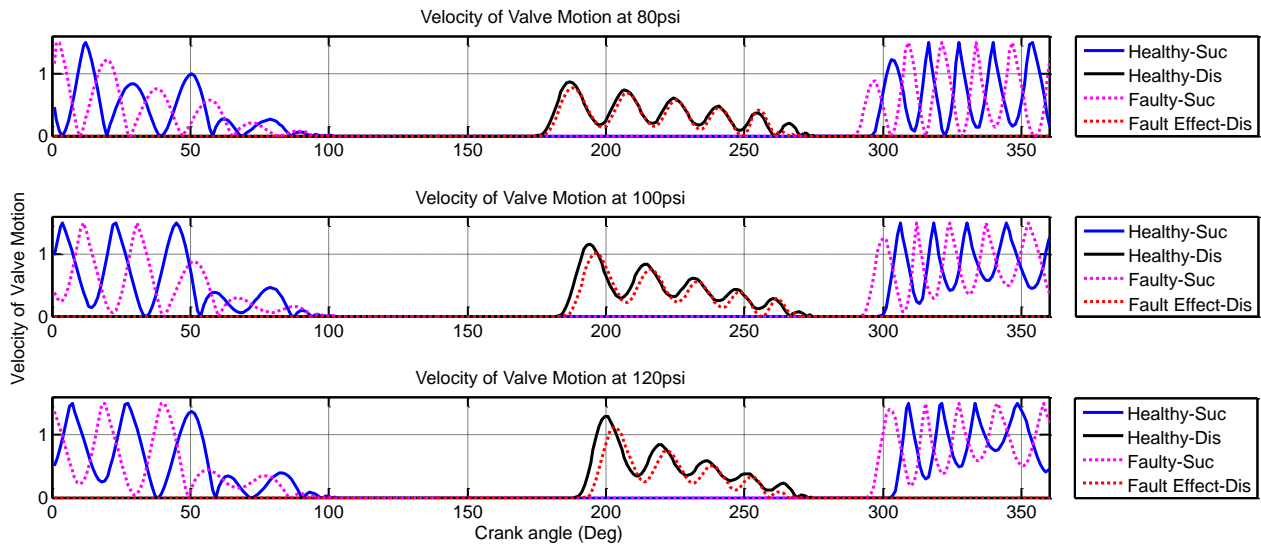
Figure 4.19 Predicted motion of suction and discharge valves for the 2nd stage with suction valve leakage

Figure 4.20 shows measured vibration valve impact signatures for discharge and suction valves opening and closing impact as a function of weakening suction springs at 2nd stage. The timing of the opening impact advances as weakening springs decrease in the predicted model. At the discharge pressure 100 psi (a healthy spring) the suction valve opens at 304.6º. For a weakened spring at 302.4º.The effects of a weakened 2nd stage suction valve spring on the 2nd stage discharge valve are shown in Figure 4.19. The discharge valve opens later due to the slightly lower than normal pressure in the cylinder. For a healthy spring the discharge valve opens at 169.6º and for a weakened suction spring it opens at 170.0º. The experimental data follows the predicted results. This shows that for the weak spring measurements, there is good relationship with the model prediction.
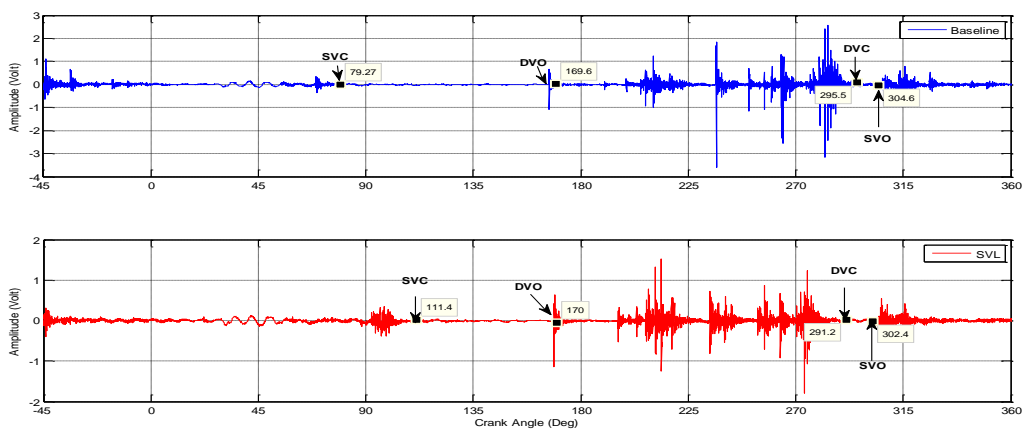
Figure 4.20 Time domain representation of vibration signal from accelerometer on the 2nd stage cylinder with worn spring suction valve

4.9.2.2   *Leaky Discharge Valve*

A leaky valve is a common fault in RCs. It causes high-temperature air to be forced across the valve surface by differential pressure, which accelerates the deterioration of the valve system (including the valve spring) and reduces compressor efficiency considerably. Experimentally, a leaky discharge valve was simulated by drilling a small diameter hole in the valve plate for the 2nd stage. The hole was 2.0 mm diameter, which was about 2.0% of the maximum free area of the valve. The leakage was introduced into the high-pressure stage where the valve works under much harsher conditions and a fault such as this have much greater impact on compressor performance. The predicted variations in valve displacements due to a leaky discharge valve are shown in Figure 4.21. The leak causes the discharge valve to open earlier, by the opening impact velocity, decreases only slightly because of pressure equalization due to the leak. The reason for this is that high pressure air flowing through the discharge leak raises the pressure in the cylinder above that which normally exists. As a result, the pressure needed to open the valve is reached sooner. The suction valve motion is also changed because the higher cylinder delays the time at which the pressure is low enough to open the valve. The final closing impact for the suction valve also occurs earlier.
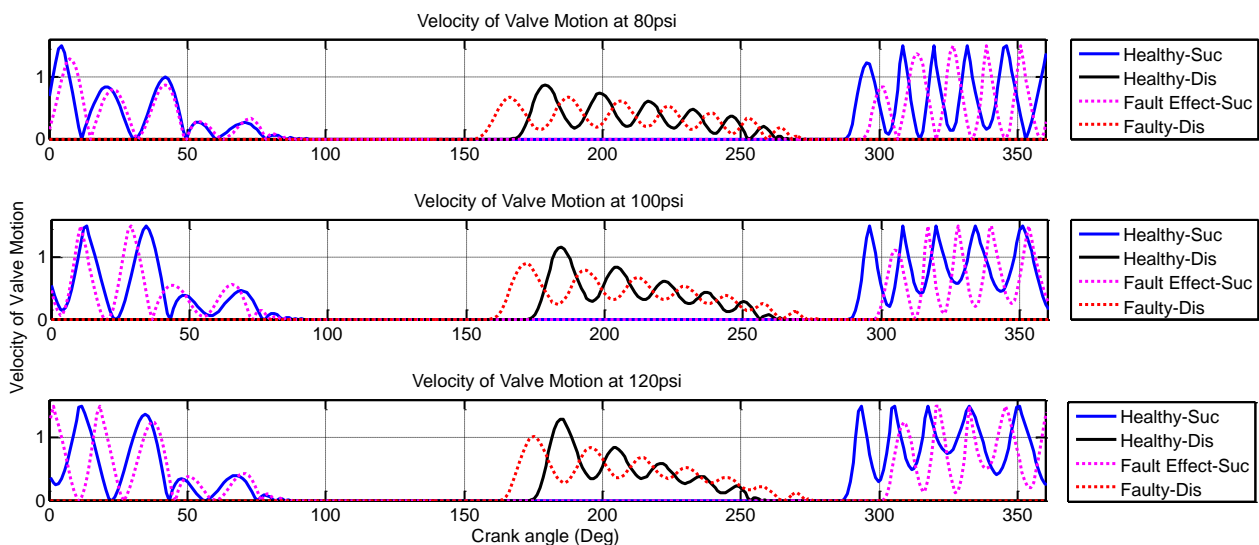
Figure 4.21 Predicted motion of suction and discharge valves for the 2nd stage with leaky discharge

Figure 4.22 shows the trace of amplitude of the vibration detected by the accelerometer for the 2nd stage valves impact signatures (for discharge and suction valves opening and closing impact as a function of discharge leakage). The timing of the opening impact advances as leakage size increase in the predicted model. The experimental data follows the predicted trend up to a leak size. When the leakage occurs in discharge valve in the 2nd stage then, during the suction stroke high pressure air flows back through the discharge leak and raises

the pressure in the cylinder above that which would exist under healthy condition. As a result, the pressure needed to open the discharge valve is reached earlier. The pressure to open the suction is reached later because the leakage delays the time at which the pressure differential across the suction valve is sufficient to open the valve. With the discharge pressure 100 psi, the healthy second stage discharge valve opening is about 169.6º. When 2.0% leakage is introduced, this becomes 156.0º and the discharge valve leakage also has an effect on the suction valve of 2nd stage, the suction valve will open later at 308.6º and closes at 103.9º while in the healthy state it opens at 306.0º and close at 103.9º.
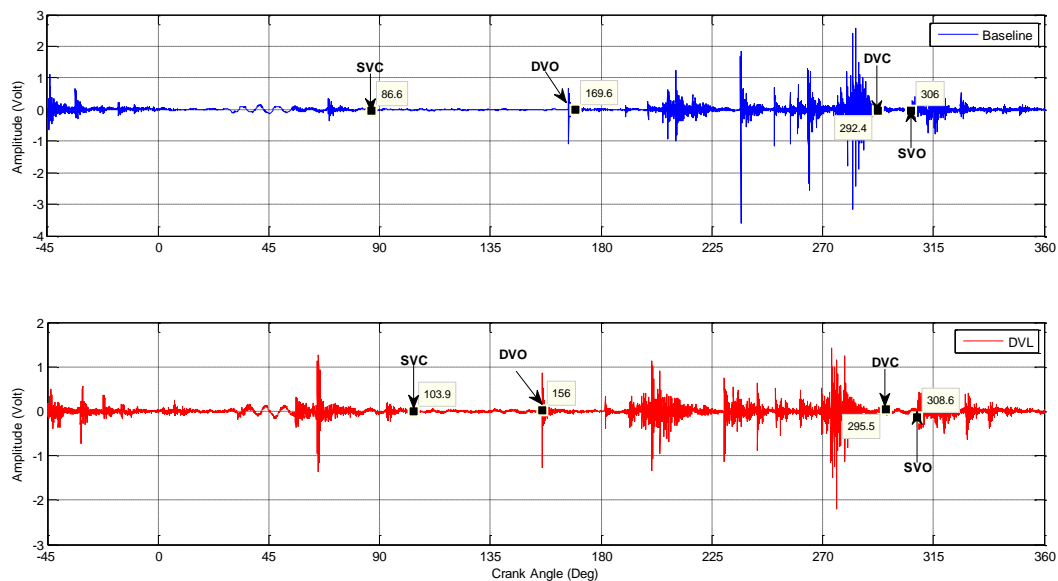
Figure 4.22 Time domain representation of vibration signal from accelerometer on the 2nd stage cylinder with discharge valve leakage

4.9.2.3   *Intercooler leakage*

Leakages in compressor pipeline systems are also very common and are often caused by vibration and thermal stress. A small leakage fault was inducted into the intercooler pipe which connects the discharge from the 1st stage to the suction of the 2nd stage of the compressor. In order to create such a fault in the compressor, the nut holding the intercooler pipe onto the second stage was loosened, as indicated in chapter 3 (Figure 3.3). The degree of fault was controlled by loosening the nut a different number of turns. The more turns, the more leakage. In this study, one complete turn is found to produce an observable change to the performance of the compressor.

Figure 4.23 shows the predicted motion of suction and discharge valves for the 2nd stage with intercooler leakage. The leakage causes a distinct delay in the opening of the 2nd stage valves (both the discharge and the suction). The reason for this is that the cylinder pressure is lower than normal during the compression and suction strokes. This suggests that intercooler leakages can be detected from the fluctuations in the pressure waveforms.
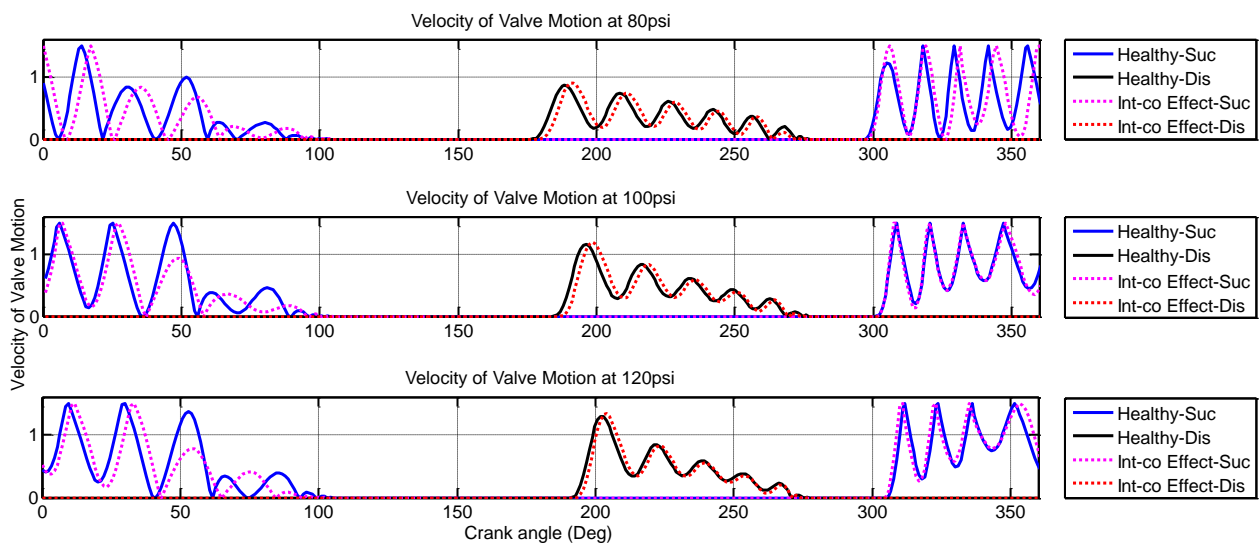
Figure 4.23 Predicted motion of suction and discharge valves for the 2nd stage with intercooler leakage

The changes in the signal can be explained from the compressor working mechanism. For the 1st stage, the change in signal due to intercooler leakage is more obvious during discharge. With intercooler leakage, the pressure inside the lower-pressure cylinder is significantly lower than that for the healthy condition. This consequently reduces the force acting to keep the 1st stage discharge valve closed: so that the discharge valve opens earlier than in the healthy condition. This causes a pressure drop compared to the healthy operation.

For the 2nd stage cylinder the presence of an intercooler leak causes the pressure over the entire cycle to be lower than for the healthy case which will result in a reduction in discharge efficiency. The lower pressure causes a delay in the opening of both the discharge and suction valves. The induced fault causes a clear variation in cylinder pressure which offers the possibility of monitoring for a leak in the intercooler by studying the pressure traces.
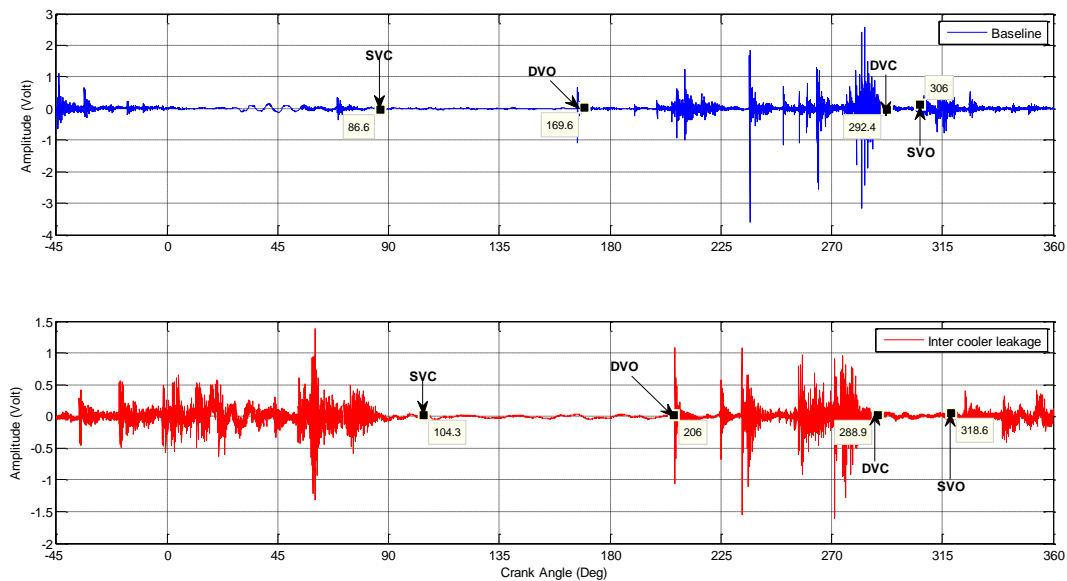
Figure 4.24 Time domain representation of vibration signal from accelerometer on the 2nd stage cylinder with intercooler leakage

Comparison of the amplitude of the vibration for a healthy compressor and one with a leaky intercooler can be seen in Figure 4.24. The amplitude for faulty operation is clearly less than that for healthy operation. Both the discharge and suction valves with intercooler leakage are delayed for 2nd stage opening and closing impact. At the discharge pressure 100 psi a healthy suction valve opens at 306.0º; for a faulty intercooler it is at 318.6º. The effects of an intercooler leakage on the 2nd stage discharge valve are shown in Figure 4.24. The valve opens later due to the slightly lower than normal pressure in the cylinder. A healthy discharge valve opens at 169.6° and for an intercooler leakage the valve opens at 206.0º.

4.9.2.4   *Loose Transmission Belt*

The transmission belt of the compressor undertakes a highly fluctuating load in longitudinal direction. The load amplitude will be more than 3 times higher during the transient start-up process. It is likely that the amplitude during acceleration will exceed the amplitude of the friction force between the wheel and the belt; hence the slip will occur in this duration. The slip will speed up the belt wearing. In addition, the friction processes are also factors which cause belt deterioration. The belt will gradually become worn in surface and deteriorated internally. These will lead to a loose belt. Figure 4.25 shows the predicted opening and closing discharge and suction valves, with and without drive belt. The loose transmission belt causes a distinct delay in the opening of the 2nd stage valves (both the discharge and the suction).

To simulate this belt fault, a loose belt was introduced by reducing the distance between the two drive pulleys from 169mm to 167mm (equivalent to a proportional change of belt length of about 0.5%) to determine whether some corresponding changes as can be seen in Figure 4.26.
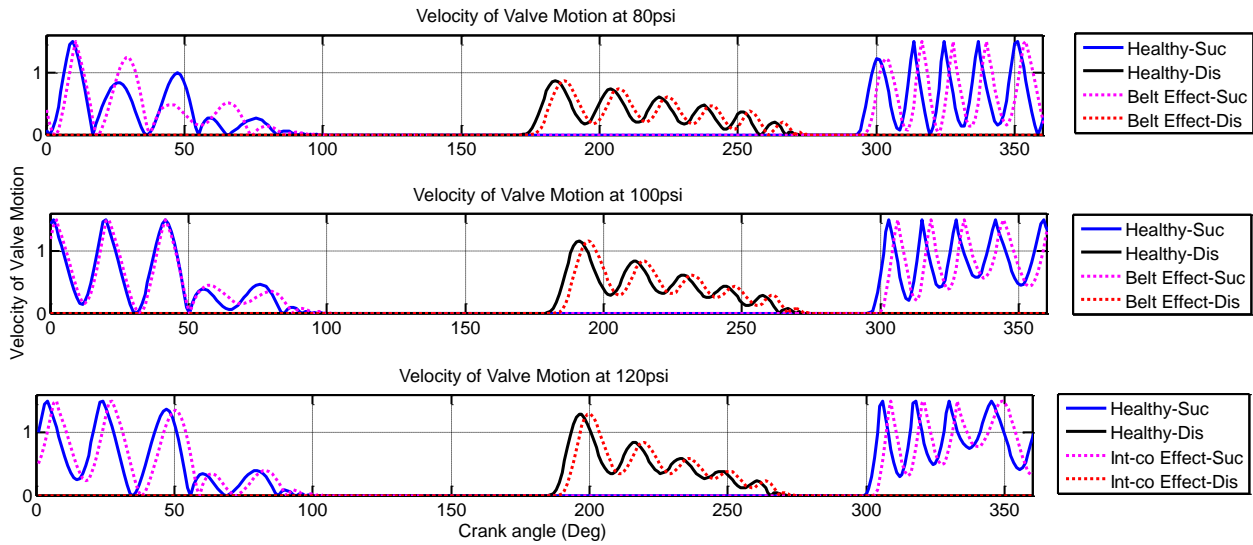
Figure 4.25 Predicted motion of suction and discharge valves for the 2nd stage with belt loosened

The measured vibration of the 2nd stage cylinders with and without a loose drive belt is shown in Figure 4.26. In the 2nd stage, the loose belt also causes the opening and closing of both the discharge and suction valves to be delayed. This delay leads to that the instantaneous pressure is less with a loose belt than for the healthy condition during the compression, while the pressure is higher during the expansion process. This delay is produced by a slippery belt and the effect is a slower compressor. A healthy discharge valve opens at 169.6° and, for a loosened belt the valve opens at 170.3° and, the suction valve will open later at 306.6° and closes at 90.45°, while in the healthy state it opens at 306.0° and closes at 103.9°.
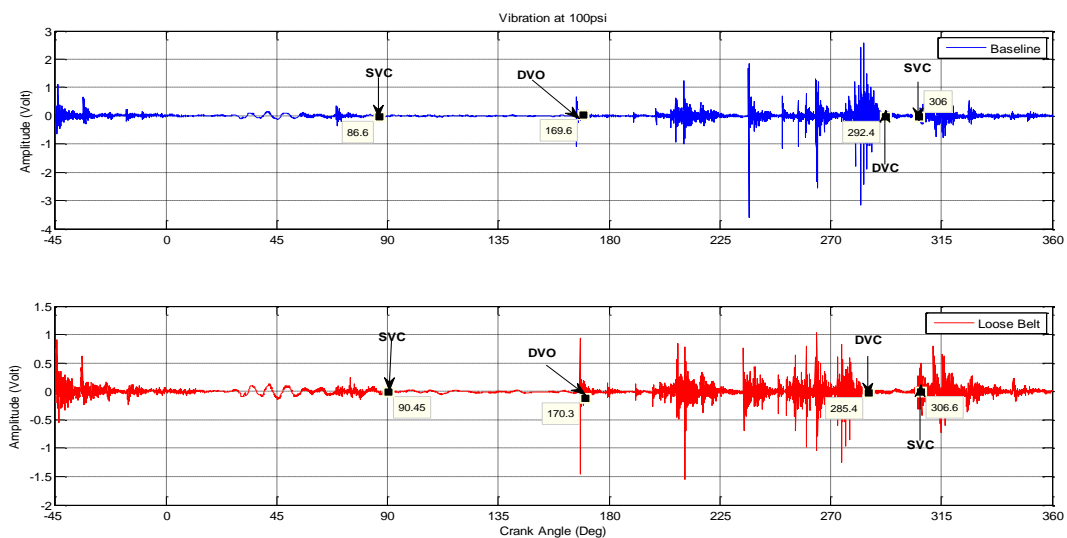
Figure 4.26 Time domain representation of vibration signal from accelerometer on the 2nd stage cylinder with belt loosens

4.10 Summary

In this chapter, the mathematical model of a RC has been developed for simulation of the working of a healthy operation and with different cases of predicted faults.

Furthermore, the equations required for mechanical dynamic system modelling were introduced and solved numerically in a MATLAB environment to predict cylinder pressures and valve motion for both the pressure and vibration.

In addition, the model was used to predict fault signatures for suction valve leakage, discharge valve leakage, intercooler leakage and a loose drive belt. These predicted results were compared with the equivalent measurements taken during the experimental tests.

Chapter 5

Detection and Diagnostics of A compressor With Different Faults Using Conventional Vibration Signals

This chapter presents the use of vibration measurement techniques for the detection of specific faults in the RC. These different operating conditions were recorded from the sensor mounted on the head of the second stage cylinder. These signals were analysed using conventional methods made in applications of time domain, frequency domain and envelope spectrum.

5.1   Introduction

The objective in this chapter is to use vibration monitoring techniques to detect faults in the valves of a RC in their early stages, and to use this information to protract the life of the compressor and protect the system from emergency shutdown. The vibration signal measured on the cylinder head of a compressor is, essentially, the combination of the responses to two main types of vibration excitation: flow induced vibration such as occurs when airflow interacts with valves or other parts of the system to cause periodic oscillations in the flow; and vibration due to, e.g., the valve plate hitting its seat when opening or closing.

The signals were generated from the vibration analysed using conventional methods in the time, frequency domains and envelope spectrum to obtain a set of effective features for detecting and diagnosis the seeded air compressor faults. The time domain analysis leads to popular statistical feature parameters such as RMS, kurtosis, peak factor and skewness. The frequency domain analysis (standard Fourier Transform) analysis produces features including amplitudes at frequencies. In the same way the envelope spectrum representations of the vibration data are also obtained features. The set of these features is evaluated in the separation of faults under different conditions.

5.2   Time Domain Analysis: Waveform Analysis

In spectral based CM, uneven signal strength indicates the severity of a fault or an abnormal condition while the corresponding vibration frequency points toward the source of a problem. Each point located on the spectral line gives a particular amplitude value and corresponding frequency location along the frequency axis. In normal conditions the vibration signals only contain the multiples of peak frequency, yet the presence of excessive amplitude or multiple harmonics undoubtedly implies a fault condition.

Figure 5.1 shows the time domain of the output signal from an accelerometer attached to a RC for different fault cases. In these cases the faults introduced were substantial and the differences in vibration signals are easily observed. However, difficulties with vibration monitoring can occur due to the presence of multiple vibration sources within the machine. The signals can combine in non-linear and possibly non-stationary ways. Also, many small changes could not be seen in the waveforms. Nevertheless, these vibration signals contain rich information of compressor health conditions.
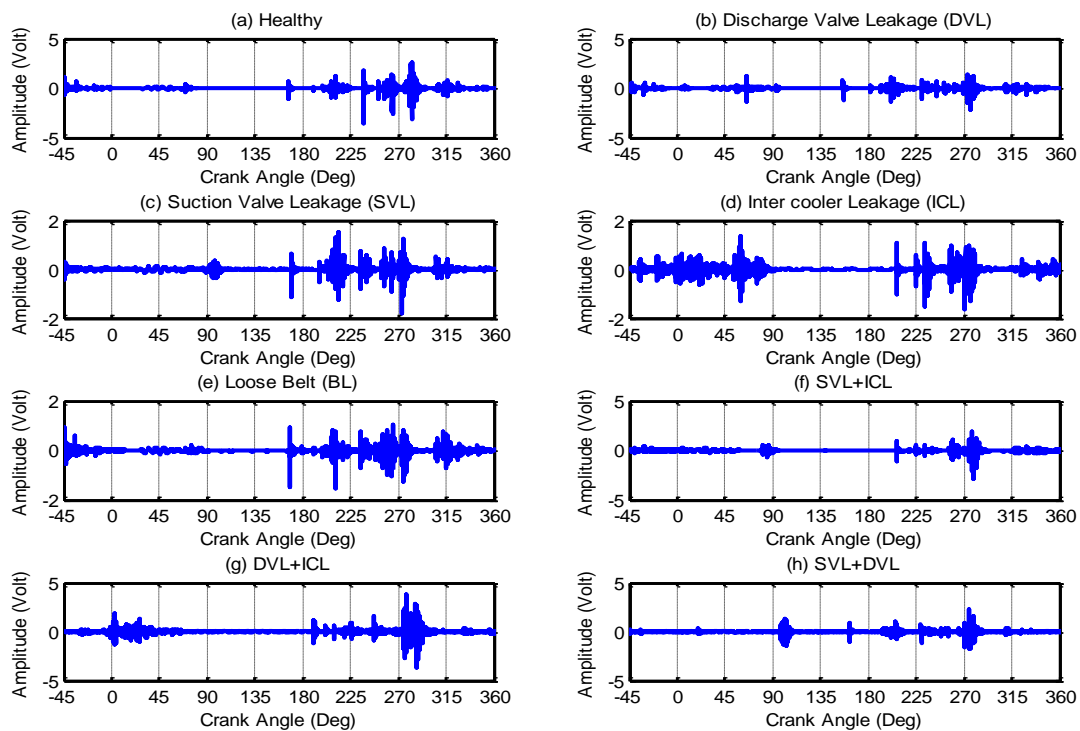
Figure 5.1 Time domain representation of vibration signal from an accelerometer on the 2nd stage cylinder head with different conditions at discharge pressure 100 psi

5.3   Features Extraction

Many possible features can be extracted from vibration signals for fault detection and diagnosis. This study explores the features derived from the time domain, frequency domain and envelope analysis, which are the most commonly used in CM.

The features extracted from raw vibration signals are the statistical measures including root mean square (RMS), peak factor, lower bound, upper bound, entropy, variance, skewness, kurtosis, maximum value and range which are given in Equations 2.3 to 2.16.

The other data set called the Fast Fourier Transform (FFT) was used to transform the vibration signal into the frequency domain from which the spectral features were obtained. Rather than using the spectrum from the raw data, an envelope spectrum is based for feature selection. As shown in Figure 5.2 envelope spectra for different cases exhibit clearly a number of discrete

components mainly from compressor working frequency 7.3Hz and its harmonics, up to 30 orders; in contrast, the spectrum from raw data show continuous spectral features which makes it more difficult to select a small number of feature components. Nevertheless, it can be seen in the envelope spectra that the amplitudes vary slightly but significantly between the different cases of faults. However, it is still not so easy to find a simple set of features to separate the fault cases completely. Thus the amplitudes of these components were all taken as candidate features and different harmonics were used for each trial run. Thus, the resultant feature dataset is a (n × s) matrix (n is the number of harmonics and s number of samples). The example has presented 30 numbers of harmonics and 24 numbers of samples with a total of (30 × 192) for eight different cases.
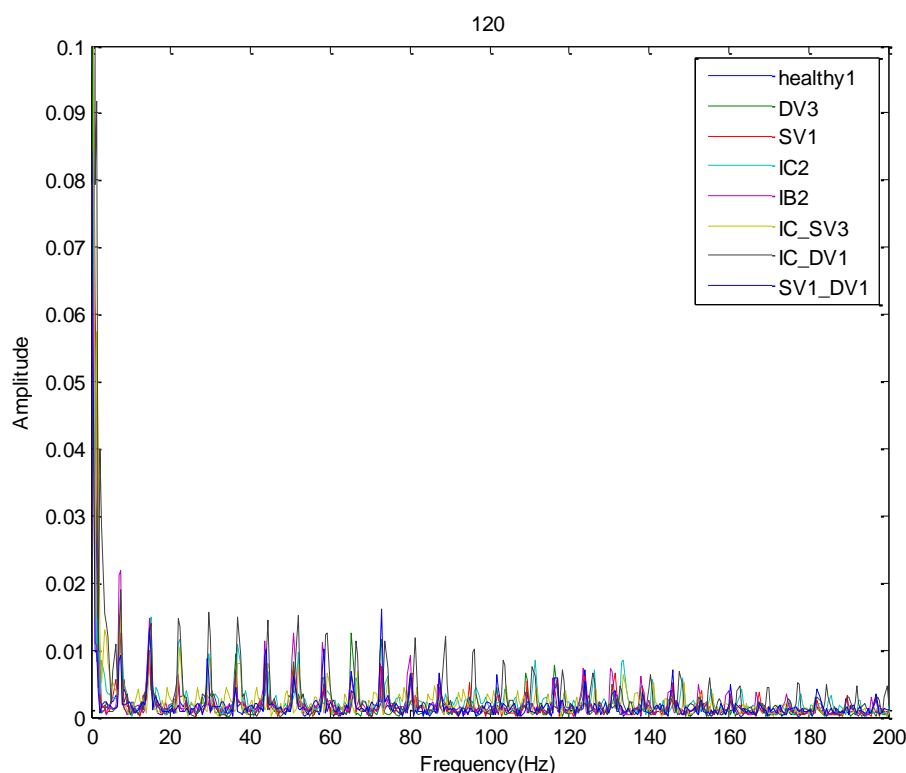
Figure 5.2 Envelope spectra of compressor vibration for healthy case and seven seeded faults

5.3.1 Signal Strength Based Diagnosis (RMS)

Figure 5.3 shows the RMS values of vibration signals at discharge pressure of 80 psi (0.55MPa) under single fault conditions such as discharge valve leakage (DVL) fault, suction valve leakage (SVL) fault, intercooler leakage (Inr-L) fault and belt loosens (BL) fault. The RMS values were calculated from the four segments of the waveform at different pressures. The RMS value of a given spectral signature is characterised as a uniform second central moment of the signal and it does not provide information about any sudden short duration isolated peaks in the signal. However as far as an overall vibration level of system is concerned, RMS is a good descriptive tool and can be used to have an insight about the system. In Figure 5.3 the cases with (DVL, SVL, BL and SVL+DVL), the RMS values for healthy and fault signals for various faults (as

mentioned above) have approximately similar patterns and with respect to magnitude or symmetry of signal, there is no significant difference between them. The only difference between fault and healthy (baseline) signals are presence of very small peaks at various frequencies along the frequency (crankshaft revolution) axis. In Figure 5.3 (DVL, SVL, BL and SVL+DVL), the RMS values both for healthy and fault signals are approximately within 0.1 ± 10%. In Figure 5.3 (Inr-L and Inr-L+ DVL) combine the multiple faults condition; the RMS value for the fault signal shows a significant increase in its strength and it almost has twice the value of Figure 5.3 (DVL, SVL, BL and SVL+DVL). It can be seen that most data points exceed the healthy signals with large deviation amplitude that implies a fault condition. Yet as an increase in RMS value of the fault signal only indicates an increase in overall magnitude of spectral line and vice versa, hence RMS analysis lacks clear information about the source of the fault. This argument is true and justified in multiple faults conditions Figure 5.3(Inr-L+SVL, Inr-L+DVL and SVL+DVL), as even though the RMS values of fault signals indicate the presence of fault, yet from the signal we cannot infer the source of the problem. The fault signal spectral line for multiple faults of intercooler and suction valve leakages shown in Figure 5.3 (Inr-L+DVL) has quite different behaviour both in terms of RMS value and the symmetry of signal. The signal contains a number of peaks having RMS values ranging from 0.175- 0.20. The significant increase in RMS value corresponds to major out-of-balance condition due to accumulation of faults in the compressor system.
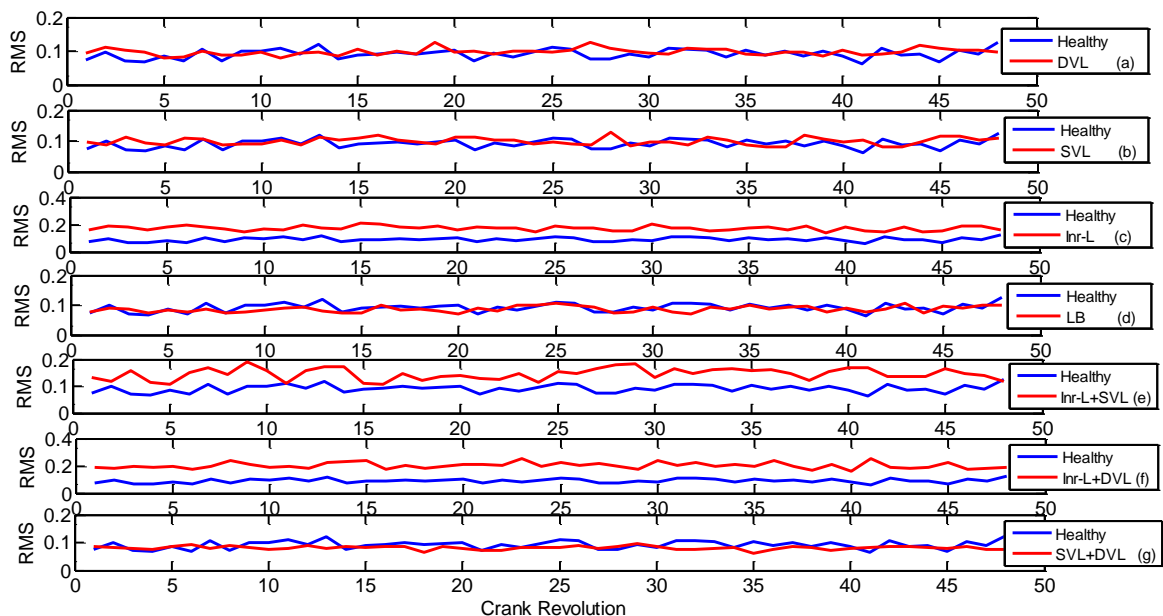
Figure 5.3 Typical RMS values of vibration RC at discharge pressure 80psi (0.55MPa)

Figure 5.4 shows the RMS values of vibration RC at discharge pressure of 120psi (0.83MPa). Here also we approximately have similar observations as discussed for Figure 5.3. It has been experimentally proven that the RMS value of vibration signals is directly proportional to the speed and load conditions of a machine. Since in Figure 5.3 and Figure 5.4, we only change

the discharged pressure conditions, hence there is no significant change in both figures and similar analyses of the results are applied here too for Figure 5.4.
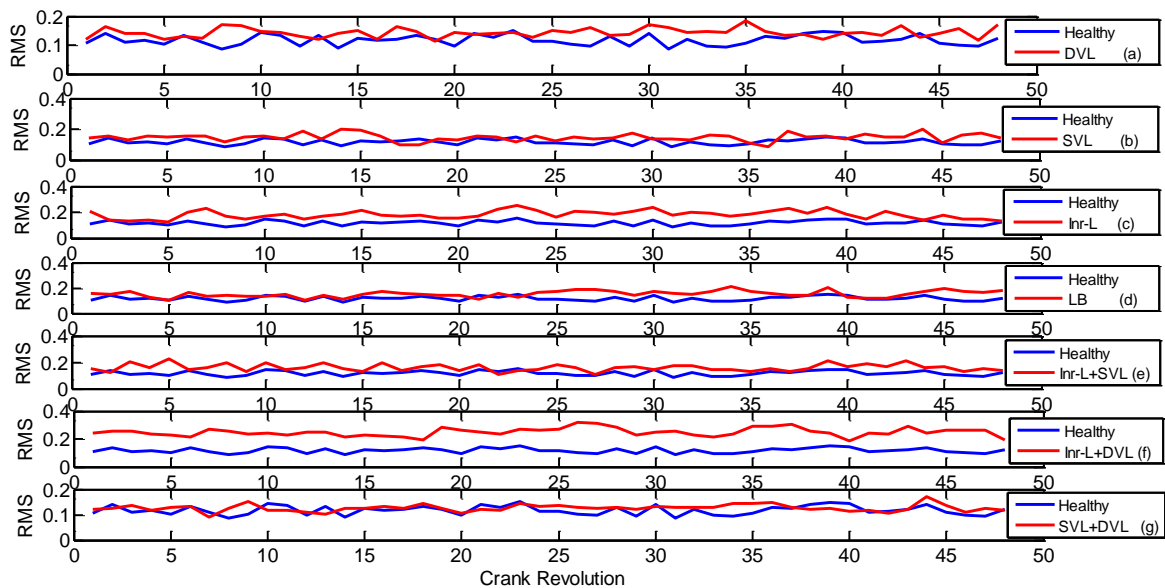
Figure 5.4 Typical RMS values of vibration RC at discharge pressure 120psi (0.83MPa)

5.3.2 Signal Structure Based Diagnosis (Kurtosis)

The kurtosis value is also a frequently used statistical parameter in CM. Compared to normal signal distribution, the kurtosis value of a signal is a measure of relative peakness or flatness of a distribution of signals on a spectral line. It is the 4 order statistical moment of the vibration signal, and as the signal gains the fourth power value, kurtosis effectively amplifies the isolated peaks in the signal. The kurtosis based diagnosis is used to measure the size of peaks and tails of distribution and it is a good indicator of the major peaks in a set of data. In case of a fault condition as the vibration levels increase the Kurtosis feature (relative peak or flatness) of a fault signal also increases and becomes prominent in spectral lines. Additionally with an increase in the kurtosis value the signal's peaks become more sharper with longer tail values while for lower kurtosis values, the peaks become more rounded in shape (Al-Arbi, 2012). Figure 5.5 and Figure 5.6 shows the kurtosis data for vibration signals for a RC at discharge pressure 80psi (0.55MPa) and 120psi (0.83MPa) respectively, under different fault conditions (as mentioned in previous section). From Figure 5.5 and Figure 5.6, it is obvious that both base line (healthy) and fault signals have isolated peaks.
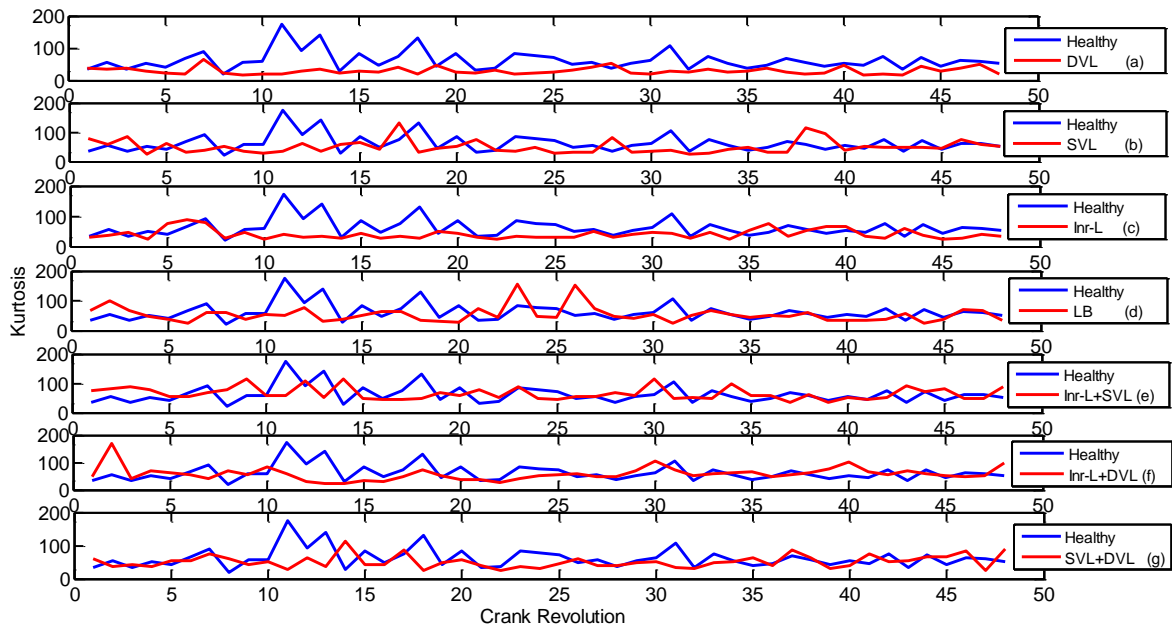
Figure 5.5 Typical Kurtosis values of vibration RC at discharge pressure 80psi (0.55MPa)

These isolated peaks are generally attributed to opposite phase angle along the frequency (crankshaft revolution) axis, i.e. where the healthy signal exhibits peakness behaviour, the fault signal becomes flat and vice versa. This behaviour is exactly in accordance with the Kurtosis feature as mentioned previously. Hence as DVL, SVL, Inr-L or combined fault conditions develop the relative peak or flatness in fault spectral signatures increases. In Figure 5.6 (DVL, SVL, Inr-L, LB, Inr-L+SVL, Inr-L+DVL and SVL+DVL) for various fault conditions) the flatness of the signal increase as fault conditions develop. Again the Kurtosis based diagnosis of a fault only gives transient behaviour of a system yet it does give information about the source of problem (Girondin, Loudahi, Pekpe, & Cassar, 2012).
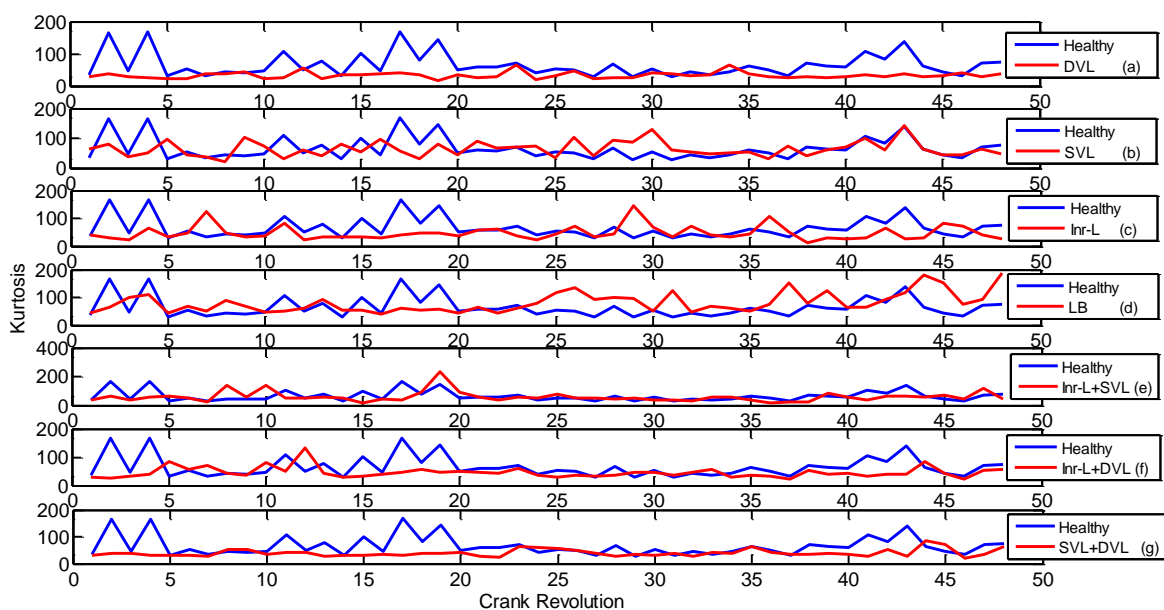
Figure 5.6 Typical Kurtosis values of vibration RC at discharge pressure 120psi (0.83MPa)

5.3.3 Signal Strength Based Diagnosis (Peak Factor)

The peak factor or crest factor is another parameter that can be used to have an insight of a system through vibration spectrum. The peak or crest factor of a spectral signal is a ratio of the peak level of the input signal to the RMS level. In contrast to RMS or kurtosis values that give an overall behaviour of a system, peak/crest factor is a useful tool to detect the discrete impulses above the background signal due to an impulsive vibration sources. From definition of peak/crest, it is obvious that as the peak levels of input signal increase, the crest factor also boosts up. The increase in peak level is generally attributed to an increase in vibration as the fault/damage grows. From Figure 5.7 and Figure 5.8 for various fault conditions such as DVL, SVL, Inr-L and BL faults, as it is obvious that crest levels of both healthy and fault signals approximately have same symmetry and similar magnitude, except few small peaks at few points along the spectral lines.
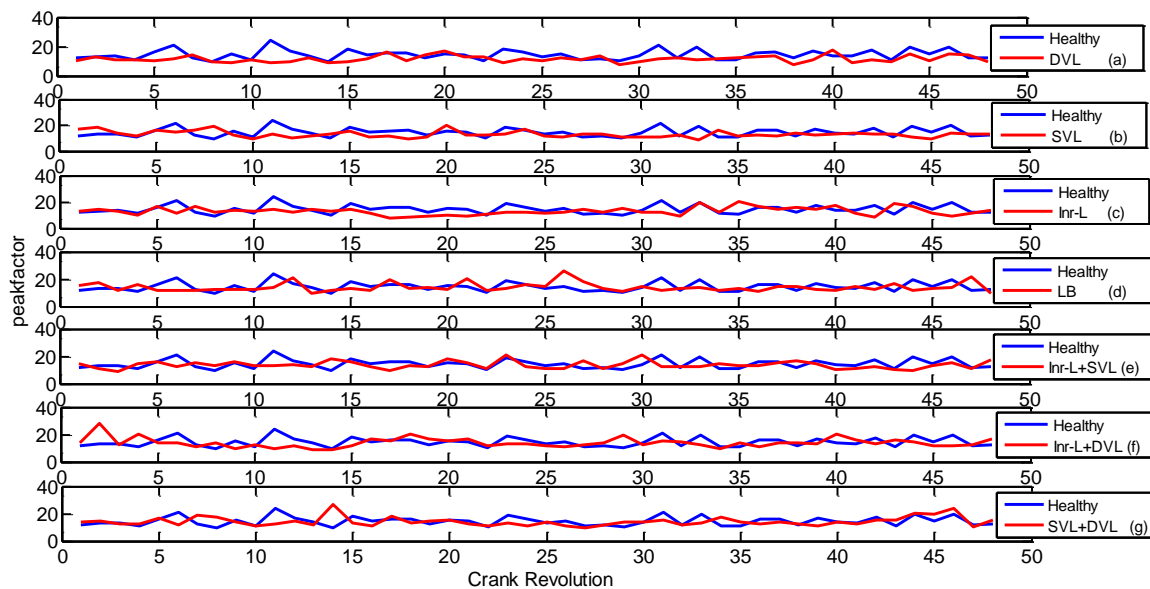
Figure 5.7 Typical Peak factor of vibration RC at discharge pressure 80psi (0.55MPa)
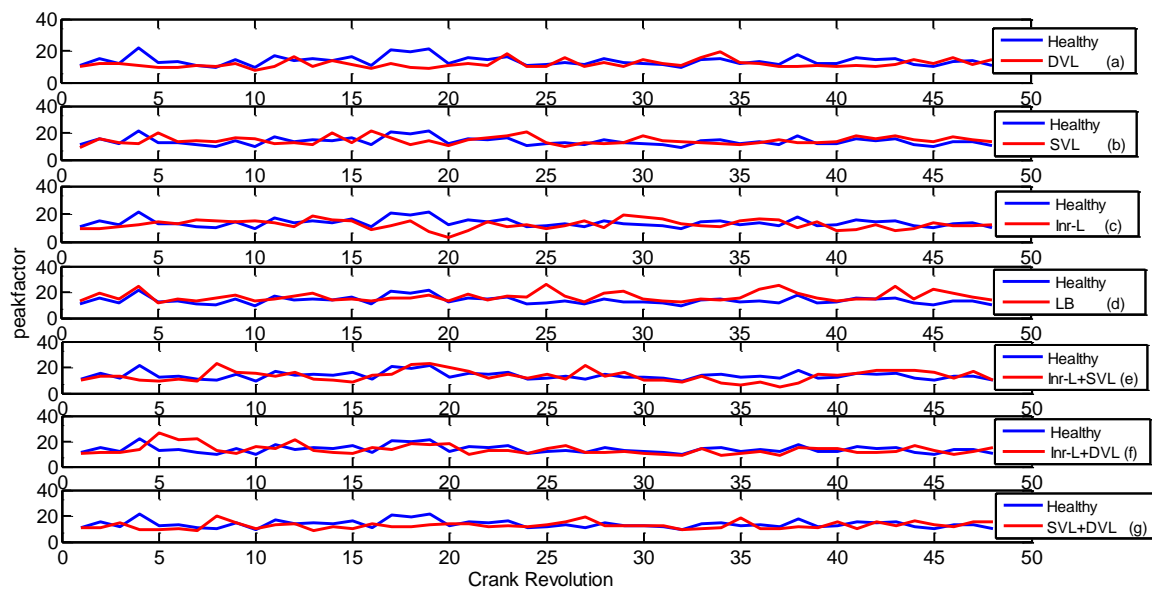
Figure 5.8 Typical Peak factor of vibration RC at discharge pressure 120psi (0.83MPa)

In addition to that, by changing the discharge pressure values for RCs, there is not any significant change in symmetry of fault and the healthy signal occurred. Only a slight decrease in magnitude for the fault signal has been observed.

5.3.4 Signal Structure Based Diagnosis (Skewness)

Similar to RMS, kurtosis and crest factor, skewness is also a statistical parameter that is used to analyse the fault condition of a machine exhibited through spectral signatures. From Figure 5.9 and Figure 5.10 it is obvious that both healthy and fault conditions (single or multiple faults conditions as discussed in previous sections) signals have very nominal skewing or tailing behaviour along the spectral lines. Both healthy and fault signals exhibit small and medium size peaks with positive skewness (i.e. tailed right) randomly along the spectral lines. Studies have shown that the presence of skewness in a spectral line corresponds to the buried fault related features and they increase as vibration increases (Liu, Wang, Golnaraghi, & Liu, 2008). Same analogy is true for spectral signatures given in Figure 5.9 and Figure 5.10. The long tailing feature is either in positive or negative direction, corresponding to the buried fault condition in compressor. Unfortunately, like most of other statistical parameters (RMS, Kurtosis, Peak/Crest etc.) skewness analysis also lacks information about the source of problem.
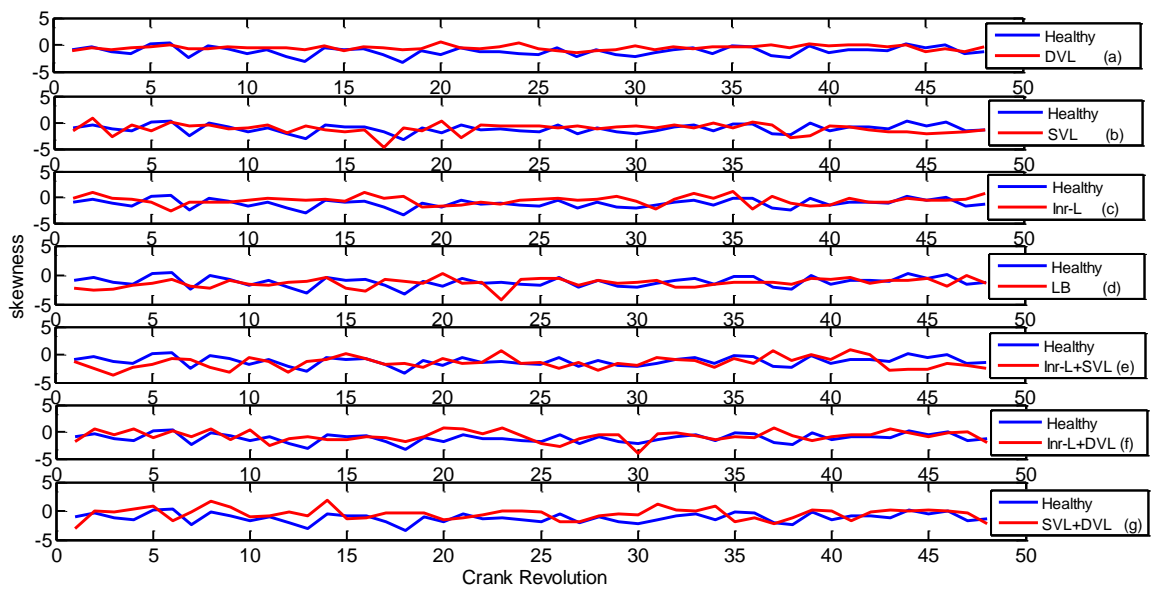
Figure 5.9 Typical Skewness of vibration RC at discharge pressure 80psi (0.55MPa)
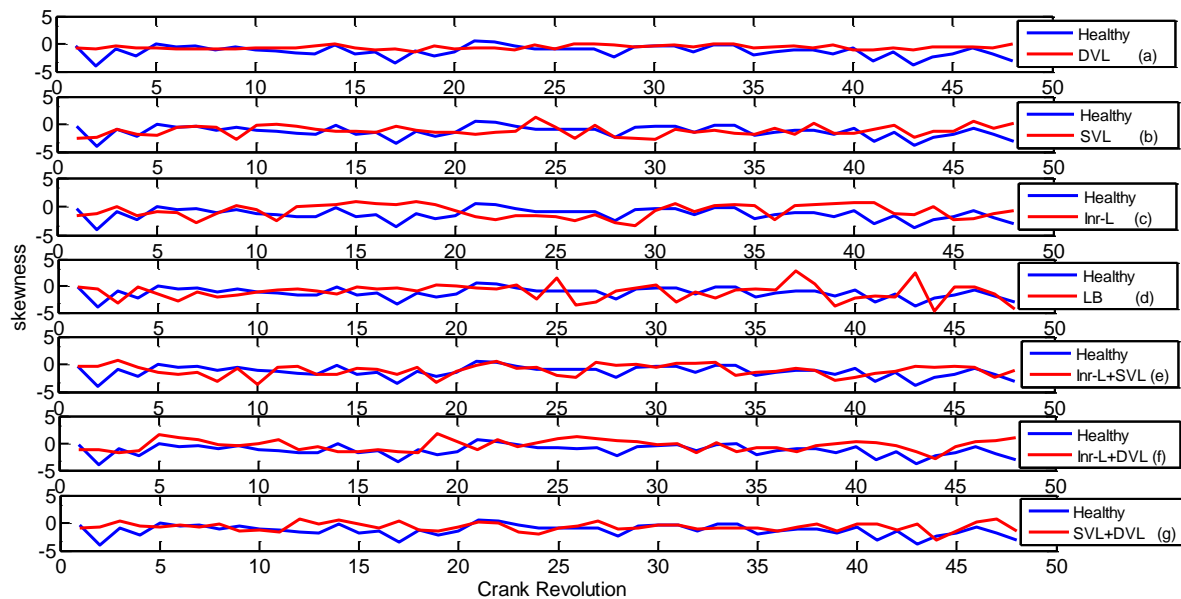
Figure 5.10 Typical Skewness of vibration RC at discharge pressure 120psi (0.83MPa)

5.4 Analysis and Results of Frequency Domain (Spectrum)

The spectral signatures of a compressor provide very useful information both in terms of frequency and fault conditions. The presence of a fault condition is directly linked to a particular frequency component and exhibited by a particular frequency signature. In Figure 5.11 to Figure 5.14, for a healthy condition and various single and multiple faults (as discussed in previous sections) spectral signatures, we observe two types of spectral features i.e.; low frequency harmonics and higher frequency harmonics. From Figure 5.11 and Figure 5.13 for fault conditions DVL, SVL and LB at 80psi and 120 psi discharge pressure, as it is

obvious that most of the peaks, both for healthy and fault signals are very similar in symmetry and have the same peak amplitudes with few exceptions (such as 10th, 17th, 18th and 23rd) as shown in Figure 5.11 (DVL). Inr-L conditions (both at 80 psi and 120 psi), a shift in behaviour of fault spectral line, both in terms of symmetry and peak amplitude is observed. In Figure 5.11 (Inr-L) between 120-140 Hz frequencies very strong peaks are observed for fault signal along with a small increase in overall magnitude of spectral line along the frequency axis.   Similar shift with fewer peak values has been observed for intercooler leakage fault condition at 120 psi discharge pressure as shown in Figure 5.13 (Inr-L). For multiple faults conditions (at 80 psi and 120 psi discharge Figure 5.12 and Figure 5.14) there are mixed spectral patterns. In frequency ranges 20-120 Hz and 150-200Hz, both healthy and fault spectral signals have approximately similar symmetry and amplitude, while between 10-20Hz and 120-150 Hz high peaks harmonics associated with intercooler fault are observed. The presence of low frequency harmonics in a compressor are generally associated to the shaft imbalance, misalignment, eccentricity cracks or bends in the shaft or similar faults conditions (Zhang, Jiang, Flatley, & Hill, 2001).
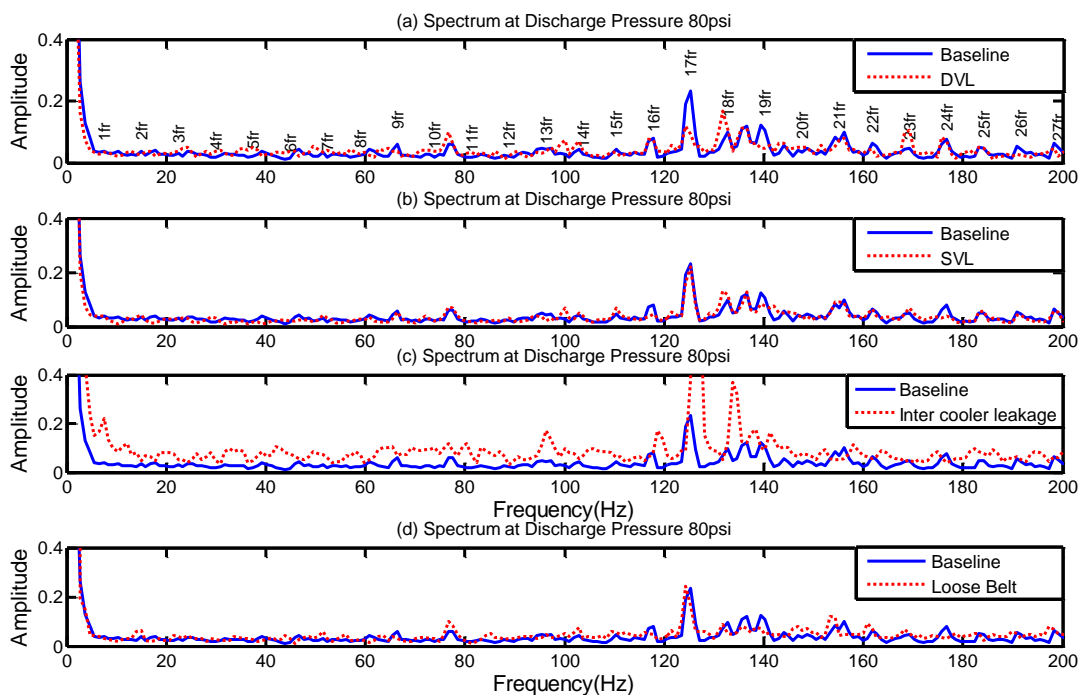
Figure 5.4 Comparison of spectra between healthy and four faults under discharge pressure 80psi
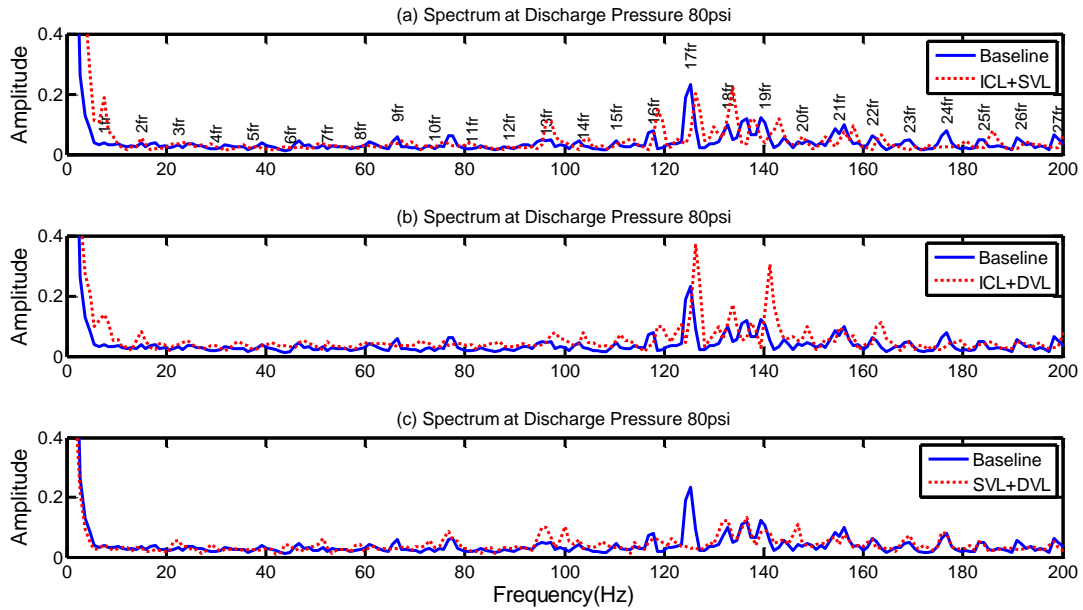
Figure 5.5 Comparison of spectra between healthy and three combined faults under discharge pressure 80psi
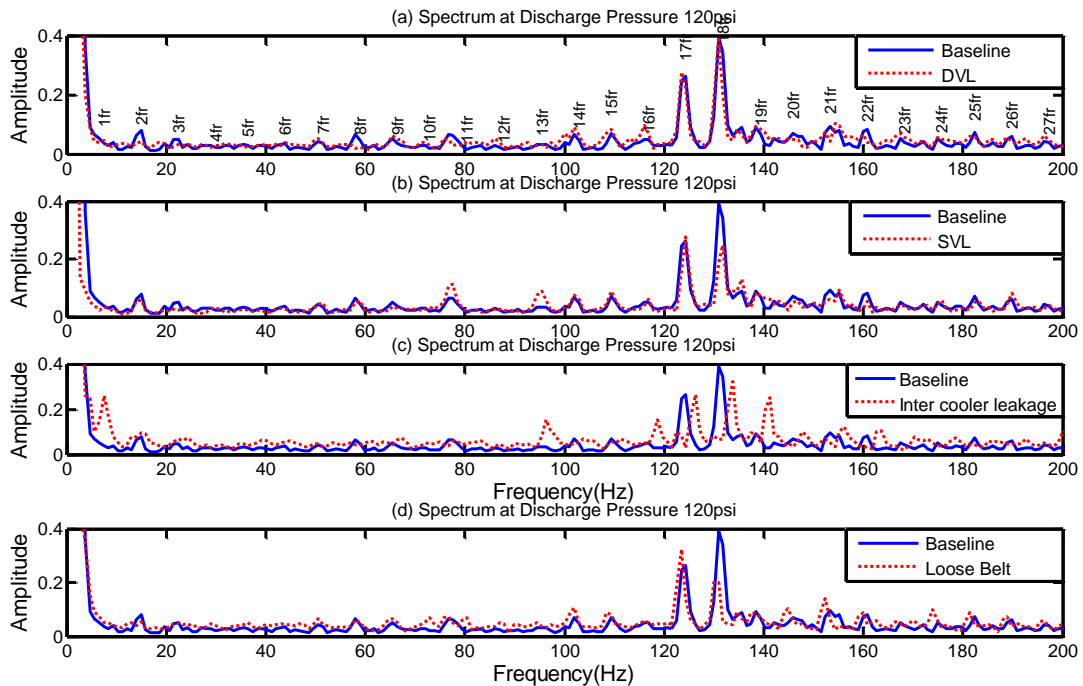
Figure 5.6 Comparison of spectra between healthy and four faults under discharge pressure 120psi
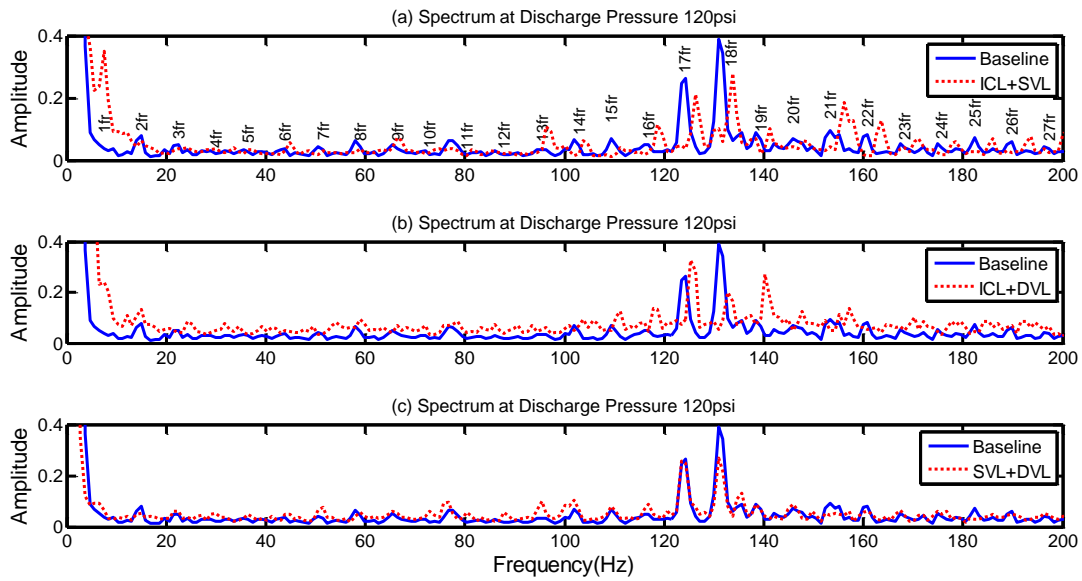
Figure 5.7 Comparison of spectra between healthy and three combined faults under discharge pressure 120psi

5.5   Analysis and Results of Envelope Spectrum

The Envelope spectrum analysis with proper window selection is a well-known signal processing technique for fault detection. In the following discussion, the windows of envelope spectra both for healthy and faulty spectral signals have been selected between 0-200 Hz frequency range at discharge pressure of pressure 80 psi and 120 psi as shown in Figure 5.15 to 5.18. Starting from Figure 5.15 (DVL) i.e. spectral signature for both healthy and DVL fault signal; a magnified image of spectral signature shows that there are 26 obvious harmonics. They are present in the selected window both for healthy and DVL fault spectral signatures. Comparing both healthy and faulty signatures, it is clear that there is a boost in amplitude of a few harmonics of fault spectral signature. These observations are summarized in Table 5.1. From Figure 5.15 (SVL, Int-L and BL) we have similar results as given in Table 5.2 to Table 5.4 respectively.

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	2, 3, 5	3	
From 40Hz to 80 Hz	8, 9, 10	3	
From 80 Hz-120 Hz	12	1	
From 120 Hz-160 Hz	17,18, 19, 20	4	
From 160 Hz-200 Hz	22,23	2	Total= 13

Table 5.1 The frequency domain with higher amplitude harmonics in DVL

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	2, 4	2	
From 40Hz to 80 Hz	6, 7, 8, 10	4	
From 80 Hz-120 Hz	15	1	
From 120 Hz-160 Hz	17, 18, 19, 21	4	
From 160 Hz-200 Hz	25	1	Total= 12

Table 5.2 The frequency domain with higher amplitude harmonics in SVL

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	1, 2, 3, 4	4	
From 40Hz to 80 Hz	6, 7, 8,10	4	
From 80 Hz-120 Hz	11, 12, 13, 14, 16	5	
From 120 Hz-160 Hz	17, 18, 19, 20	4	
From 160 Hz-200 Hz	22, 23, 24	3	Total= 20

Table 5.3 The frequency domain with higher amplitude harmonics in Inr-L

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	1, 2, 3, 4, 5	5	
From 40Hz to 80 Hz	6, 7, 8	3	
From 80 Hz-120 Hz	11, 12, 13, 14, 16	5	
From 120 Hz-160 Hz	17, 18, 19, 20, 21	5	
From 160 Hz-200 Hz	22, 23, 24, 26	4	Total= 22

Table 5.4 The frequency domain with higher amplitude harmonics in BL

From the above data tables, it is shown than that compared to healthy signals signatures, there is a significant increase in amplitude and hence envelope of fault signals. From the data table it is clear that amplitude harmonics in a belt loosening fault has maximum increase in amplitude, while the envelope corresponding to the SVL fault harmonics has the least variation compared to healthy signals envelope. The difference in respective envelops corresponds to the buried fault conditions of the compressor.
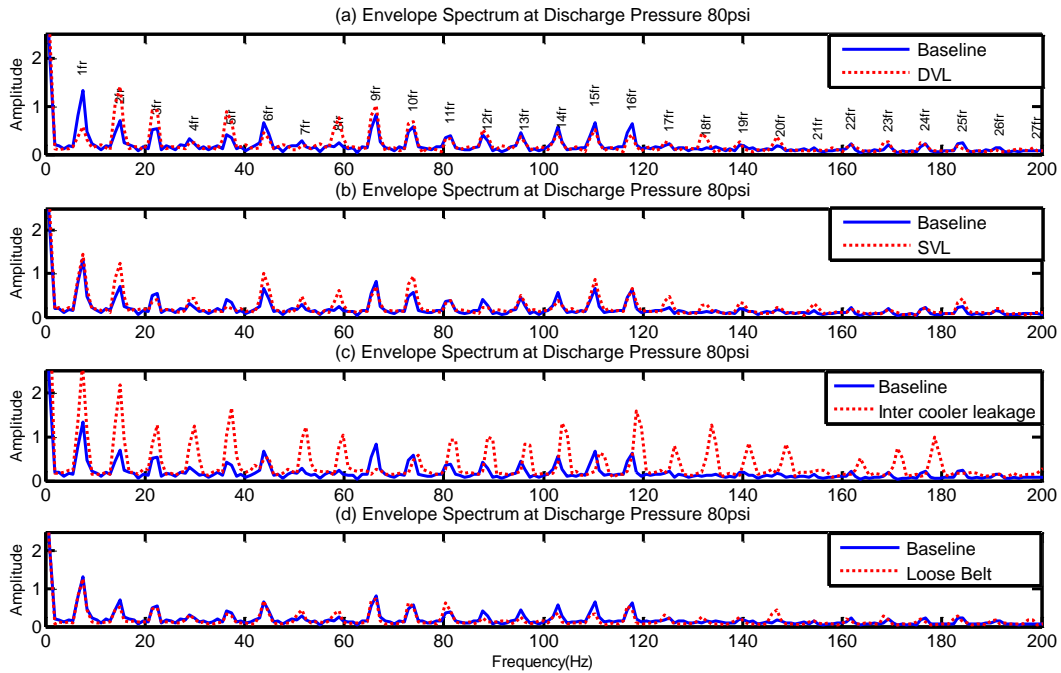
Figure 5.8 Comparison of envelope spectra between healthy and four faults under discharge pressure 80psi

In Figure 5.15 (DVL, SVL, Inr-L and BL) we have spectral signatures with a single fault. The next step is to analyse the envelope of spectral signature for combined multiple faults conditions as shown in Figure 5.16 (DVL, SVL and Inr-L). Similar to previous sections we have following data Tables 5.5 to Table 5.7 (given below) for accumulated are combined multiple faults condition. From the data table and spectral signatures we observe that compared to individual fault, the envelopes of accumulated faults have stronger variation. In contrast to healthy signals the multiple accumulated faults harmonics have very strong amplitude values (Figure 5.16 (DVL) and Figure 5.17 (SVL)).

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	1, 2, 3, 4, 5	5	
From 40Hz to 80 Hz	6, 7, 8	3	
From 80 Hz-120 Hz	11, 12, 13, 14, 15, 16	6	
From 120 Hz-160 Hz	17, 18, 19, 20, 21	5	
From 160 Hz-200 Hz	22, 23, 24, 26	4	Total=23

Table 5.5 The frequency domain with higher amplitude harmonics in accumulated Inr–L and SVL

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	1, 2, 3, 4, 5	5	
From 40Hz to 80 Hz	6, 7, 8, 9,10	5	
From 80 Hz-120 Hz	11, 12, 13	3	
From 120 Hz-160 Hz	17, 18, 19, 20, 21	5	
From 160 Hz-200 Hz	22, 23, 24, 25, 26	5	Total=23

Table 5.6 The frequency domain with higher amplitude harmonics in accumulated Inr-L and DVL

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	2, 4, 5	3	
From 40Hz to 80 Hz	6, 7, 8, 10	4	
From 80 Hz-120 Hz	12, 14	2	
From 120 Hz-160 Hz	17, 18, 19, 20, 21	5	
From 160 Hz-200 Hz	22, 23, 24, 25, 26, 27	6	Total=20

Table 5.7 The frequency domain with higher amplitude harmonics in accumulated SVL and DVL

Based on a similar argument, if we examine the envelopes of spectral signature at a discharge pressure of 120 psi both for individual (Figure 5.17 (DVL, SVL, Inr-L and BL)) accumulated faults conditions as shown in Figure 5.17 and Figure 5.18, we have similar results (as shown in tables 5.8 to Table 5.14) i.e. an increase in amplitude of fault signals and corresponding change in symmetry of the respective envelopes are observed.
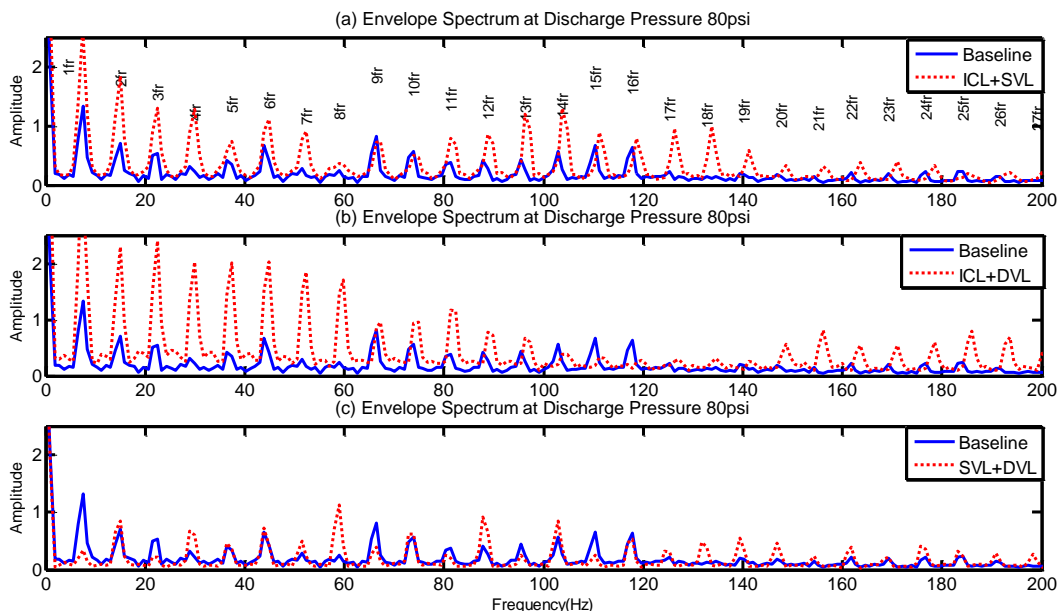
Figure 5.9 Comparison of envelope spectra between healthy and three combined faults under discharge pressure 80psi

In case of fault condition as moving parts strike, impacts are produced; these impacts modulate the associated impact areas frequency signals by adding energy to the signal that

results in an increase in amplitude and RMS value of the signal. Along with amplitude, the increase in amplitude modulated frequency signals also depend on the relative phase angle of modulated signals. In case of in phase fault signals, the modulated effect accumulates and vice versa. This argument justifies the symmetry of envelopes having very high peak harmonics as shown in Figure 5.16 (DVL, SVL) and Figure 5.18 (DVL, SVL)) and very low peak harmonics (Figure 5.16 (Inr-L) and Figure 5.17 (Inr-L)) values for accumulated fault conditions. The above discussion implies that an envelope spectrum is a good tool to measure a fault condition at a particular passing frequency as shown in fault spectral signatures of Figure 5.15 to Figure 5.18.

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	3, 6	2	
From 40Hz to 80 Hz	8, 9, 10, 11	4	
From 80 Hz-120 Hz	15, 16, 17	3	
From 120 Hz-160 Hz	19, 20	2	
From 160 Hz-200 Hz	22, 23	2	Total=13

Table 5.8 The frequency domain with higher amplitude harmonics in DVL at 120 psi

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	2, 4	2	
From 40Hz to 80 Hz	5, 6, 7, 9, 10	5	
From 80 Hz-120 Hz	11, 12, 13, 14, 15	5	
From 120 Hz-160 Hz	16, 17, 18, 19, 20	5	
From 160 Hz-200 Hz	21, 22, 23, 24, 25, 26	6	Total=23

Table 5.9 The frequency domain with higher amplitude harmonics in SVL at 120 psi

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	2, 3, 4, 5	4	
From 40Hz to 80 Hz	7, 9, 10	3	
From 80 Hz-120 Hz	11, 12, 13, 14, 15	5	
From 120 Hz-160 Hz	16, 17, 18, 19, 20	5	
From 160 Hz-200 Hz	21, 22, 23, 24, 25,	5	Total=22

Table 5.10 The frequency domain with higher amplitude harmonics in Inr-L at 120 psi

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	1, 2, 3, 4, 5, 7, 8 , 9, 10	9	
From 40Hz to 80 Hz	11, 12, 13, 14, 15, 16, 17, 18, 20	9	
From 80 Hz-120 Hz	21, 22, 23, 24, 25, 26, 27, 28, 29	9	
From 120 Hz-160 Hz	30, 31, 32, 33, 34, 35, 36, 37, 38, 39	10	
From 160 Hz-200 Hz	40, 41, 43, 45, 46, 48	6	Total=44

Table 5.11 The frequency domain with higher amplitude harmonics in LB at 120 psi

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	1, 2, 3, 4, 5, 7, 8 , 9, 10, 11	9	
From 40Hz to 80 Hz	11, 12, 13, 14, 15, 16, 17, 18, 21	9	
From 80 Hz-120 Hz	23, 24, 25, 26, 27, 28, 29, 30 31	9	
From 120 Hz-160 Hz	32, 33, 34, 35, 36, 37, 39, 41	8	
From 160 Hz-200 Hz	42, 44, 45, 46, 47, 49	6	Total=41

Table 5.12 The frequency domain with higher amplitude harmonics in Inr-L and SVL at 120 psi

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	1, 2, 3, 4, 5,	5	
From 40Hz to 80 Hz	6, 7, 8, 9, 10	5	
From 80 Hz-120 Hz	11, 12, 13, 14, 15, 16	6	
From 120 Hz-160 Hz	18, 19, 20, 21, 22	5	
From 160 Hz-200 Hz	23, 24, 25, 26, 27	6	Total=27

Table 5.13 The frequency domain with higher amplitude harmonics in Inr-L and DVL at 120 psi

Frequency range	Harmonics order	No of peaks	
Under 40 Hz	2, 4, 5	3	
From 40Hz to 80 Hz	6, 8, 10	3	
From 80 Hz-120 Hz	12, 13, 14	3	
From 120 Hz-160 Hz	17, 19, 20	3	
From 160 Hz-200 Hz	21, 23, 24, 26	4	Total=16

Table 5.14 The frequency domain with higher amplitude harmonics in Inr-L and SVL at 120 psi
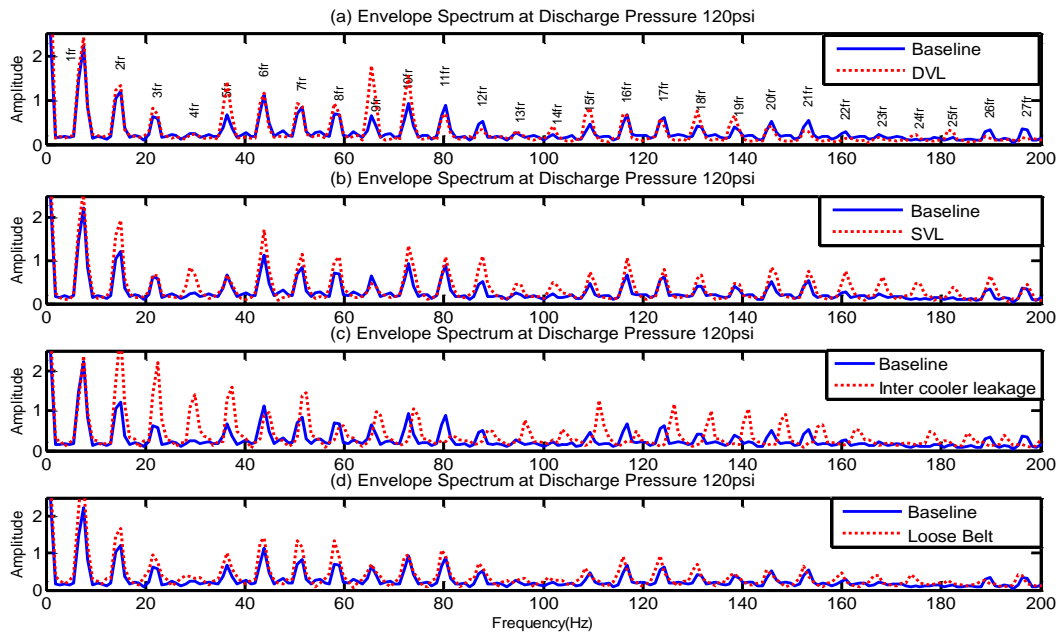
Figure 5.10 Comparison envelope spectrums for healthy and four different faults for RC at discharge pressure 120psi
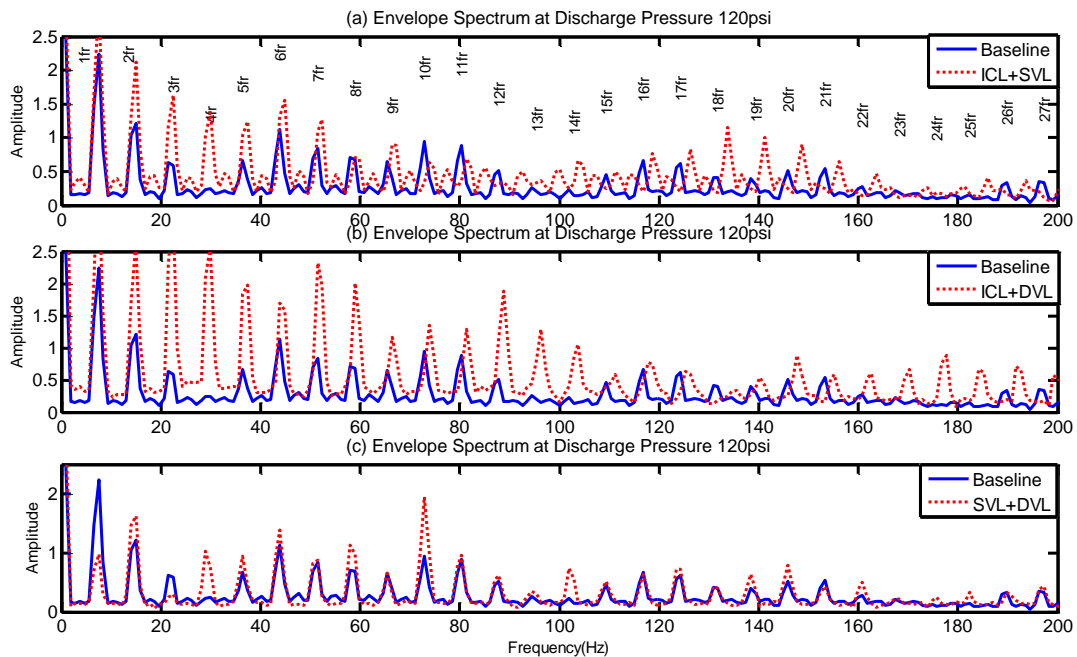
Figure 5.11 Comparison envelope spectrums for healthy and three combined faults for a RC at discharge pressure 120psi

5.6   Summary

The test results show that this envelope spectrum signal processing technique is an effective RC fault detection method, which is particularly useful for non-stationary feature extraction

and analysis. The presented envelope spectrum analysis technique in this chapter is to achieve the fourth research objective in Section 1.4. (To develop a more effective signal processing technique for feature extraction in RC fault detection).

Chapter 6

Condition Monitoring and Fault Diagnosis Using Artificial Intelligence Techniques and Principal Component Analysis

To extract effective CM information from non-stationary vibration signals that result from a number of nonlinear effects in RCs and noises, three typical AI based classification techniques: NNs, SVMs and RVMs are overviewed first. Then new approaches of combining them with key AI optimisation technique are introduced: GAs to enhance their classification performances in terms of effective feature selection, efficient model training and accurate classification. In addition, conventional statistical techniques are applied where PCA is investigated.

6.1   Introduction

Due to the complex dynamic mechanisms and wide operating conditions of the RCs, vibration signals exhibit highly nonlinear and nonstationary characteristics. In addition, there always exist various measurement noises in the measured vibration signals. Therefore, it is difficult to find effective techniques such spectrum analysis, envelop analysis and time domain statistics that can be used to achieve the condition of the compressor effectively, which motivate the search for new techniques.

AI techniques are useful for determining the relationship data with input variables, of applications where conventional method such as statistics are too complicated or not valid. Non-linear systems are an example of applications which can be evaluated by AI techniques, and which are difficult or impossible to analyse with conventional techniques.

The value of AI can be understood by comparing it with natural human intelligence as follows (Wang, 2003);

AI is more permanent, natural intelligence is perishable from a commercial standpoint since specialists leave their place of employment or forget information. AI, however, is permanent as long as the computer systems and programs remain unchanged.

AI offers ease of duplication and dissemination. Transferring knowledge from one person to another usually requires a long process of apprenticeship; even so, expertise can never be duplicated completely.

AI being a computer technology is consistent and thorough. Natural intelligence is erratic because people are unpredictable, they do not perform consistently.

AI can be documented. Decisions or conclusions made by a computer system can be more easily documented by tracing the activities of the system. Natural intelligence is difficult to reproduce, for example, a person may reach a conclusion but at some later date may be

unable to re-create the reasoning process that led to that conclusion or to even recall the assumption that were a part of the decision.

The CM and fault diagnosis of rotary machinery have moved in recent years from traditional techniques to AI techniques (Filippetti, Franceschini, Tassoni, & Vas, 2000). AI techniques, such as ES, FLS, ANN, GA and SVM have been applied in fault diagnosis of very complex time varying and non-linear system, where accurate mathematical models or conventional data analysis methods are difficult to be built. These techniques use association, reasoning and decision making processes as would the human brain in solving complicated diagnostic problems. AI techniques are good candidates for the automation of the rotary machinery diagnostic procedures (Awadallah & Morcos, 2003).

The general problem with AI techniques is that a large number of parameters have to be set adequately in order to obtain acceptable results. However, there are no clear rules how to set these parameters. Yet these parameters determine the success of the training and eventual AI applications.

In addition, AI techniques usually give good indication on the fault types but provide little explanation on the mechanisms of physical system. It means that the locations and natures of faults cannot be determined.

The contribution of this research is to improve the performance of AI approaches. This chapter investigates the development of a novel approaches that combine GA with PNN, SVM and RVM. The GA is an efficient tool to identify the large number of parameters associated with these techniques.

Different types of vibration features are examined using these novel combined approaches. The resulted optimal features are then correlated to different types of faults.

In addition, a conventional multivariate technique: PCA has also been investigated using the time domain vibration features for detection and diagnosis from a RC. Furthermore, the studies of Q-contributions and D-contribution have applied to determine suitable features which allow full classification of different simulated faults. The exploration of PCA based methods because they not previously been used for RC fault CM. Moreover the results from PCA will serve as the benchmarks for evaluating that of AI techniques.

6.2   Neural Networks for Fault Classification

The use of neural networks presents the possibility of reducing or altogether eliminating the need for complex mathematical systems that warrant the usage of large resources and time. Data is typically obtained for neural networks on a real time basis and can then be analysed using time series techniques. The resulting power spectral densities provide the amorphous basis that can be utilised in any given form for the application of neural network systems later. The power spectral densities are obtained through the application of Fast Fourier Transforms (FFTs). A back propagation method is often used in any neural network in order to allow machine learning to occur. This also allows the construction of a fault diagnosis system using

neural network methods alone (Asakura, Kobayasi, Hayashi, & Xu, 1997). Mindful that neural networks require only one major input from the machine i.e. vibration patterns but these patterns are often composed of wide ranging frequencies and the resulting vibration signals are also wide ended. This is also provided under Shannon's sampling theorem whereby vibration monitoring for fault diagnosis requires a consistently high sampling rate. In turn this means that sample sizes are often large and there is superfluous data available leading to high dimensionality. In order to deal with this situation, there is often a need for extensive pre-processing before actual ANNs are applied for fault diagnosis (Rafiee, Arvani, Harifi, & Sadeghi, 2007).

6.2.1 Theory of Neural Networks

A neural network can be seen akin to a biological neural network that interconnects neurons in order to form a complex learning and interaction mechanism. NNs employed for applications such as CM and fault diagnosis are pronounced as ANNs. The basic layout of the network is analogous to any biological neural network where there are distinct levels and layers of neurons between the inputs and the outputs. The number of layers and their arrangements of neurons as well as the inputs and outputs tend to determine the complexity of the overall neural network model. The basic contention behind a neural network is to employ a connectionist approach to solving real world problems by allowing the neural network to interact and learn as it develops. The overall structure of the neural network never remains consistent and tends to evolve as the neural network learns new things (Montavon, Orr, & Muller, 2012).

In the simplest terms, a neural network can be seen as the composition of various mathematical functions that can be defined from the input X to the output Y as:

$$f: X \rightarrow Y$$

In this case, the distribution could be over X or over both X and Y combined. In addition to the representation using functions, it must be taken to note that these functions may be subservient to rules that allow for learning. In addition, a neural network model is composed of a set of such functions that map the input to the output. These functions and the set of such functions could be acquired by varying the various inputs and outputs to see how the overall neural network model reacts. Moreover, the connection between various nodes could also be varied for investigation along with the number of nodes in addition to other parameters that may be exclusive to each neural network model (Iovine, 2012).

One of the more important parts of the ANN framework are the network itself that connects the various nodes / neurons together. Unlike other simpler networks, the elements / neurons in a neural network have weighted connections to each other than keep on changing as the neural network learns new things. The simplest neural network can be seen as being composed of three distinct layers that are connected by a network of neurons. The outermost layer is composed of neurons that accept inputs only and move them to the second layer of

neurons. Commonly, the second layer of neurons is also known as the hidden layer since the arrangement of this layer is not completely known. The second layer of neurons is followed by a third layer of neurons that act as the outputs of the neural network system. Diagrammatically, this situation can be represented as:
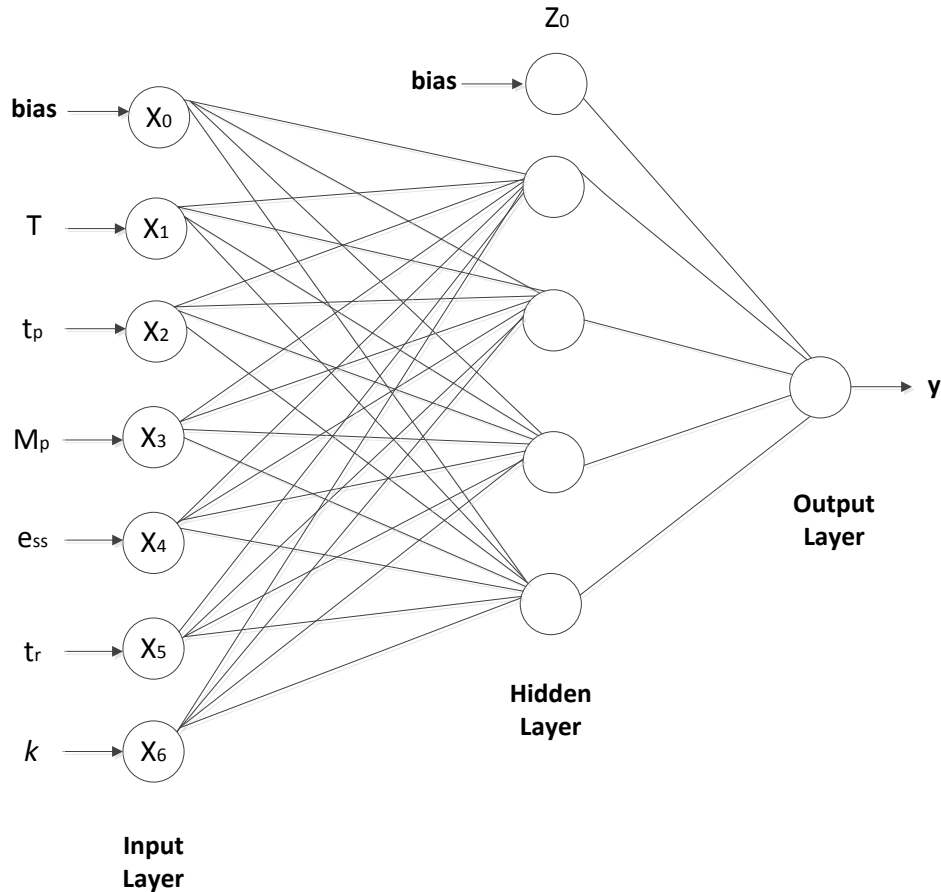
Figure 6.1 Representation of a typical simple neural network (X. He, Hua, Lin, & Liu, 2011)

It must be taken to note that when more complex neural networks are considered, the second layer shown in the diagram above would consist of many levels. The number of levels and their arrangement would in turn determine the overall complexity of the neural network being considered.

In order to describe neural networks in detail, mathematical functions are relied upon. Any neuron's network function tends to depend on other functions which in turn depend on other functions. The dependencies between various functions allow for the creation of a network structure. A typical neural network structure can be represented mathematically in the form:

$$f(x) = K(\sum w_i g_i(x))$$
Equation 6.1

Where:

x is the input variable

$f(x)$ is the output function

$g_i(x)$ is a function in the hidden layer that depends on other functions

$w_i(x)$ is the weight of $g_i(x)$

K is the activation function

In terms of complex neural networks, the function $g_i(x)$ could depend on the function $h_i(x)$ with its own weight(s) and $h_i(x)$ could depend on a number of other functions with different weights. Typically, neural networks can be either feed forward or recurrent depending on how the networks are arranged (Miller, 2012).

6.2.2 A Probabilistic Neural Network

One application of neural networks has been PNNs that are essentially feed forward neural networks that are composed of a pattern layer in addition to the other layers in a neural network. A PNN relies on inputs that have no pattern associated to themselves. Initial development in PNN came from Specht who developed a high efficiency PNN that was used to deal with classification problems (Specht, 1988). The basic contention of a PNN is to produce approximations as means of estimating the probability density function related to the various categories found in any classification problem. PNNs are derived from ANNs using the Bayes Decision Theory as a basis, which provides for a probabilistic approach to emerge. It is common to utilised supervised training with PNNs in order to aid the development of probability density functions that are used in the pattern layer. The utilisation of PNNs offers the distinct advantages of rapid training, an innate parallel arrangement that provides for simpler convergence as well as inherent flexibility that allows the removal and addition of training samples without a need for wide ranging retraining efforts (Gorunescu, Gorunescu, El-Darzi, Gorunescu, & K., 2005).

The input layer of a PNN is followed by the pattern layer and then the category layer. The PNN relies on the difference between the input vector and the training input in order to compute the closeness of the input vector to the training input vector. On the other hand, the second layer of the PNN tends to add together the various inputs to produce the probabilities of each connection taking place. The produced probabilities are then screened through a transfer function that tends to classify the greatest probability from the rest. The highest probability is then assigned "1" while the rest of the probabilities are assigned "0" to identify what possibility is likelier to occur than the others.

6.2.3 PNN Architecture and Theory of Operation

The PNN is essentially a feed forward type of ANN. The simplest and basic implementation of a PNN bases itself on nonparametric probability density functions as well as Bayes' classification rule. The model method of training for PNNs is based on a single pass through all of the provided training data sets. The faster movement through all of the available data sets tends to make the PNN faster to train compared to other neural network possibilities although it does require more memory to accommodate the larger set of training patterns and their data sets. The evolution of cheaper and more effective computer memory over the decades has meant that larger data storage is of little concern when implementing PNNs nowadays (Tripathy,

Maheshwar, & Verma, 2010). PNNs are four layered feed-forward neural networks that are designed to be capable of estimating the most suitable classifier. It is common to utilise a Gaussian activation function such that the classified pattern is placed in either one region with an output of "1" or in another region with a classification of "0". There are no third possibilities in a PNN output function. Moreover, PNNs are not related in any form or manner to the normal distribution method. A typical PNN is shown in the diagram below:

INPUT
LAYER

PATTERN
LAYER

SUMMATION
LAYER

OUTPUT
LAYER
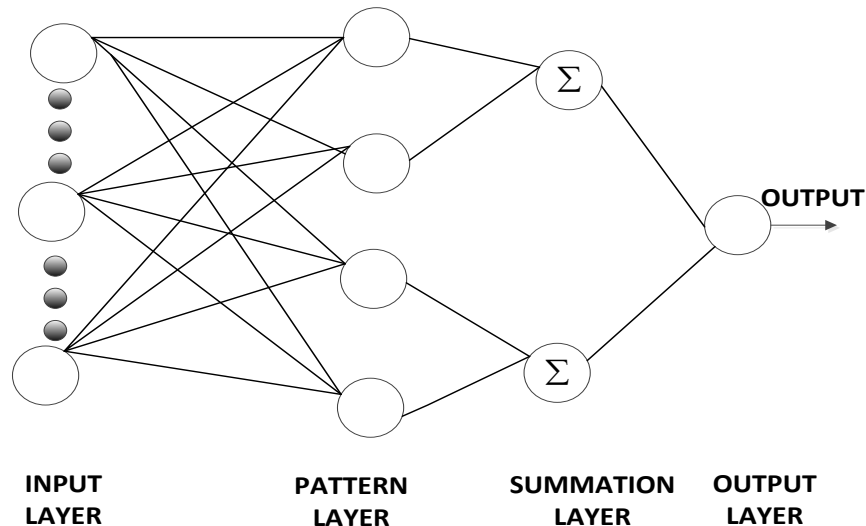
Figure 6.2 Typical PNN structure (Tripathy, et al., 2010)

PNN operates in a non-linear fashion and operates with non-parametric methods as a pattern recognition algorithm. A PNN tends to define a probability density function for the various data classes using information gathered from the training sets as well as based on the kernel width. The outputs of any PNN can be seen as Bayesian probabilities which classify that the inputs are members of the output classes with some given probability. Any PNN is therefore based exclusively on the Bayes classification technique and can be expressed mathematically as (Masters, 1995):

$h_i c_i f_i(x) > h_j c_j f_j(x)$ Equation 6.2

Where:

h_i and h_j are the preceding probabilities

c_i and c_j are the costs of improper classification

f_i and f_j are the probability density functions

In any PNN, the preceding probabilities are unknown since it is hard to determine if the input sample would be from the training set data used to evolve the subject PNN. The best possible approach to this problem is to utilise the training set data to estimate such probabilities. This form of evaluation is based on Parzen's method of probability density function estimation (Shaffer, Rose-Pehrsson, & McGill, 1999).

6.2.4 Classification Theory of PNN

Classification is one of the central tenets of PNNs. This scheme allows PNNs to operate and learn faster than other types of ANNs. Moreover, classification provides for the centralised scheme that converts the inputs to the outputs based on learning from the training sets. Classification can be understood better if it is looked into from a mathematical perspective for any typical PNN.

If a vector x is considered such that it has m dimensions for the input of a PNN, let us assume that it could be classified into either of two categories. These categories can be classified better as C_1 and C_2 such that there can be no other category where the vector x could be classified. In a similar manner, the probability density functions for the categories C_1 and C_2 could be labelled as $F_1(x)$ and $F_2(\mathrm{x})$. In such a case, L_1 would represent the cost function or the loss linked to improperly classifying the input vector for category C_1. On the other hand, L_2 would represent the cost function or the loss linked to improperly classifying the input vector for category C_2. Similarly, P_1 would be the prior probability that x belongs to the category C_1. Conversely, P_2 would be the prior probability that x belongs to the category C_2.

Implementing the Bayes decision rule, it could be surmised that x could only belong to the category C_1 if:

$$\frac{F_1(\mathrm{x})}{F_2(\mathrm{x})} > \frac{L_1 P_2}{L_2 P_1}$$
Equation 6.3

Alternatively, x could only belong to the category C_2 if:

$$\frac{F_2(\mathrm{x})}{F_1(\mathrm{x})} > \frac{L_1 P_2}{L_2 P_1}$$
Equation 6.4

It is also worth mentioning that for a number of different cases of PNNs, the prior probability as well as the loss function are either equal or are nearly equal. Using this condition on the classification mechanisms presented above makes it clear that classification depends in large part on probability density functions exclusively. Hence, it could be surmised that probability density functions derived from the training patterns can be utilised with success to classify input vectors for a given PNN (Goh, 2002).

The Parzen window technique is used in PNNs in order to estimate the class dependent probability functions. Attentive that such estimate are non-parametric in nature. The class dependent probability functions created in this manner are then further used for classification in any of the classification categories present using the Bayes decision rule. This technique allows the determination of whether an input vector pattern could be classified into one of the provided categories. This information is then further combined with the relative occurrence of each category so that the PNN can select the category that most likely espouses the provided input vector pattern. It must be taken to note that Bayes decision rule as well as Parzen window applications have been used for decades and are well established so as not to leave out any doubt. Mathematically, the Parzen window estimate for the probability distribution function for the category C_1 can be expressed as(Goh, 2002):

$$F_1(x) = \frac{1}{(2s)^{\frac{m}{2}} W^{mn}} \sum_{n=1}^{n} exp\left[\frac{-(x-x_j)^T(x-x_j)}{2w^2}\right]$$

Equation 6.5

Where:

n are the various data training sets used

m is the dimension of the input vector x

j is the pattern being utilised

W is the smoothing parameter being utilised

With reference to the basic architecture of a PNN above, the first layer is the input layer which is composed of m different input variables that make up the input vector x. Neurons present in the first layer act to distribute the various inputs of the vector x to the second layer. The second layer, known better as the pattern layer, has as many neurons as are present in the input layer. These patterns in the pattern layer are represented better by each neuron being present for each input of the input layer. Neurons present in this layer are assigned various weights that depend on the various training sets utilised. The summation of the Parzen window estimate's exponential terms is carried out by the summation layer which has a lower number of neurons when compared to the previous layer. Instead, the summation layer is composed of as many neutrons as there are categories since each neuron tends to represent each category. It is important to note that summation layer connections are essentially limited at 1 so that the outputs from the pattern layer are summated together only. These values are then carried into the final output layer that has only one neuron element that performs the classification. The neuron in the output layer of the PNN tends to provide a binary value resulting from the probability density function with the greatest output value. In essence, the highest value obtained through this mechanism provides the optimal classification that the PNN can provide for the input pattern (Bolat & Yildirim, 2003).

6.3 SVMs for Fault Classification

Learning requires that errors be reduced utilising either better equipped training sets or using more structured statistical approaches. The SVM method is essentially a statistical technique that relies on producing a robust statistical learning method from statistical theory. The structural risk minimisation offered by the SVM approach relies essentially on minimisation of the error's upper bound limits (Boser, Guyon, & Vapnik, 1992; Cortes & Vapnik, 1995; V. N. Vapnik, 1995). One of the biggest advantages offered by SVM is that it can handle a large number of features with ease since the training for SVM approaches is carried out to facilitate such handling. The training of SVM techniques ensures that the dimensions of the available vectors do not impact the SVM performance in any significant way at all. Traditional classification methods do not offer this advantage and are instead affected in terms of their performance when dealing with large feature sets. The application of SVM to exhaustive classification problems has been preferred over other methods given the efficiency and speed offered by SVM when compared to other methods. In turn, this particular characteristic of

SVMs tends to make them more suited to fault classification problems since there are large feature sets that must be searched and processed in order to produce the required results. The apt application of SVM ensures that the processing does not omit any features or faults that would cause a false diagnosis (A. Widodo & Yang, 2007).

In a similar manner, the SVM approach also boasts better generalisation characteristics than other methods used for fault diagnosis with large feature sets. The inherent nature of the SVM approach ensures that the improper classification of features is minimised to the most optimal using training data sets. In contrast, the more conventional approaches to learning require that the amount of error be minimised on the data set instead of on the learning technique itself (Cristianini & Shawe-Taylor, 2000).

In describing the SVM emphasis is on the engineering and physics. If required, details of the mathematical methods can be found in, e.g. (Z. Chen & X . Lian, 2010; Gunn, 1998; A. Widodo, Yang, B-S. , 2007).
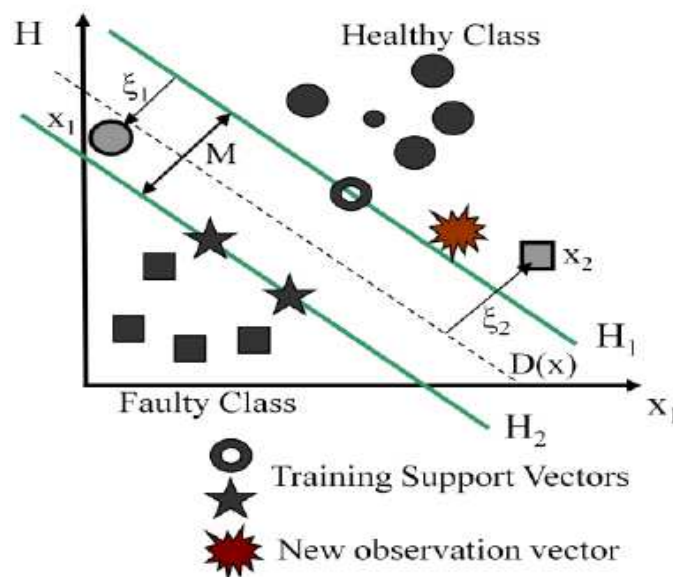
Figure 6.3 Classification of binary classes using SVM

Consider Figure 6.3, showing only two kinds of training samples: ● and ■ . Where ● represents healthy and ■ represents faulty. H is the classifier hyperplane dividing the two groups of samples; x1 and x2, are the data points closest to H; H1 and H2 are parallel to H and pass through x1 and x2 respectively. Consider a planar classification task where, optimally, the set of vectors should be separated by the hyperplane without error. The distance separating the closest points of the two classes (distance between H1 and H2) is defined as the margin (Chapelle, Haffner, & Vapnik, 1999). The task is to maximize the margin (minimise the error bound) to give best performance. Note that this problem is linear.

In standard form the separating hyperplane must satisfy the following constraints:

$$y_i(w.x_i + b) \geq 1 \qquad i = 1,2,…,n \qquad\qquad\qquad \text{Equation 6.6}$$

Where: x_i is the set of training samples, $w.x_i$ is the dot product, n is the number of samples, b is a scalar measure of the distance of H$_2$ from the origin, and w is the normal vector to the hyperplane. Here the samples are assumed be in only one of two classes: healthy or faulty. For the healthy class $y_i = +1$, and faulty class, $y_i = -1$.

However, in most real situations such an ideal hyperplane does not exist. To find the optimum solution the standard technique is to relax the constraints on (8.1) by introducing a slack variable, $\xi_i \geq 0$. This slack variable is said to represent the noise in the system. The solution to this problem requires the application of advanced but relatively well-known mathematical techniques. The calculation is converted into the equivalent Lagrangian dual problem and the learning task is reduced to minimising the primal Lagrangian with respect to w and b:

$$L(w, b, \propto) = \tfrac{1}{2}\|w\|^2 + \sum_{1=1}^{n} \propto_i y_i (w.x_i + b) \qquad \text{Equation 6.7}$$

Where $\propto_i$ are Lagrangian multipliers.

Finding the optimal values for a_i allows w to be expressed in terms of $\propto_i$ which allows the solution of (6.7) to be found. The optimal values for $\propto_i$ give the decision function:

$$f(x) = \text{sgn}(\sum \propto_i y_i (x_i.x_j + b)) \qquad \text{Equation 6.8}$$

This section refers to a linear problem in which the training samples, ● and ■ , were separable both in the original input space and in the feature space (hyperspace). However, with multiple dimensions, the features in the original input space will not normally be separable. Nevertheless a suitable choice of a so-called kernel function to be used in the decision function will separate the features in hyperspace.

$$f(x) = \text{sgn}(\sum \propto_i y_i (\varphi(x_i).\varphi(x_j) + b)) \qquad \text{Equation 6.9}$$

The importance of this is that the analysis performed in hyperspace becomes linear. The kernel function is written $K(x_i.x_j) = \varphi(x_i).\varphi(x_j)$. There are now standard kernel functions and this work uses the very popular polynomial function (Z. Chen & X. Lian, 2010):

$$K(x_i.x_j) = (x_i.x_j + 1)^P \qquad \text{Equation 6.10}$$

The SVMs algorithm was designed to separate two classes as binary classifiers. In the real problem, however, we find more than two classes for examples: in fault diagnosis of RCs there are a number of faults (see chapter 3) as mechanical approaches that address a multi class problem as a single 'all together' optimisation problem exist (Weston, Watkins, & 1999), but are computationally much more expensive than solving several binary problems. Therefore, a variety of techniques for decomposition of the multi-class strategy into several binary using SVMs as binary classifiers have been proposed, and several widely used are given below:

➢ **One-Against-All (OAA)**

The SVM multi class classification was implemented early for OAA approach. The flowchart of working process of this model is shown in Figure 6.4. The construction of k SVM models where k is the number of classes. The ith SVM is trained with all of data examples in the ith class with negative labels and positive labels. Thus l training data $(x_1, y_1), \dots (x_l, y_l)$, where $(x_i) \in R^n$,
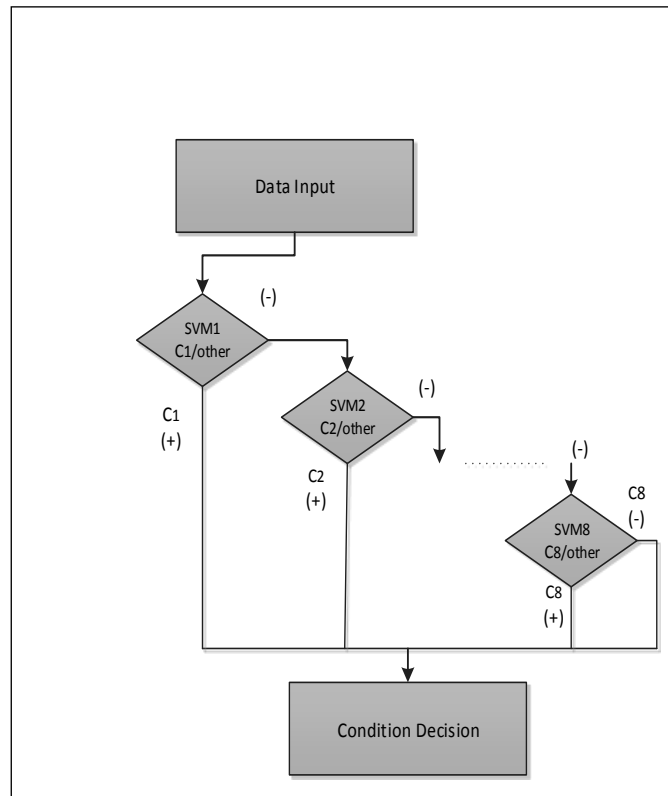
Figure 6.4 Flow chart of working process of multi–class SVM-OAA

$i = 1 \ldots, l$ and $y_i \in [1, \ldots, l]$ is the class of x_i, the ith SVM solve the following problem (A. Widodo & Yang, 2007):

$$\frac{1}{2}\left\|W^i\right\|^2 + C\sum_{i=1}^{l} \xi_j^i (W^i)^T \qquad \text{Equation 6.11}$$

$$(W^i)^T \phi(x_j) + b^i \geq 1 - \xi_j^i \quad \text{if } y = 1 \qquad \text{Equation 6.12}$$

$$(W^i)^T \phi(x_j) + b^i \leq -1 + \xi_j^i \quad \text{if } y \neq 1 \qquad \text{Equation 6.13}$$

$$\xi_j^i \geq 0, \; j = 1 \ldots, l \qquad \text{Equation 6.14}$$

Where the training data x_i is mapped to a higher-dimensional space by function ϕ and C, is the penalty parameter.

Minimizing Equation (6.11) means we would like to maximize $2/\|W_i\|$, the margin between two groups of data.

When data is not separable, there is a penalty term $\sum_{i=1}^{l}\xi_j^i$, which can reduce the number of training errors.

➢ **One-Against-One**

Another approach for multi class SVM is called (OAO). In this model, for k classes will result in $k[k-1]/2$ binary classifiers as shown in the flowchart in Figure 6.5. In this method the number of classifiers is larger than the number of (OAA). For training the data from the ith and the jth classes it is solved as following:

$$\frac{1}{2}\left\|W^{ij}\right\|^2 + C\sum_t \xi_t^{ij}(W^i)^T \qquad \text{Equation 6.15}$$

$(W^{ij})^T \phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij} \quad \text{if } y_t = i$ <div style="float:right">Equation 6.16</div>

$(W^{ij})^T \phi(x_t) + b^{ij} \leq -1 + \xi_t^{ij} \quad \text{if } y_t = j$ <div style="float:right">Equation 6.17</div>

$\xi_{jt}^{ij} \geq 0, \quad j = 1 \dots, l$ <div style="float:right">Equation 6.18</div>

Once the classifiers are created. It tested by some of the strategies to make a decision. One of these strategies if sign $(W^{ij})^T \phi(x_t) + b^{ij}$ states x is in the ith class, then the vote for ith class is added by one. If not, then x is predicted in the class using the largest vote and the jth is added by one.
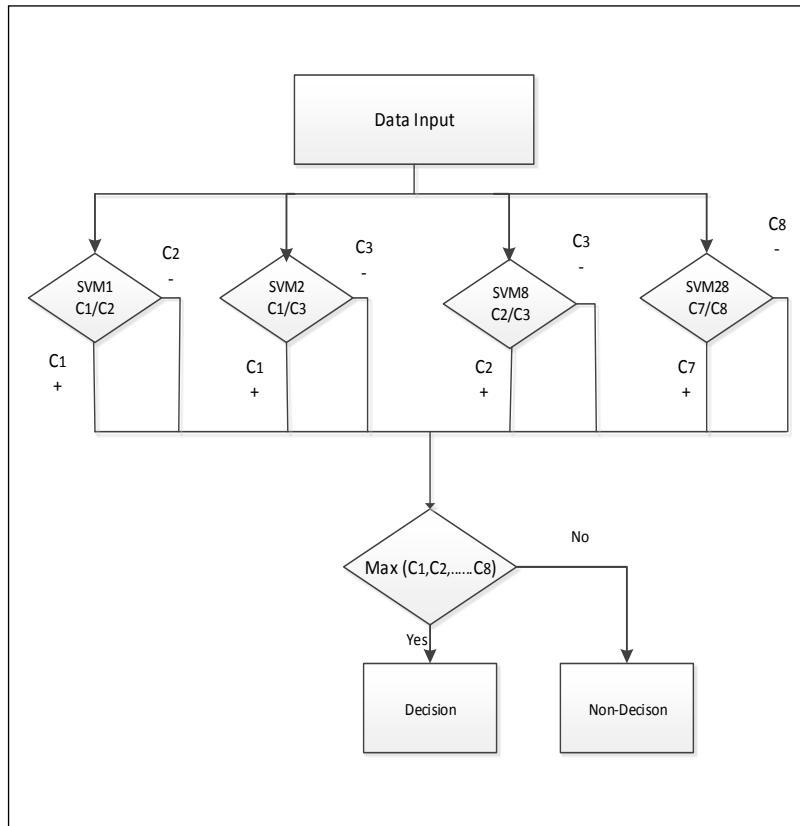
Figure 6.5 Flow chart of working process of multi-class SVM-OAO

6.4   *RVMs for Fault Classification*

RVM has also been used for classification. Consider a two-class problem with training points $X = [x_1, \dots x_N]$ and corresponding class labels $t = [t_1, \dots t_N]$ with $t_i = [0,1]$. Based on the Bernoulli distribution, the likelihood (the target conditional distribution) is expressed as:

$p(t\backslash W) = \prod_{i=1}^{N} \sigma((y(x_i)))^{t^i} [1 - \sigma((y(x_i)))]^{1-t_i}$ <div style="float:right">Equation 6.19</div>

where $\sigma(y)$ is the sigmoid function:

$\sigma(y(X)) = \frac{1}{1+\exp(-y(x))}$ <div style="float:right">Equation 6.20</div>

Unlike the regression case, however, the marginal likelihood $p(t\backslash \alpha)$ can no longer be obtained analytically by integrating the weights from "(1)", an iterative producer has to be used (D. G. Tzikas, Wei, Likas, Yang, & Galatsanos, 2006).

Let α_i^* denote the maximum a posterior (MAP) estimate of the hyper parameter α_i. The MAP estimate for the weights, W_{MAP}, can be obtained by maximising the posterior distribution of the class labels given the input vectors.

$$J(w_1, \dots w_N) = \sum_{i=1}^{N} \log p(t_i \setminus w_i) + \sum_{i=1}^{N} \log p(w_i \setminus \alpha_i^*) \qquad \text{Equation 6.21}$$

The first summation term denotes the likelihood of class labels and the second term denotes the prior parameters, w_i. In the resulting solution, only those samples associated with non-zero coefficients w_i (relevance vectors) will contribute to the decision function.

The gradient of the objective function J with respect to w is:

$$\nabla J = -A^* W - \varphi^T (f - t) \qquad \text{Equation 6.22}$$

whereby $f = [\sigma(y(x_1)) \dots \sigma(y(x_N))]^T$, matrix φ has elements $\emptyset_{i,j} = K(x_i, x_j)$.

The Hessian of J is: $H = \nabla^2(J) = -(\varphi^T B \varphi + A^*) \qquad \text{Equation 6.23}$

where $B = \text{diag}(\beta_1, \dots \beta_N)$ is a diagonal matrix $\beta_1 = \sigma(y(\beta x_i))[1 - \sigma(y(\beta x_i))]$.

The posterior is approximated around W_{MAP} by a Gaussian approximation with covariance

$$\sum -H \setminus W_{MAP}{}^{-1} \qquad \text{Equation 6.24}$$

and mean: $\mu = \sum \varphi^T Bt \qquad \text{Equation 6.25}$

6.5   GA Based Classification and Feature Selection

6.5.1 GA based Optimisation

When using ANNs, mindful that the processing mechanism relies on comparing the current provided input to the vast array of possibilities that have been used as the learning data sets. It is not always possible to utilise the large array of possibilities, although it is still required (Sen, 2001). For example, for a rotating mechanical system, a large number of features are required in order to allow for a robust fault determination scheme to operate. However limitations, in terms of both availability of computational resources as well as available time, force the usage of a set of features that are representative of the entire set of features. The representative set of features need to be chosen with care so as to make them comprehensive enough for usage yet compact enough for consuming an optimised amount of computational resources and time. A holistic solution cannot be used for neural networks since the options are too many and the amount of resources and time are typically limited. The feature selection process to be used must be brief yet comprehensive enough whilst retaining the original purpose behind the neural network implementation (Saxena & Saad, 2006).

In order to retain the original intent behind the neural network, to speed up the computation required and to optimise the final outputs, GAs are used. Real time and practical applications of neural networks require implementing a large number of possibilities that cannot be captured in demanding and practical situations. The application of GAs to ANNs has been successfully demonstrated for a number of different applications. In essence, GAs serve as the classifiers required to simplify and optimise the practical operation of ANNs. The application of GAs to neural networks using a reduced set of characteristic features has provided highly

optimised results. It must also be mentioned that the application of GAs to ANNs has also required the usage of optimised parameters for ANNs since it aids the performance of the overall classification of required and desired features (L. B. Jack & Nandi, 2000; Samanta, 2004).

Conventionally the application of choosing optimal feature sets has relied on two major optimisation methods namely gradient based searches that rely on calculus methods and dynamic programming (Tang, Man, Kwong, & He, 1996). Dynamic programming may be more effective than gradient based searches but are still not practicable enough for use with neural networks given the large amounts of data to be processed. GAs allows the simplified solution of problems with a high dimensional character such as those present when choosing between a large numbers of varying features. The exhaustive set of features for vibration monitoring and fault diagnosis would consist of features that have no relation to fault diagnosis and are hence not required. One possible solution for this situation is to utilise allied methods to reduce the problem's dimensionality, such as the utilisation of principal component analysis (PCA) as well as other similar methods (Vlachos, Domeniconi, Gunopulos, & Kollios, 2002).

6.5.2 GA based Classification

A GA is a search technique used in computing to find a solution in optimization problems (Tang, et al., 1996). It applies the principles of evolution found in nature to the problem of finding an optimal solution. In a "GA", the problem is usually encoded in a series of bit strings that are manipulated by the algorithm. Based on the process of GA optimisation, a feature selection scheme for fault classification can be conducted based on a process as following:

➢ **Chromosome Encoding**

A binary chromosomal representation is adopted for the problem in this work. The length of the chromosome depends on the number of inputs features to the AI appoach required in the solution set. For both first and second simulations tested with features from time domain and envelope spectrum, the number ranged between 10 and 20.

➢ **The Creation of an Initial Population**

The initial generation was generated with randomly selected features, each of which has a numerical value. The maximum number is determined by the size of the feature space available. When computation starts, the number is put into binary form more easily allow the generation of a new population of the parameters and performing GA operations.

➢ **The Fitness Evaluation**

At this step in the development of the NN model, and calculation of the classification error in the objective function, the binary form is translated back to decimal numbers. The objective function is defined as:

$$\mathrm{Err} = \sum_{i=1}^{N\,max}(t_i - \bar{t}_i)^2 \qquad\qquad \text{Equation 6.26}$$

Where $\bar{t}_i$ is the predicted class by PNN and t_i is the labelled class. N max is the size of the dataset. The program terminates when Err is equal to or less than the set minimal level, or

when it becomes equal to the maximum generation number (N_{gen}) set in the computational code developed for the study.

➢ **The Selection Operation**

This is the operation whereby the process is guided towards ever-better solutions by the GA. Different algorithms can be used to select the best individual values from the estimated population, such as: Roulette wheel selection, Tournament selection, Remainder selection and Uniform selection. Tournament selection has been used in this experiment as it is efficient and easy to implement. In tournament selection method, n individuals are selected randomly from the larger population, and the selected individuals compete against each other. The individual with the highest fitness wins and will be included as one of the next generation population. The number of individuals competing in each tournament is referred to as tournament size. Tournament selection also gives a chance to all individuals to be selected and thus it preserves diversity, although keeping diversity may degrade the convergence speed.
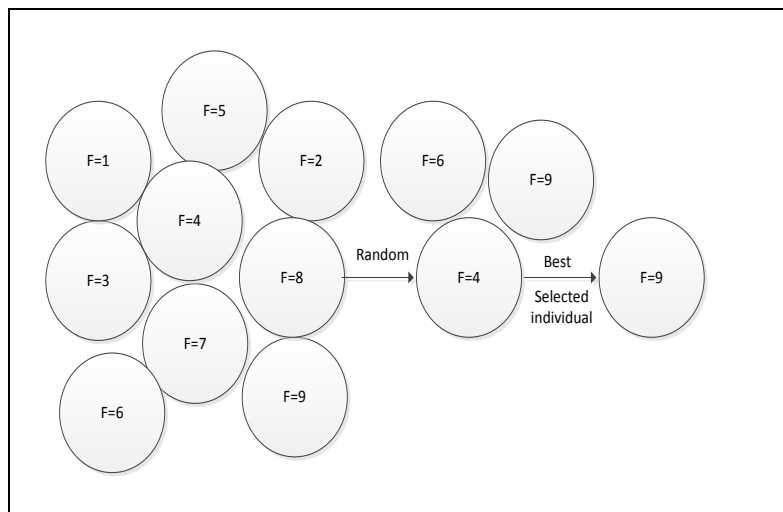
Figure 6.6 Selection strategies with tournament mechanism

In this example, the tournament size T_s is set to three, which mean three chromosomes competing with each other. Only the best chromosome among them is selected to reproduce. In tournament selection, larger values of tournament size lead to higher expected loss of diversity (Blickle & Thiele, 1995; Razali & Geraghty; Whitley, 1989). The larger tournament size means that a smaller portion of the population actually contributes to genetic diversity, making the search increasingly greedy in nature. There might be two factors that lead to the loss of diversity in regular tournament selection; some individuals might not get sampled to participate in a tournament at all while other individuals might not be selected for the intermediate population because they have lost a tournament.

➢ **The Crossover Operation**

This stage produces the next generation by using a simulated mating process. This is performed by two parents creating offspring which consist of genetic material from both parents. This operation combines features of existing solutions in an attempt to create a better

solution. For example, elements of existing solutions can be combined in a crossover operation as shown:

Parent 1	0	1	1	0	1	0	1	0	1	0

Parent 2	1	0	1	1	1	1	0	1	0	0

Offspring 1	0	1	1	1	1	1	0	1	0	0

Offspring 2	1	0	1	0	1	0	1	0	1	0

The probability of the crossover was assumed to be 0.80 in this experiment.

➢ **The Mutation Operation**

Once the children are generated during crossover, the mutation operator is applied to each child. Each gene has a user-specified mutation probability of being mutated. In this experiment the initial probability of mutation was 0.01, this was sometimes increased and/or decreased randomly to fit. Again the location of mutation was randomly determined every time.

After crossover	0	1	1	1	1	1	0	1	0	0
After mutation	0	0	1	1	1	1	0	1	0	0

After crossover	1	0	1	0	1	0	1	0	1	0
After mutation	1	1	1	0	1	0	1	0	1	0

6.5.3 GA-ANN

The initial weight values and thresholds of PNN are optimized based on GA. The optimal solution is acquired from the whole searching space of GA to generate the initial weight values and thresholds for PNN. The search is then conducted in negative gradient direction in PNN to reach the target value. PNN optimization based on GA flowchart is shown in Figure. 6.7.

1) Determine the topological structure of PNN, including the numbers of layer and hidden nodes.
2) Decode according to the scope of each value, decoded chromosomes are the initial weight values and thresholds of the PNN.
3) Conduct selection, crossover and mutation; adopt the corresponding operator to calculate the next generation.
4) Calculate the fitness value.
5) If the fitness value is larger than the previous set value, then the genetic operation is stopped. If not, repeat phase (3) while the individual with the largest fitness value is chosen as the initial weight values and thresholds of the PNN.

6) Substitute the initial weight values and thresholds into PNN and conduct its process to calculate the model error. The iteration is terminated when model error meets the demand of set value.
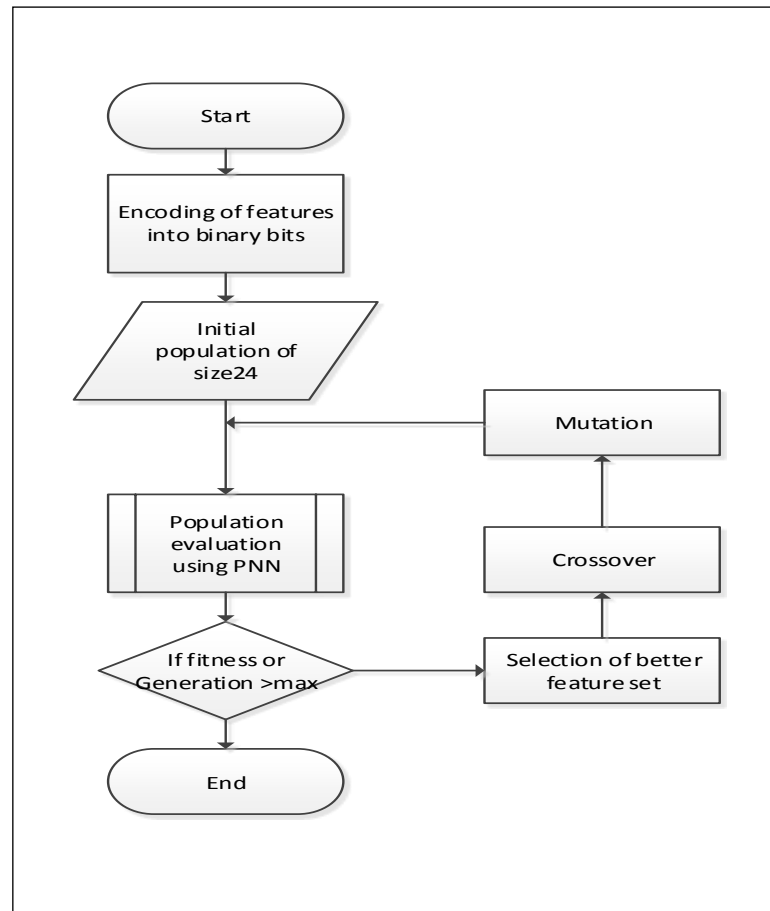
Figure 6.7 Flow chart of PNN optimization based on GA (GA-PNN)

6.5.4 GA-SVM

The selection of the SVM parameters plays an important role in the performance of SVM. To design an effective SVM model, values of parameters in SVM have to be chosen carefully in advance (Awadalla, Abdien, Rashad, Ahmed, & Al Abri, 2014; Chou, Wu, & Chen, 2010). The SVM with RBF includes the following parameters:

A. Regularization parameter C, which determines the trade-off cost between minimizing the training error and minimizing the complexity of the model (Guarnaccia & Neri, 2013; Volos & Neri, 2012);

B. Parameter sigma (σ) of the kernel function which defines the non-linear mapping from the input space to some high-dimensional feature space, which constructs a non-linear decision hyper surface in an input space;

C.   Parameter ξ is non-negative slack variables and provides the minimum training error.

In this study, GA is used to determine the optimal values of C, σ and ε that assure highest predictive accuracy and generalization ability simultaneously. The proposed model is named SVM-GA. Since three parameters should be optimized, the chromosome should comprise three parts, C, σ and ξ as shown in Figure 6.8, where the binary coding system was used to represent it.

b_c^1	$...b_c^j...$	b_c^{nc}	b_σ^1	$...b_\sigma^k...$	$b_\sigma^{n\sigma}$	b_ξ^1	$...b_\xi^l...$	$b_\xi^{n\xi}$
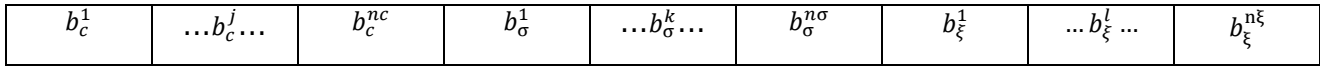
Figure 6.8 The chromosome comprises three parts C, σ and ξ

6.5.5 GA-RVM

A GA-based parameter method was used to automatically obtain the optimal features of the RVM classifier. The classification rate of the training data is the fitness function of the GA. Therefore, the optimum features are achieved when a minimum error is detected by the classifier to complete the generation. Also for the GA-RVM-OAO method, the population size is considered to be equal to 20. The initial range is taken to be within [0, 2] for all individuals.

6.6   PCA based Detection and Diagnosis

A primary objective of PCA is for dimensionality reduction or data compression to achieve efficient data analysis. To this end, PCA results in a new smaller subset of variables with minimal loss of information, compared with the original data. Based on this unique characteristic, PCA is used for the classification of different cases with the new subset of variables and hence early identification of abnormalities in the data structure, i.e. detection of faults.

The PCA creates a covariance matrix (or correlation matrix) by transforming the original correlated variables into a new set of uncorrelated variables. Let the variables describing the machine being investigated be the m–dimensional data set: $X = x_1, x_2, x_3, ... x_m$, the PCA decomposes the observation vector, X, into a set of new directions P as (J. F MacGregor & T Kourti, 1995):

$$X = TP^T = t_1 P_1^T + t_2 P_2^T + \cdots + t_m P_m^T = \sum_{i=1}^{m} t_i P_i^T \qquad \text{Equation 6.27}$$

Where P_i is an eigenvector of the covariance matrix of X. P is defined as the principal component loading matrix and T is defined to be the score matrix of the principal components (PCs).

The loading matrix helps identify which of the variables contribute most to individual PCs, whilst the score provides information on sample clustering and identifies transitions between different operating conditions.

The expectation with PCA is that the original variables are sufficiently well correlated that only a relatively small number of the new variables (PCs) account for most of the variance. In this case no essential information is lost by using only the first few PCs for further analysis and equation (6.27) can be expressed as (Wise & Gallagher, 1996):

$$X = TP^T + E = \sum_{i=1}^{k} t_i p_i{}^T + E \qquad \text{Equation 6.28}$$

where E represents a residual error matrix. For example, if only the first three PCs represent a sufficiently large part of the total variance, E will be calculated by:

$$E = X - [t_1 p_1{}^T + t_2 p_2{}^T + t_3 p_3{}^T] \qquad \text{Equation 6.29}$$

Therefore, by evaluating the amplitude of E, fault detection can be implemented. If the system is normal, E is small, whereas it becomes large if the system is faulty.

However, in certain applications such as process monitoring, when a plant malfunctions, original variables have minimal impact on the first few PCs, but dominate the higher orders. Thus in process engineering, use of these higher order components may be needed to provide the necessary diagnostic information (J. F MacGregor & T Kourti, 1995). In this way E can be very useful to measure these changes.

6.6.1 PCA Model Based Detection

PCA based fault detection is usually based on two detection indices: Hotelling's T^2 statistic and Q statistic.

Hotelling's T^2 statistic is a measure of the major variation of measurement variation and detects new data if the variation in the latent variables is greater than the variation explained by the model or baseline condition. For a new measurement feature vector x, the T^2 statistic detection can be found from:

$$T^2 = x^T P \lambda^{-1} P^T x \leq T_\alpha^2 \qquad \text{Equation 6.30}$$

Where the $100(1 - \alpha)\%$ control limit for T_α^2 is calculated by means of a F-distribution as (J.F MacGregor & T Kourti, 1995):

$$T_\alpha^2 = \frac{k(m-1)}{m-k} F(k, m-1; \alpha) \qquad \text{Equation 6.31}$$

Where $F(k, m-1; \alpha)$ is an F-distribution with k and $(m-1)$ degrees of freedom, with chosen level of significance α, k is the number of PC vectors retained in the PCA model, and m is the number of samples used to develop the model. The Q statistic, also represented as SPE, is the squared prediction error. It is a measure of goodness of fit of the new sample to the model. The Q statistic based detection can be done by:

$$SPE = \|(I - PP^2)x\|^2 \leq Q_\alpha \qquad \text{Equation 6.32}$$

The $100(1 - \alpha)\%$ control upper limit Q_α (Kourti & MacGregor, 1996) is:

$$Q_\alpha = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \qquad \text{Equation 6.33}$$

Where:

$$\theta_i = \sum_{j=a+1}^{m} \lambda_j^i \qquad \text{Equation 6.34}$$

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \qquad \text{Equation 6.35}$$

6.6.2 Contribution Plots of T^2 and Q Statistic

Once an abnormal factor has been detected using the T^2 or SPE, it is important to diagnose the event to find its cause. The contribution of the measurement variable and time periods to the deviation observed in the Q and T^2 statistics can be used to help suggest an assignable cause. The Q-contribution plot, which is a bar graph representing Q residual contribution (the significance of each variable on the index) versus variable numbers for certain sample, can be used to diagnose the fault. When the T^2 or SPE breaks the threshold, the contribution of the individual variables to the T^2 or SPE can be identified, and the variable making a large contribution to the T^2 or SPE is indicated to be the potential fault source. In general, the Q-contribution plot helps to reduce the possible fault.
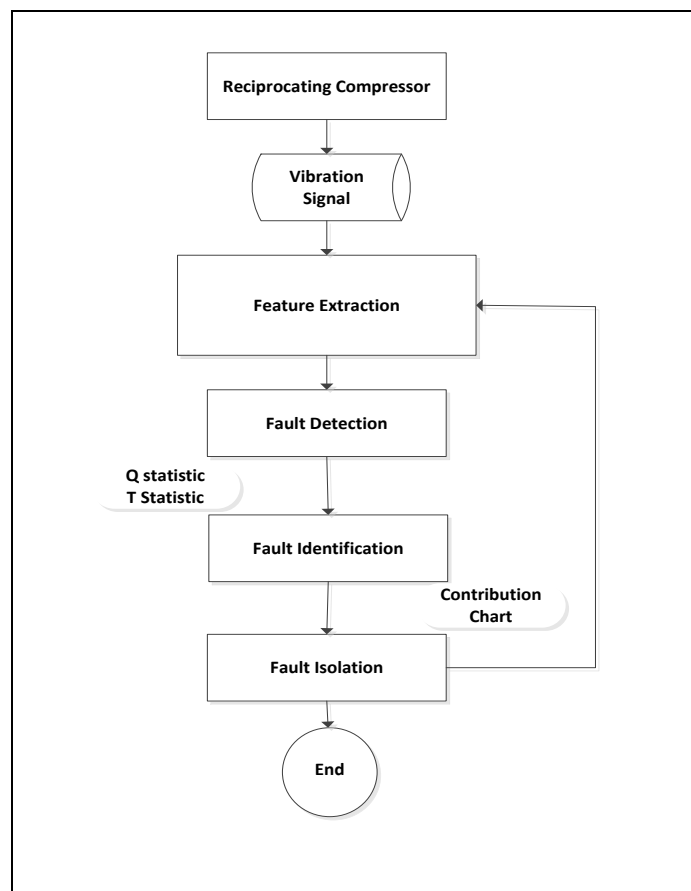
Figure 6.9 Flowchart diagram of fault diagnosis processing

Figure 6.9 the flowchart diagram shows that once a fault is successfully detected, the source of the fault can be identified and isolated. Using the distributions, confidence limits for the two statistics can be obtained. For the monitoring of new batches, the process data of the new batch $X_{new}(JK \times 1)$ is projected onto the model.

$$c_j\left(T_k{}^2\right) = t_{new,k}^T S_k{}^{-1}[x_{new,kj}P_j{}^T]^T \qquad \text{Equation 6.36}$$

Where $c_j\left(T_k{}^2\right)$ is the contribution of the jth variable *to* $T_k{}^2$, $x_{new,kj}$ is the jth element of $x_{new,kj}(JX1)$, and $P_j{}^T$ is the jth row of the loadings matrix $P(JXR)$.

The contribution $c_j(Q_k)$ of process variable j at time k to the Q statistic can be calculated as following:

$$c_j(Q_k) = e_{kj}{}^2 \qquad\qquad \text{Equation 6.37}$$

where e_{kj} is the jth variable of $x_{new,k}$ (Lee, Yoo, & Lee, 2004).

6.7   Summary

In this chapter, AI techniques, ANN, SVM, and RVM, are proposed to use to classify machine conditions. Features are first extracted from machine raw vibration data using envelope spectrum, and then the dimension of the features will reduce by GA. With the selected features, ANN, SVM, and RVM classification models are built to classify different machine conditions. In addition, this work proposes an approach to face the fault detection using statistical techniques, concretely, PCA. Data are collected from the test rig for normal conditions in order to calculate the PCA model and the thresholds of the T^2 and Q statistics, used in order to detect the faults. Finally, the contributions of the Q-plot and D- plot, will present in a way which allows it to use with any latent variable component or regression model to detect a specific progress variable to find effective features for fault classification and identify the variables related to the faults.

Chapter 7

Features Selection and Fault Classification of A Reciprocating Compressor Using A Genetic Algorithm and A Probabilistic Neural Network

In this chapter, a healthy case and seven different fault classification schemes were tested to determine the optimal scheme for use. Testing with a PNN classifier using features in the time domain returned the poorest results with correct classification rates of 43% only. Comparably, a PNN classifier using harmonic components from envelope analysis provided better results with the peak classification rate going up to 100% with 30 input features. In contrast, when GA-PNN classifiers based on features in the time domain were used, the best classification rate achieved was only 48% with 10 input features. Next, GA-PNNs with features from envelope analysis were used that provided a highest classification rate of 99% with 110 input features. Also the GA optimized the features for both PNN in time domain and envelope spectrum. Finally, system verification from time domain and envelope analysis was carried out to ensure the robustness of the system in use.

7.1   Introduction

The main contribution of this work is to develop novel approach that combines GA with PNN; the fault classification has been achieved using a PNN approach. The application of features from both time-domain analysis and envelope spectrum in CM is investigated through the approach. The selection of input features and the classifier parameters are optimized using a GA-based approach. The feature from both time-domain analysis and envelope spectrum are used to distinguish between normal and defective RC.

The vibration signals were collected from the accelerometers on the two-stage, single-acting Broom Wade TS9 RC, which has its two cylinders in the form of a "V" (see Figure 3.5 in Chapter 3), and which delivers compressed air at up to 0.8 MPa (120 psi) to a horizontal air receiver tank with a maximum working pressure of about 1.38MPa (200psi), and was then sampled at rate of 55.56 kHz (this enables the high frequencies associated transient events to be collected); the data length is set at 118833 samples. The time duration of data points = number of samples ÷ sampling frequency so the real time duration of the samples is 2.14 sec. these operations repeated 3 times for 12 different discharge pressure. The each segment of data includes more than four working cycles of the compressor.

7.2   PNN for Classification

For each given vector in the input datasets, the corresponding target vector was created in a second matrix to have the same size as the data vector and containing the values 1,2,3,4,5,6,7, and 8 corresponding to healthy and faulty elements in the data vector respectively, used during the PNNs training. This then gave, for the time domain a (1×1920) and Envelope spectrum is (1×192) matrix containing the target data. These are all the input information assembled to train the PNNs.

The data sets have an equal number of samples from normal and faulty RC conditions. Both the data set and target vector were divided into two subsets of equal size by taking every other vector values, of which one was for training the PNN and the other for testing the trained network. In particular, it is implemented in following:

Two data sets were therefore created. In the first dataset from the time domain was divided into two equal subsets. For the training the PNNs giving a training set of (10×960) were used and the rest (10×960) were used for testing. The second from the envelope was also divided into two equal subsets. For the training the PNNs giving a training set of (30×96) were used and the rest (30×96) were used for testing. Then, the data was simulated and the spread parameter of the PNN is identified to be 0.04 by a trial and error procedure. Finally, viewing the classifications results by plotting the figures with different features in both the time and envelope spectrum. Figure 7.1 shows a flow diagram of the PNN procedure.
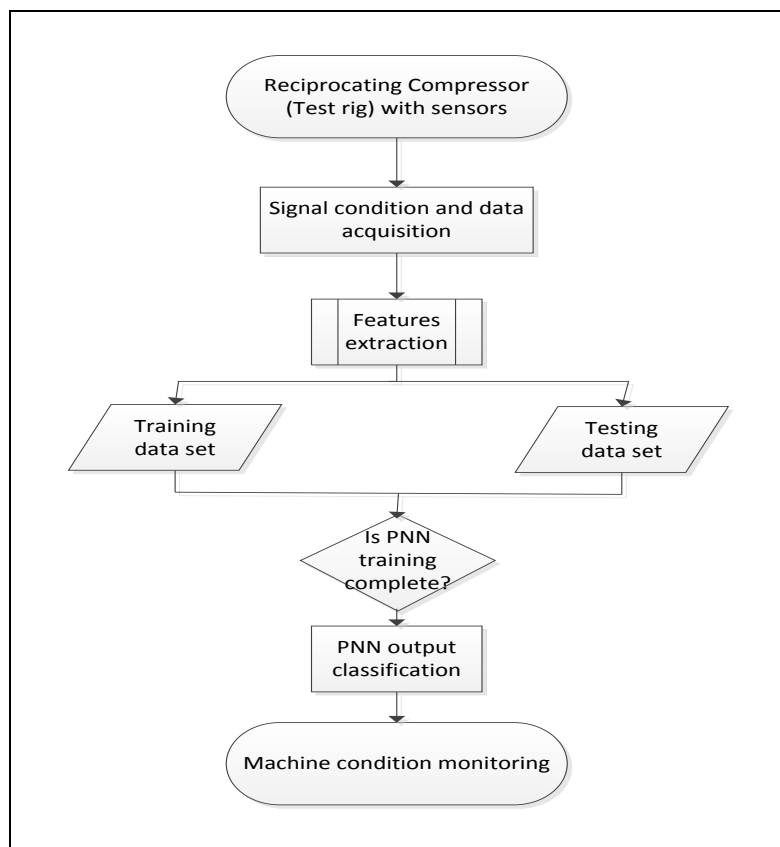
Figure 7.1 Flow diagram of the proposed procedure

7.3   GA Implementation and Simulation

The proposed feature selection scheme is evaluated based on the vibration dataset from the RC test rig in Section 5.1 to 5.3. Also, a MATLAB program is developed to implement PNN and GA. The features from signal waveform, spectrum and envelope spectrum for classification are investigated through the use of PNN, GA and their combination as following.

In the fitness function the same procedure described in section 6.5.2 is repeated in this part for testing the trained network. However, the spread parameter σ of the PNN has been optimised by GA automatically when GA feature selection is involved.

GA has been used to select the optimal features that related to a particular classifier and simultaneously find the optimal spread value of PNN. GA chromosome for this feature selection is a binary chromosome, the length of which is based on the number of features available for selection and the range of the spread parameter of PNN. Typically the chromosome ranges from 10 to 20. Multipoint crossover is implemented with the locations of the crossover points determined randomly. The population is made to crossover after they were paired according to their fitness. After that the chromosomes are sorted in the order of decreasing fitness values.

The fitness function is the minimum of error in classifying the eight classes represented by values 1 for healthy, 2 for fault 1, 3 for fault 2, 4 for fault 3, 5 for fault 4, 6 for fault 5, 7 for fault 6 and 8 for fault 7. Mutation is a multipoint bit-flip mutation based on a pre-specified probability of 0.1. The location of mutation is randomly set in every mutation.

Obviously in this approach, after a new generation of offspring was obtained, the fitness function of all parents and offspring was evaluated and the ones with the highest fitness were carried forward to the next generation.

7.4   PNN and GA-PNN Results and Discussions

For comparison, four sets of PNNs have been studied to evaluate the effectiveness of different types of features and GA feature selection. The first one is the time domain feature based PNN; the second one is the envelope spectrum feature based PNN; the third one is the time domain feature based PNN with GA feature selection; and the fourth is the envelope spectrum feature based PNN with GA feature selection.

7.4.1 Performance of PNNs without GA Feature Selection

7.4.1.1   Performance of PNN Classifier using Features in the Time-domain

A total of ten input features were employed to provide a correct classification rate of 43% as shown in the table provided below. Even with the use of a high number of input features (n=10), the correct classification rate remained low.

Number of input features	Features name	Correct rate	Sigma
10	Max, Range, UB, LB, Kurtosis, Variance, Peak factor, Skewness, Kurtosis, Entropy	43	0.04

Table 7.1 Performance of PNN classifier with different feature combinations from the time domain

7.4.1.2 Performance of PNN Classifier using Harmonic Components from the Envelope Spectrum

For PNN classification through the envelope spectrum, differing sets of input features were used, ranging between 10 and 110 input features that were derived from peak values at shaft frequency and its high order harmonics in the spectrum of envelope signals.

Number of input features	Correct rate
10	92%
20	98%
30	100%
40	99%
50	93%
70	93%
90	93%
110	83%

Table 7.2 Performance of PNN classifier with different feature combinations from the envelope spectrum with sigma=1

Table 7.2 shows that the maximum correct classification rate achieved is 100% with the first 30 input features whereas the lowest classification rate is as low as 83% from the first 100 input features. It is notable that as the number of input features increases beyond 30, the correct classification rate tends to decrease. The signal to noise ratio from the harmonic peaks in the frequency spectrum is low enough to warrant incoherent feature selection at higher numbers of input features. This is detailed in the table below along with Figure 7.2 and 7.3 depicting PNN classifiers with 10 and 30 input features respectively.
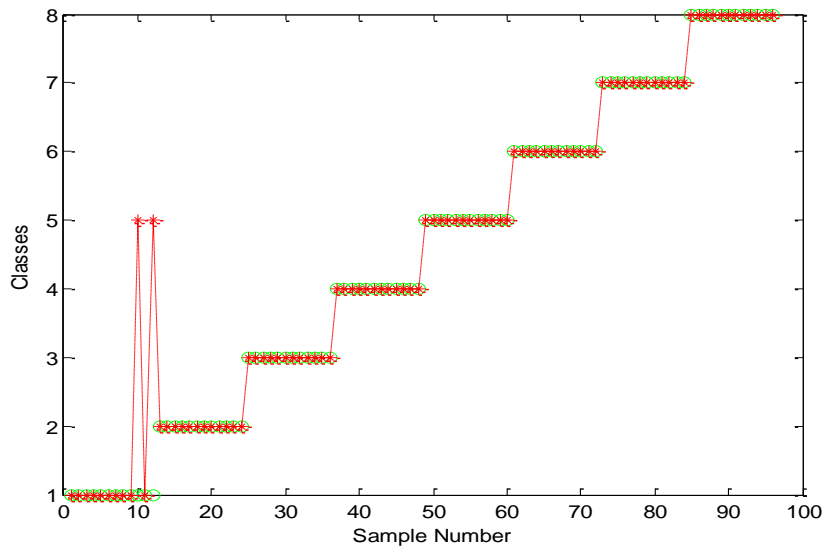
Figure 7.2 PNN classifier with 10 input harmonics
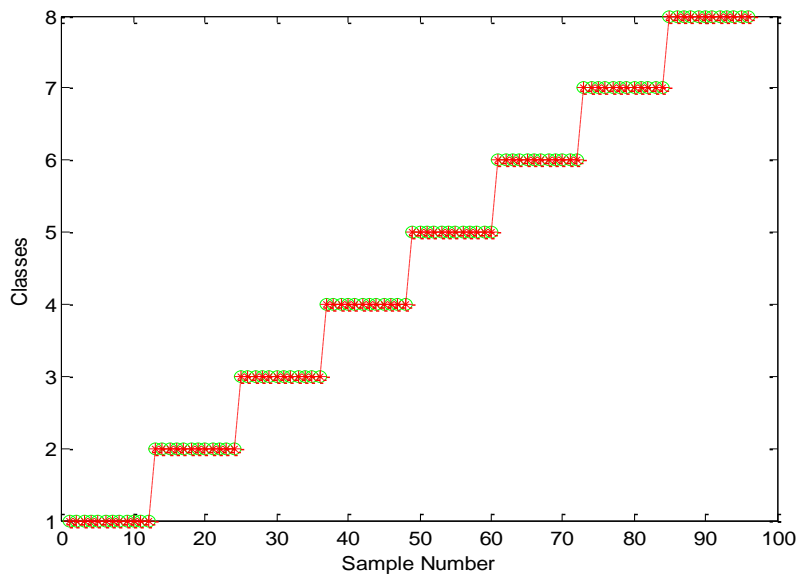
Figure 7.3 PNN classifier with 30 input harmonics

In addition, Table 7.3 depicts the effect of varying the sigma as the number of input features is kept constant. When only 30 input features are utilised, as in the Table 7.3, changes in the sigma value lead to corresponding changes in the correct classification rate. The variation in the correct classification rate with sigma sees a rapid increase followed by relative stabilisation and then a decrease. This indicates that the sigma value needs to be optimised for the use since various sigma values can trigger increases or decreases from the optimal value.

Sigma	Correct rate
0.05	79%
0.09	98%
0.2	98%
0.5	99%
0.9	100%
1.02	100%
1.05	99%
1.1	99%
1.3	97%

Table 7.3 Performance of PNN classifier with input 30 feature combinations from the envelope spectrum and different sigma values

7.4.2 Performance of PNNs with GA Feature Selection

In order to investigate the performance of PNNs with GA feature selection, investigation was carried out using features in the time domain as well as features from the envelope spectrum. It is noteworthy that the sigma for classification and feature selection was generated for these runs through GA to obtain optimal values. The intention behind GA application was to establish the efficacy of each system's application to GA feature selection and to compare current results to previous findings.

7.4.2.1 Performance of GA-PNNs based on Features in the Time Domain

Table 7.4 below depicts the classification results obtained using PNN with feature selection using GA with a high mutation rate (0.01) and a high crossover (10). For this case a combination of 5, 6, 7 and 10 features was utilised with the GA selection so as to achieve the best possible correct classification rate. For instance, using five selection features (n=5), it was discovered that the combination of Maximum, Kurtosis, Variance, Kurtosis and Entropy provided the best correct classification rate. It is notable from the results that using as many as ten features provided only slight improvements in the correct classification rate that rose from a low of 45% to a high of 48%.

Run NO	Selected features	Feature name	Correct rate (%)	Minimum fitness	Sigma
1	7	Max, Range, UB, LB, Kurtosis, Skewness, Entropy	47	0.5278	0.45
2	6	LB, Variance, Skewness, Kurtosis, Entropy	47	0.526	0.46
3	10	Max, Range, UB, LB, Kurtosis, Variance, Peakfactor, Skewness, Kurtosis, Entropy	48	0.5182	0.67
4	6	UB, LB, Kurtosis, Skewness, Kurtosis, Entropy	46	0.5356	0.55
5	5	Max, Kurtosis, Variance, Kurtosis, Entropy	45	0.5451	0.34

Table 7.4 Performance of GA-PNN classifier time domain feature selection

(mutation rate 0.01), Generation 30, population 24, crossover 10, chromosome 10

Figure 7.4 Minimum and mean fitness function with GA-PNN features selection from the time domain (Run 2)
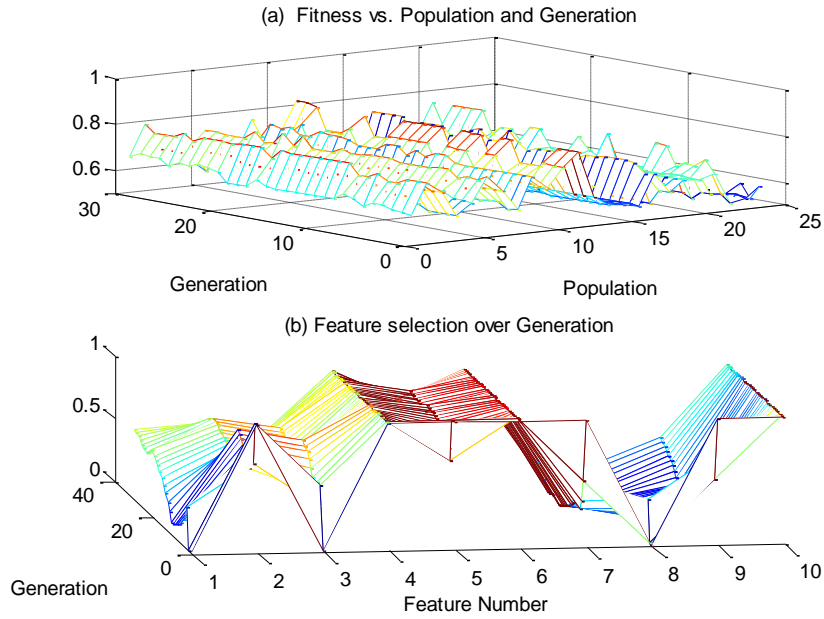
Figure 7.5 Fitness function value and GA-PNN features selection from the time domain (Run 2)
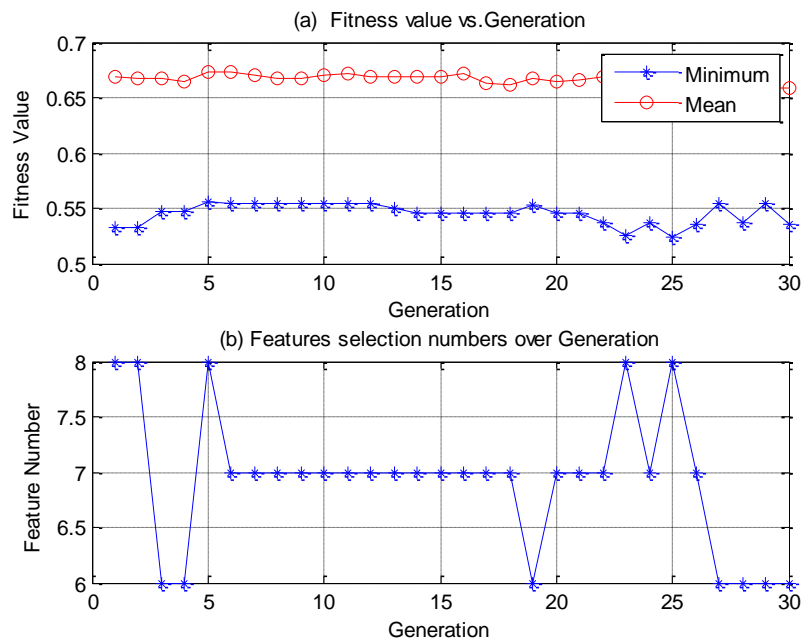
Figure 7.6 Minimum and mean fitness function with GA-PNN features selection from the time domain (Run 4)
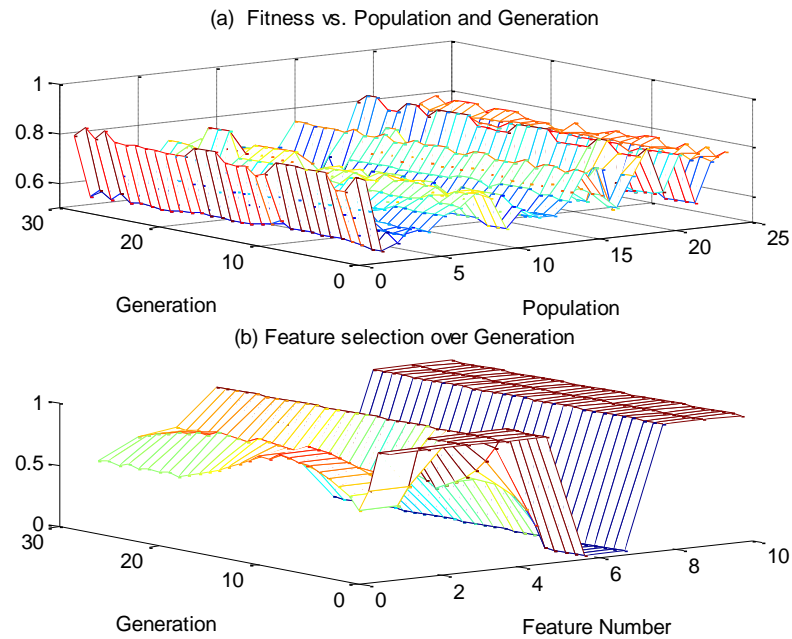
Figure 7.7 Fitness value and GA-PNN features selection from the time domain (Run 4)

The mean fitness value and the minimum fitness value for run 2 (n=6, correct classification rate=47%) and run 4 (n=6, correct classification rate 46%) are shown in Figure 7.4 to Figure 7.7 above respectively. Mindful that a total of 10 features were used for each run but the GA opted the selected features to provide the best classification rates. Run 2 displays a stable mean fitness value with local minima available in the minimum fitness values. In comparison, run 4 displays a stable mean fitness value while the minimum fitness value tends to be stable initially, but then tends to display fluctuations as immediate minima and maxima. It is observed that GA was not able to select the number of features probably due to the small number of input variables. Although, it selected specific features such as kurtosis and entropy in each run with different input variables.

Table 7.5 below depicts the classification results obtained using PNN with feature selection using GA with a low mutation rate (0.001) and a low crossover (4). For this case a combination of 5, 6 and 7 features was utilised with the GA selection so as to achieve the best possible correct classification rate. The best correct classification rate achieved was still 48% using either 5 or 6 features while the lowest correct classification rate remained at 46% using all the other combinations. The default input features remained 10 and it was up to the GA to decipher the best possible classification rates through GA selection. Again, the GA was not able to optimise the features to ideally number, but it could determine the good features which repeated in each run such as LB, kurtosis and entropy.

Run NO	Selected features	Feature name	Correct rate (%)	Minimum fitness	Sigma
1	5	LB, Kurtosis, Peakfactor, Skewness, Entropy	48	0.5191	0.43
2	6	UB, LB, Kurtosis, Variance, Peakfactor, Entropy	48	0.5226	0.44
3	7	Max, Range, UB, LB, Peakfactor, Kurtosis, Entropy	46	0.5373	0.53
4	6	Range, LB, Kurtosis, Variance, Skewness, Entropy	46	0.5399	0.22
5	5	LB, Peakfactor, Skewness, Kurtosis, Entropy	46	0.5356	0.58

Table 7.5 Performance of GA-PNN classifier time domain feature selection

(mutation rate 0.001), Generation 30, population 24, crossover 4, chromosome 10
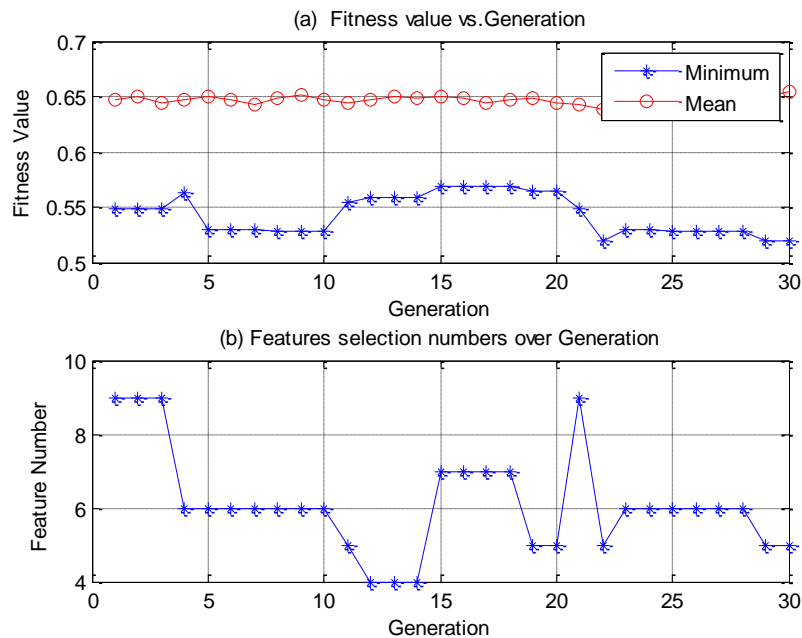
Figure 7.8 Minimum and mean fitness function with GA-PNN features selection from the time domain (Run 1)

THE USE OF ADVANCED SOFT COMPUTING
FOR MACHINERY CONDITION MONITORING

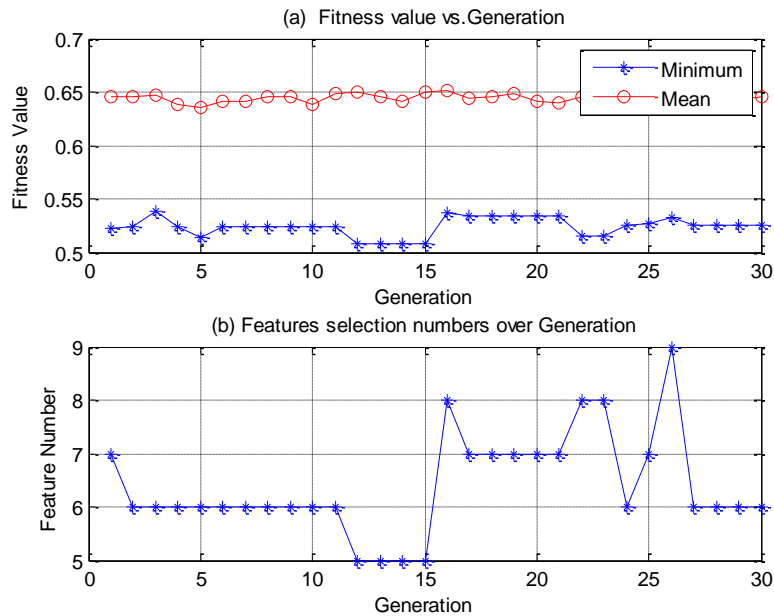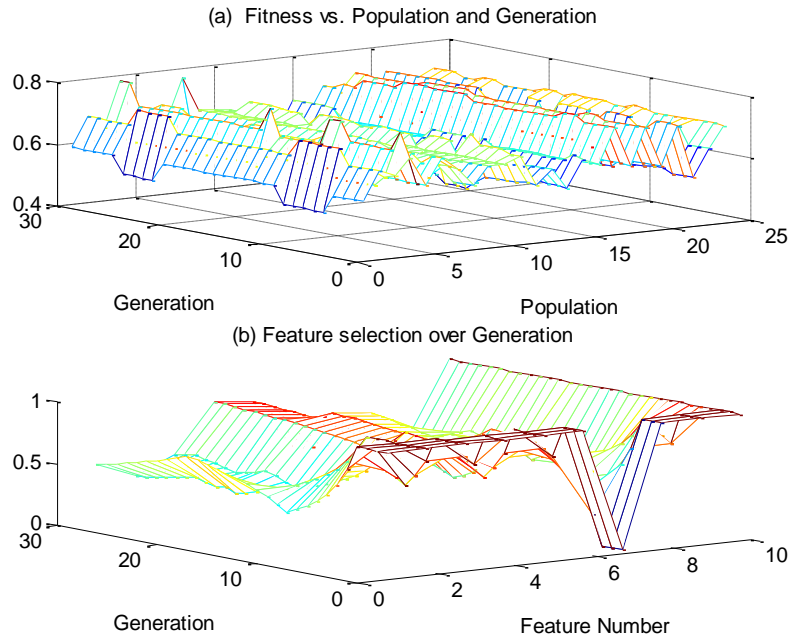Figure 7.9 Fitness value and GA-PNN features selection from the time domain (Run 1)
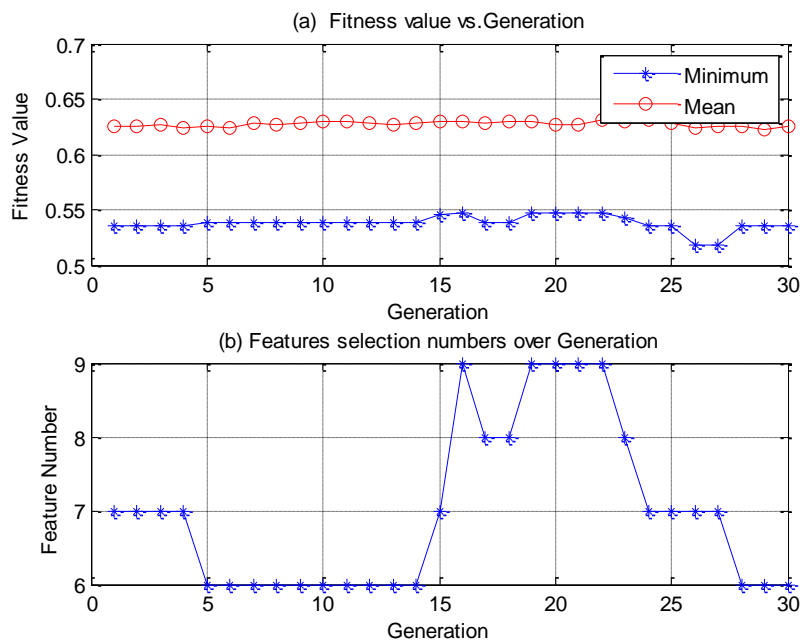
Figure 7.10 Minimum and mean fitness function with GA-PNN features selection from the time domain (Run 5)
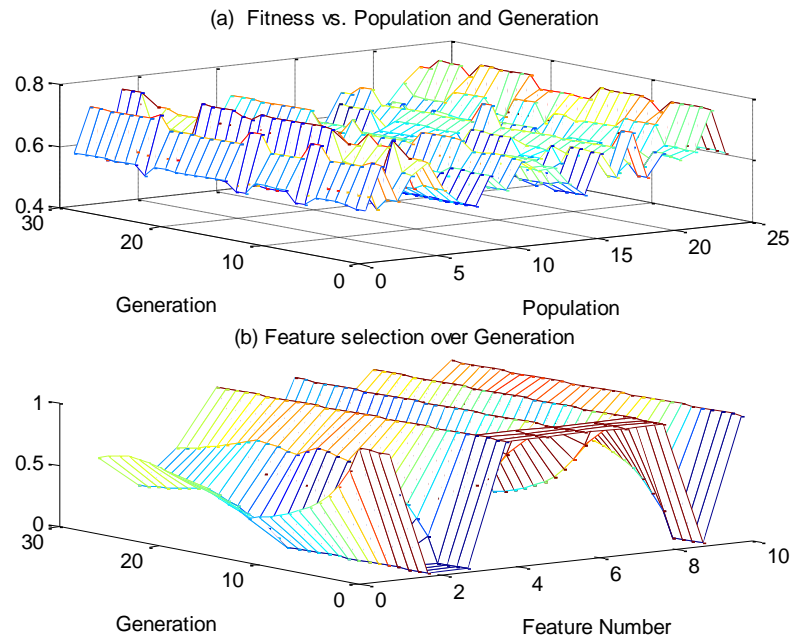
Figure 7.11 Fitness value and GA-PNN features selection from the time domain (Run 5)

The mean fitness value and the minimum fitness value for run 1 (n=5, correct classification rate=48%) and run 5 (n=5, correct classification rate 46%) are shown in Figure 7.8 to Figure 7.11 above respectively. Run 1 displays a stable mean fitness value with the presence of sizable fluctuation in the minimum fitness values. In comparison, run 5 displays a stable mean fitness value while the minimum fitness value tends to be stable initially but then tends to display fluctuations as a local minimum.

This behaviour is comparable to the use of high mutation rates and high crossover when the results were similar. The significant variation of the mutation rate and crossover seem to have little effect on the fitness of the generation during GA selection.

7.4.2.2   *Performance of GA-PNNs with Features from the Envelope Spectrum*

Table 7.6 below depicts the classification results obtained using PNN with feature selection using GA with a high mutation rate (0.7) and a high crossover (10). For this harmonic orders were selected from peaks in the frequency spectrum from the harmonics of the compressor's working frequency. The results in the table below show fluctuation in the achievement of correct classification rates with the highest correct classification rate of 99% (n=110) and the lowest correct classification rate of 93% (n=10). Sigma for the current runs was obtained through GA optimisation alone. Another noteworthy observation is the achievement of high correct classification rates following n=30 and higher.

No of input features	Harmonic order selected	Fitness	Sigma	Correct rate
10	2,3,4,5,6,7,8,10	0.0729	1.12	93%
20	2,3,6,7,8,9,12,14,20	0.0625	0.69	94%
30	1,2,4,6,7,8,9,10,11,12,13,14,15,16,18,20,23,25,26,27,29	0.0313	1.26	97%
40	1,2,3,4,5,6,7,9,14,16,17,20,26,28,29,31,32,33,34,35,36,38,40	0.0417	2.1	96%
50	3,5,6,8,10,11,12,13,17,18,20,22,26,28,29,30,31,35,37,39,40,42, 45,47,50	0.0625	2.14	94%
70	2,3,4,5,6,8,9,10,15,20,21,22,23,27,30,32,33,36,37,39,40,41,43, 45,46,48,49,50,52,53,57,59,61,63,64,65,67,68,69,70	0.0208	0.92	98%
90	7,9,15,16,17,18,19,20,21,24,26,27,28,30,35,36,37,39,40,41,47, 48,49,51,52,53,55,57,58,60,61,62,63,66,68,69,70,71,72,75,79,8 1,82,84,85,86,88	0.0313	2.83	97%
110	2,6,8,9,10,14,15,16,17,19,20,24,25,33,35,37,38,39,40,42,46,48, 50,51,52,55,58,61,65,66,68,69,70,75,76,78,79,80,81,82,83,85,8 6,87,88,89,90,92,93,94,95,98,99,102,103,106,108	0.0104	3.91	99%
120	2,5,6,7,8,9,11,12,13,14,15,17,18,20,22,23,24,25,26,27,2 8,33,36,38,39,40,43,44,45,46,47,48,52,53,59,69,7,79,80, 81,82,83,84,85,86,87,89,91,93,94,95,96,99,100,101,105, 107,108,113,114,115,116,117	0.0208	4.39	98%

Table 7.6 Performance of GA-PNN classifier using features in the envelope spectrum, generation no=30, population=40,crossovor=10,mutation=0.7

Figure 7.12 Minimum and Mean fitness function with GA-PNN features selection from the envelope spectrum with 10 input harmonics

Figure 7.13 Fitness value and GA-PNN features selection from the envelope spectrum with 10 input harmonics

Figure 7.14 Minimum and mean fitness function with GA-PNN features selection from the envelope spectrum with 50 input harmonics

Figure 7.15 Fitness value and GA-PNN features selection from the envelope spectrum with 50 input harmonics
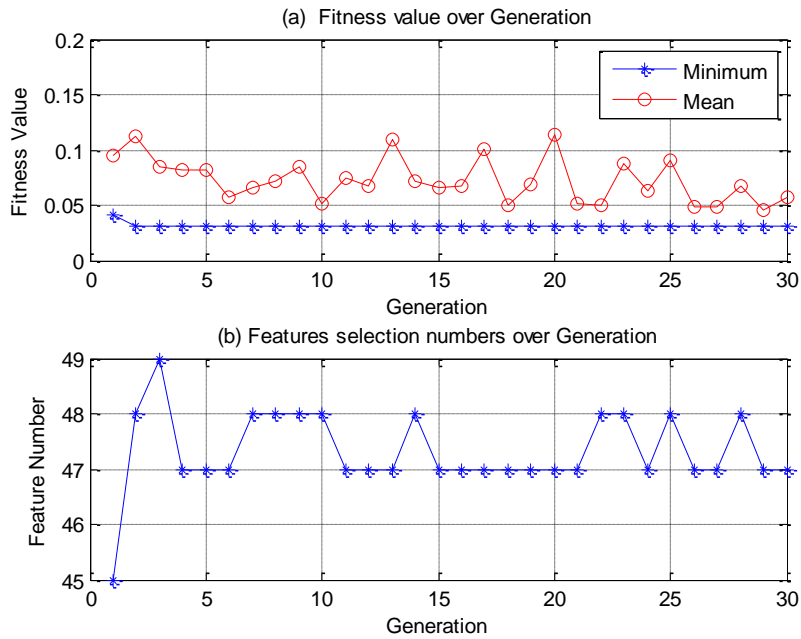
Figure 7.16 Minimum and mean fitness function with GA-PNN features selection from the envelope spectrum with 90 input harmonics
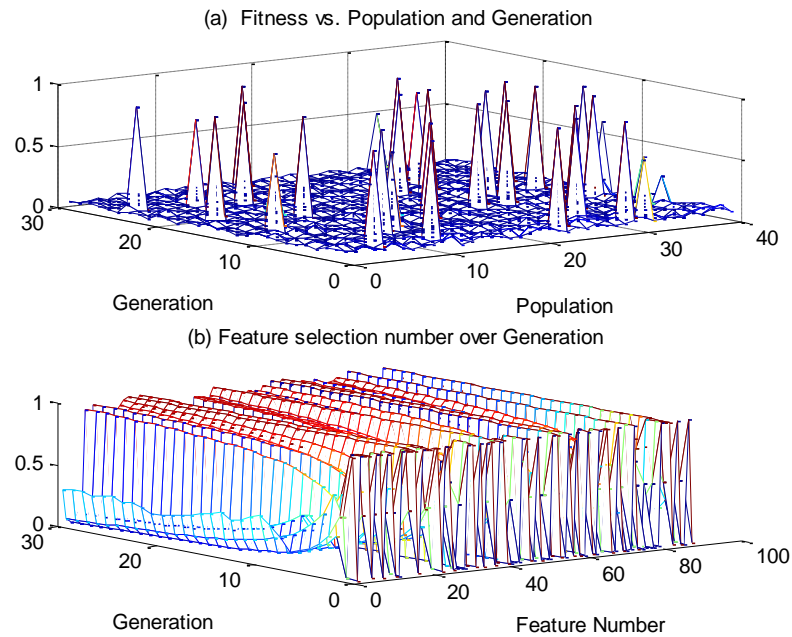
Figure 7.17 Fitness value and GA-PNN features selection from the envelope spectrum with 90 input harmonics

Comparison of the fitness values against generation in Figure 7.12 to Figure 7.14 above depict relative stability in the minimum fitness value compared to the mean fitness values. Figure 7.12 shows rapidly decreasing and then stabilised minimum and mean fitness values (n=10) while Figure 7.13 shows fluctuating mean fitness values with relatively stable minimum fitness values with a local maximum towards the end (n=50). In contrast, Figure 7.14 depicts a constantly fluctuating mean fitness value while the minimum fitness values are completely stable (n=90). This indicates that as the number of input features is increased, the minimum fitness values tend to stabilise while the mean fitness values tend to fluctuate a lot. Moreover, it should be noted that a significant reduction of the number of features is achieved by GA optimisation. As depicted in Table 7.4 and shown in Figure 7.13, Figure 7.15 and Figure 7.17, an input harmonics set of 10, 50 and 90 variables to features subsets with less than half features in all PNNs training. Moreover, the features selected in each run with different input harmonics can be seen. This will allows the online implementation of PNN to be more efficient.

Table 7.7 below depicts the classification results obtained using PNN with feature selection using GA with a low mutation rate (0.001) and a low crossover (2). The results in the table below show fluctuation in the achievement of correct classification rates with the highest correct classification rate of 99% (n=90) and the lowest correct classification rate of 92% (n=10). Comparable to before, high correct classification rates were achieved after n=20 with an exception at n=50.

No of input features	Harmonic order Selected	Fitness	Sigma	Correct rate
10	2,3,4,6,7,8,10	0.0833	0.79	92%
20	1,3,4,5,6,7,8,11,12,13,15,20	0.0625	0.62	97%
30	1,2,4,8,9,10,11,13,14,20,21,22,23,25,26,27,28,29	0.0313	0.89	96%
40	3,4,6,7,8,9,10,12,13,14,15,18,20,23,26,28,29,30,35, 38,40	0.0208	1.37	98%
50	2,3,4,6,9,10,12,13,15,16,17,20,21,22,23,25,26,27,28, 29,30,31,32,35,36,37,38,45,46,48	0.0729	1.02	93%
70	1,2,4,5,6,7,9,11,13,15,18,20,22,24,25,28,29,30,36, 39,40,41,42,43,46,48,50,54,56,57,59,62,64,66,68,69	0.0208	2.9	98%
90	1,3,5,7,10,11,14,15,17,18,19,20,23,24,26,27,28,34,3 5,36,37,39,40,42,44,45,48,52,53,55,56,59,60,63,66, 67,69,72,78,79,80,81,82,83,84,90	0.0104	2.64	99%
110	2,6,7,11,12,13,19,22,23,27,28,30,31,34,37,38,41,42, 44,46,48,53,54,55,57,63,65,66,74,77,78,79,81,82,83, 84,85,86,87,89,90,96,99,102,106,107,110	0.0208	2.61	98%
120	4,5,6,7,9,11,12,13,15,17,18,20,21,23,24,25,28,29,31, 35,36,37,38,40,42,44,46,47,50,52,53,54,56,57,58,61, 62,66,67,69,71,72,74,75,76,77,78,81,83,84,88,91,92, 94,96,97,98,102,105,107,109,110,111,113,116,117, 118,119,120	0.0208	3.83	98%

Table 7.7 Performance of GA-PNN classifier using features in the envelope spectrum, Generation no=30, population=40,crossovor=2,mutation=0.001
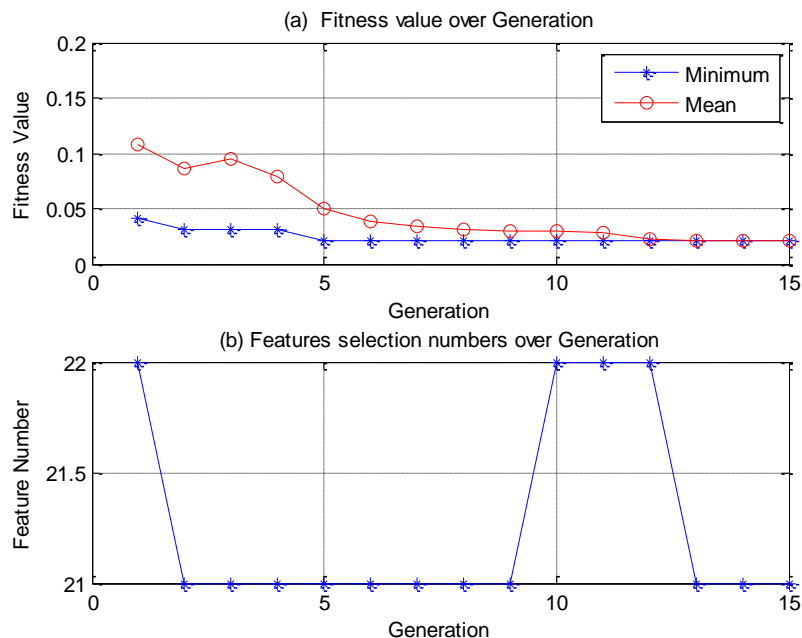
Figure 7.18 Minimum and mean fitness function with GA-PNN features from the envelope spectrum with 40 input harmonics
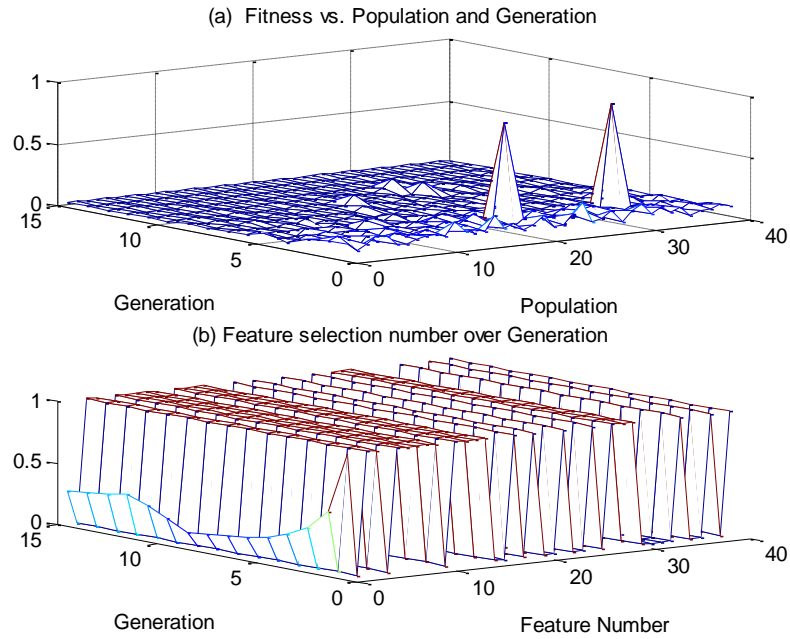
Figure 7.19 Fitness value and GA-PNN features selection from the envelope spectrum with 40 input harmonics
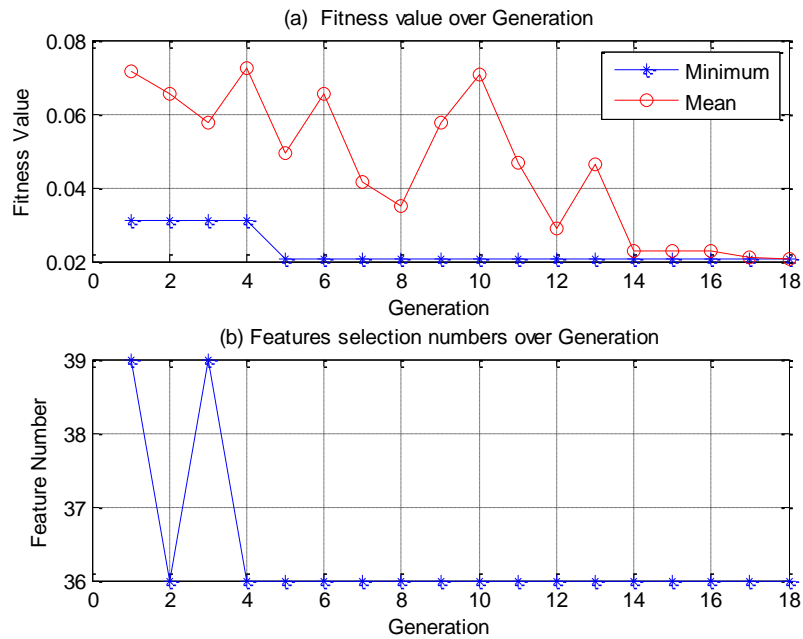
Figure 7.20 Minimum and mean fitness function with GA-PNN features selection from the envelope spectrum with 70 input harmonics
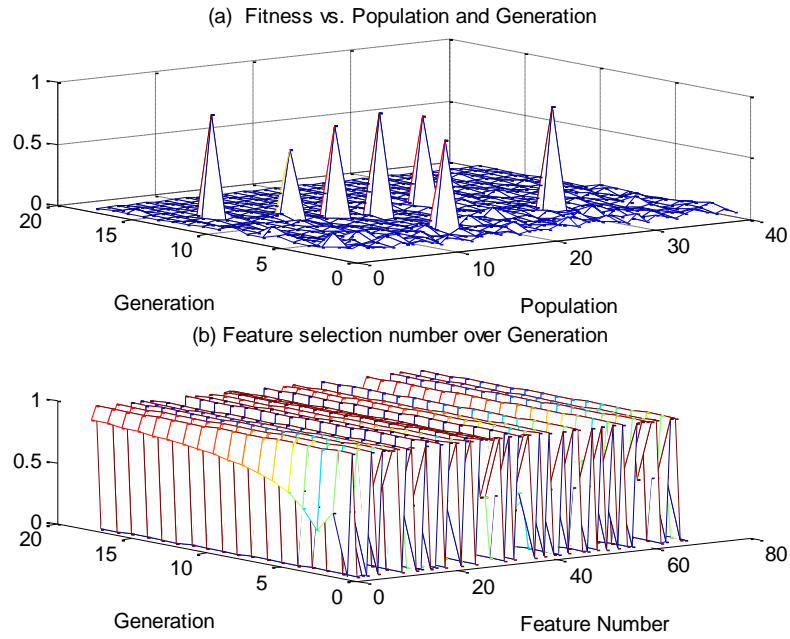
Figure 7.21 Fitness value and GA-PNN features selection from the envelope spectrum with 70 input harmonics
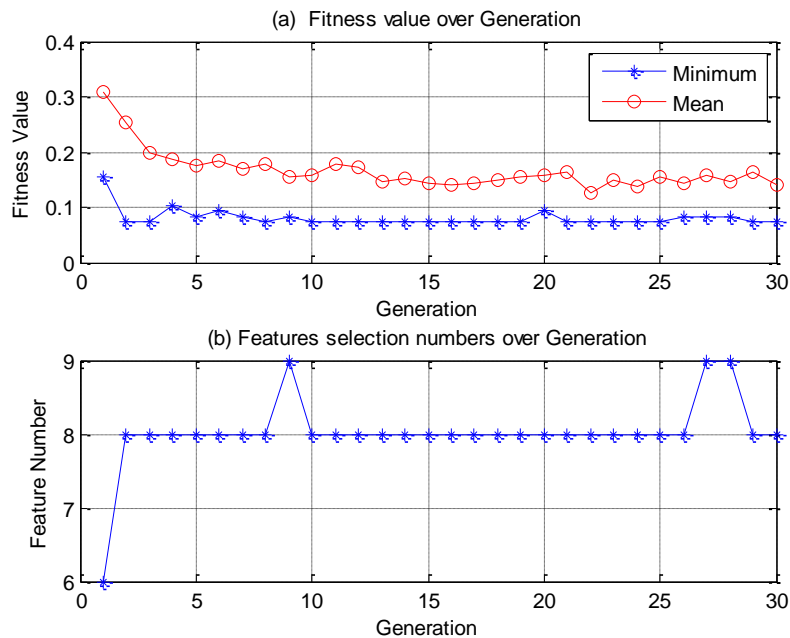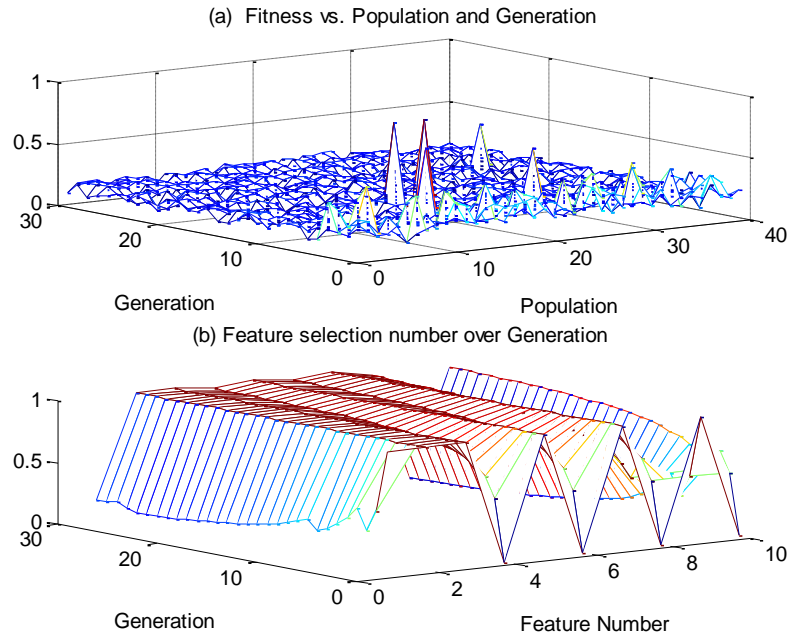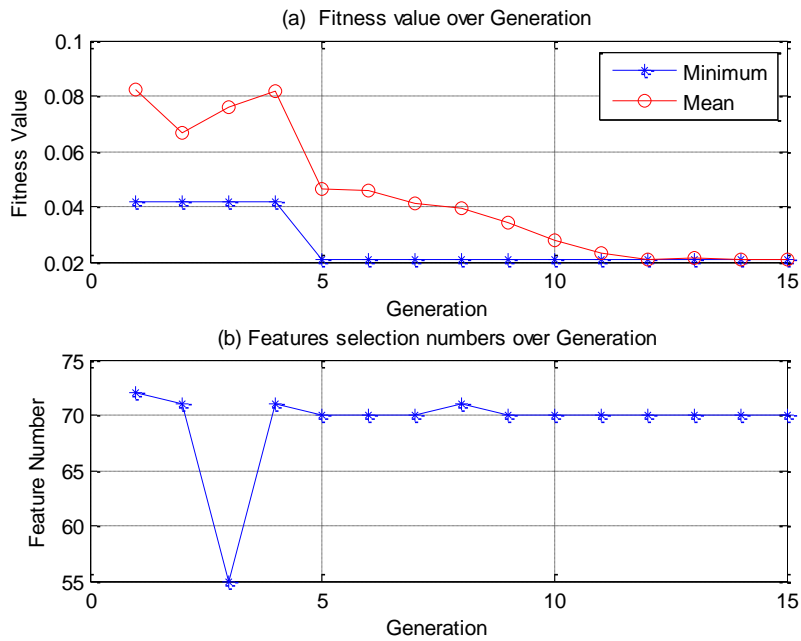

Figure 7.22 Minimum and mean fitness function with GA-PNN features selection from envelope spectrum with 120 input harmonics

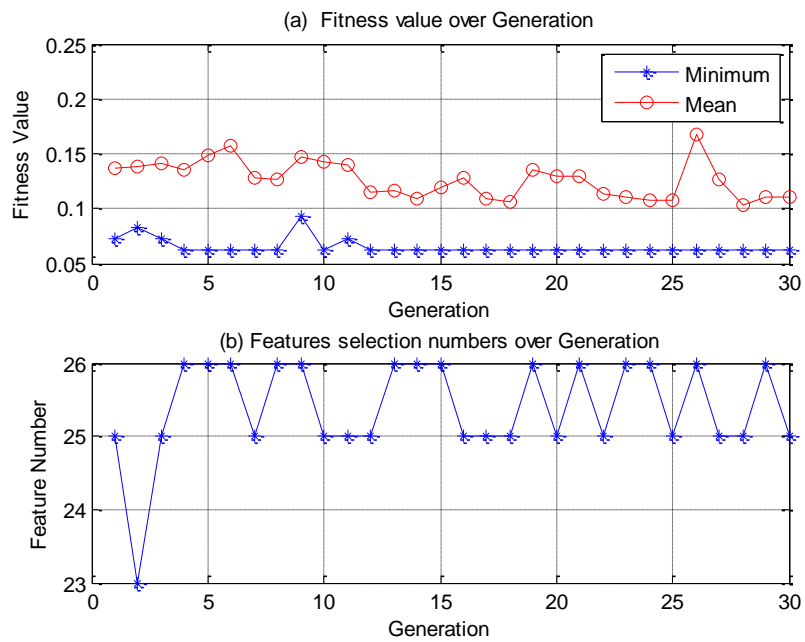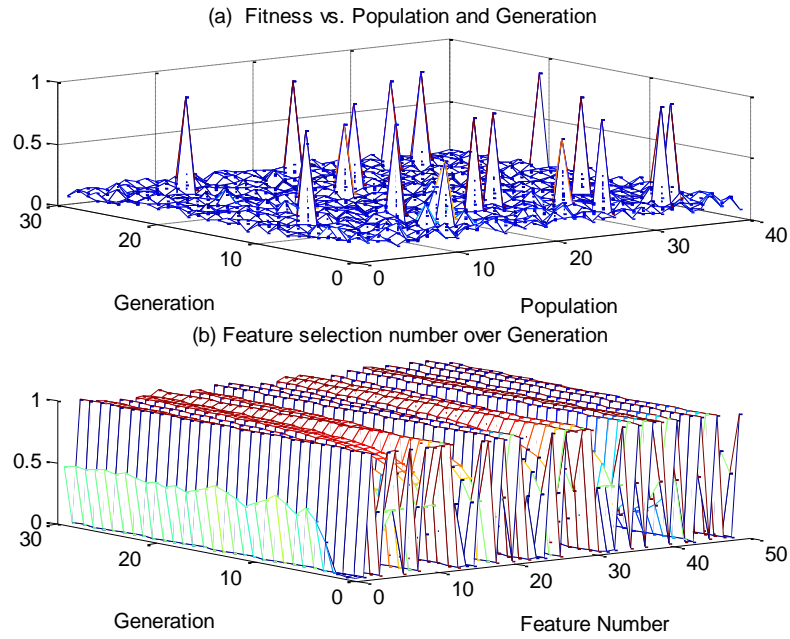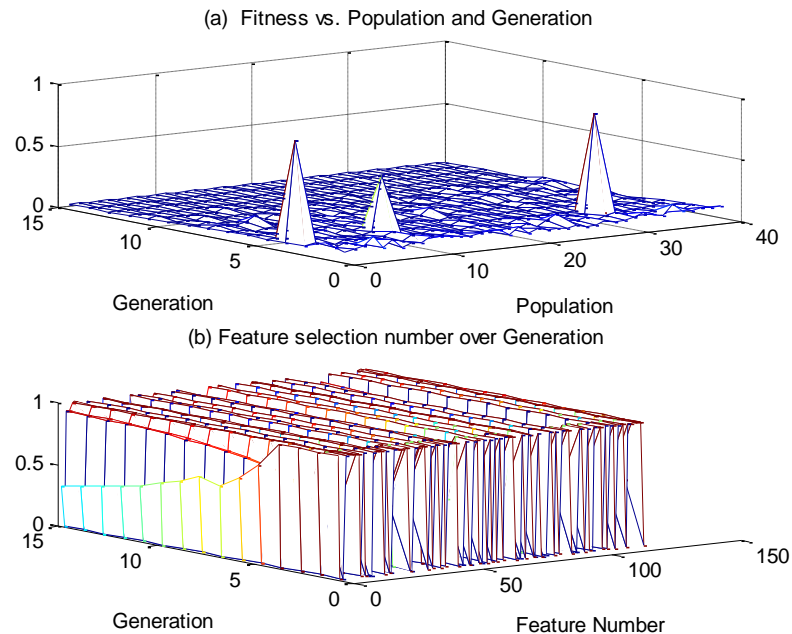**Figure 7.23 Fitness value and GA-PNN features selection from the envelope spectrum with 120 input harmonics**

Comparison of the fitness values against generation in Figure 7.18, Figure 7.20 and Figure 7.22 above depict convergence between the minimum and the mean fitness values. Figure 7.19 shows a gradual convergence between minimum and mean fitness values over generation (n=40) unlike Figure 7.20 where the convergence is late and the mean fitness values tend to fluctuate a lot before convergence (n=70). In comparison, Figure 7.22 depicts relatively gradual convergence between the mean and minimum fitness values with one local maximum for each (n=120). In addition, the number of features selected in this case seems the same features selected in previous case with different parameters.

7.4.3 GA System Verification with Feature from Time domain and Envelope Spectrum

In order to verify the robustness of the tested systems, GA system verification was carried out for both time domain and envelope spectrum. The concept is that any one input feature can produce a classification when tested with the desired system. For time domain with GA, feature 10 (entropy) was provided to the system in order to ascertain system response. Figure 7.24 below depicts the situation when entropy is provided as an input to the system allowing correct classification of fault and selected via GA as seen in Figure 7.25 below.
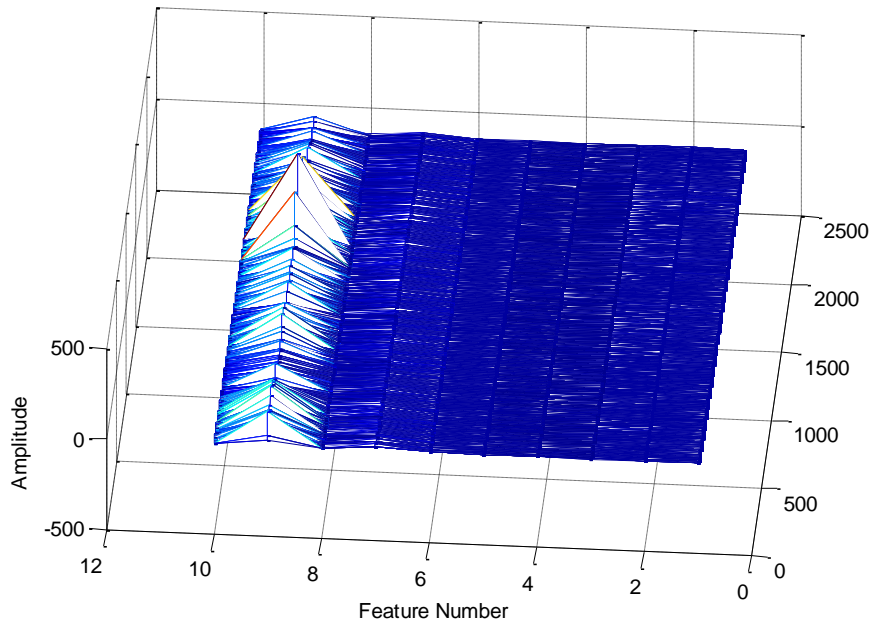
Figure 7.24 PNN feature extraction via PNN and time domain
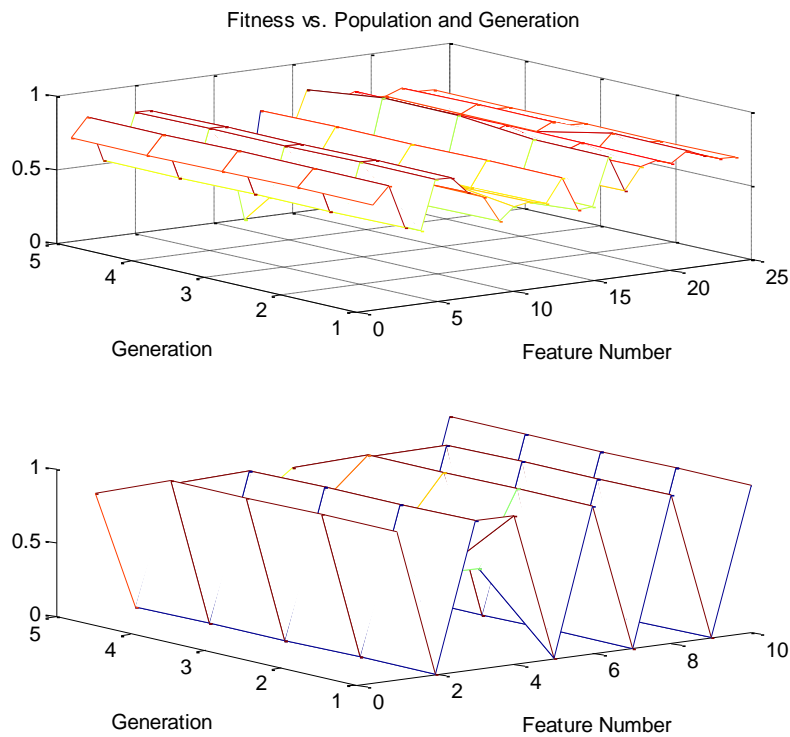
Figure 7.25 GA-PNN features selection from the time domain

In a comparable manner, verification was carried out for GA-PNN features selection from the envelope spectrum using (feature number 10) as a system input. This is depicted in Figure 7.26 below while the results of correct classification are shown in Figure 7.27 below where the feature number 10 has been selected via GA.

This proves that the system being used for both time domain and envelope spectrum with GA-PNN feature selection is robust and is performing as expected.
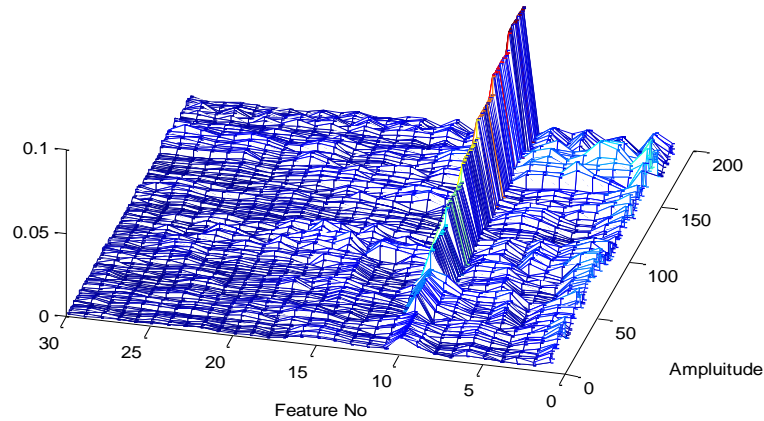
Figure 7.26 PNN feature extraction via PNN and envelope spectrum
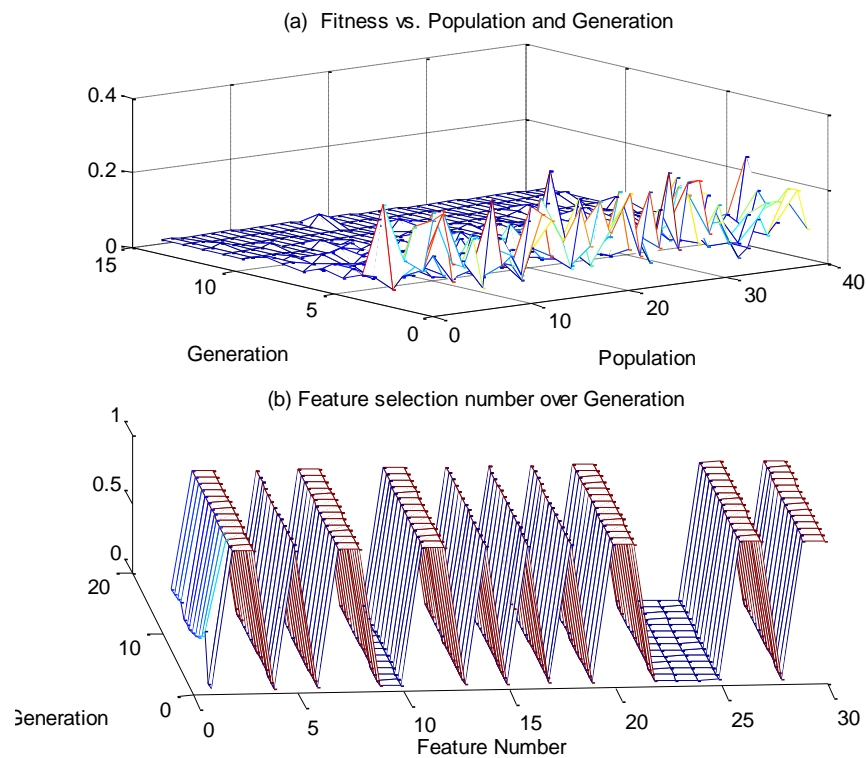
Figure 7.27 Fitness value with GA-PNN features selection from the envelope spectrum with 30 input harmonics

7.5   GA-PNN Discussion

The results shown above make it clear that GA results tend to vary as the parameters are varied. For example, when the crossover and mutation are small, the GA results end up in local minima but when the crossover and mutation are larger, the GA results are different. This

tends to indicate that GA needs to be run a few times to discover a convergent behavior for the system being studied. The system needs to be such that it need not be quick enough to end up with a local minimum and not slow enough to make optimal solution achievement a great challenge. In essence, this also indicates that the results need to fluctuate over several generations before a stable character is reached.

The features used in this study have been selected to produce significant separation. Features that were not selected tended to produce poor separation. For instance, peak value was not selected since it is not truly representative as there is large variance between various data sets. However, when RMS is selected, it is representative since the variance between data sets is sizable yet not too significant.

However, full classification could not be reached since the information or the candidate features are not sufficient enough to produce a full separation. Moreover, some of the data presents outliers that are hard to classify resulting in lack of full classification.

On another note, it is clear from the results of the current study that time domain GA-PNN feature selection is far inferior to GA-PNN feature selection through envelope spectrum. The best correct classification rate achieved through the former is only 48% compared to 99% for the latter. This speaks volumes for the efficacy of classification, achieved through either system. More research needs to be conducted in GA-PNN feature selection through envelope spectrum in order to deal with surfacing issues such as full classification achievement.

7.6   Summary

It has been conclusively shown that GA results will tend to vary as the parameters in use are varied. Hence, GA needs to be run at least a few times before a stable output can be achieved. Typically, results require fluctuation over a few generations to ensure a stable character. Moreover, the study used only features that could provide significant separation in order to reduce any erroneous classifications. On another note, a full classification could not be achieved given that the features in use could not generate full separation and because outliers caused classification problems. The results of the study conclusively indicate that GA-PNN feature selection through Envelope analysis is far superior to all other methods tested especially methods relying on time domain feature selection.

Chapter 8

Fault Identification and Diagnosis for Reciprocating Compressors Based on Support Vector Machined and Genetic Algorithms

This chapter focuses on the development of an advanced signal classifier for a RC using vibration signals. SVM have been applied, trained and tested for feature extraction and fault classification.

The results show that the model behaves well, and classification rate accuracy is up to 100% for binary classes. However, the SVMs for multi classification using features from the envelope spectrum returned 86.49% correct classification rate from two input features as the maximum rate from various input features. In contrast, when GA-SVM classifiers based on features in the envelope spectrum were used, the best classification rate achieved was 88% with 18, 20 and 25 input features. Also the GA optimized the features for SVM in envelope spectrum and the results show that one feature has been selected from 18, 20 and 25 input features and 10 features from 40 input features. Finally, it discusses how compressor faults could be detected and diagnosed using those approaches.

8.1   Introduction

In this study SVMs have been applied to a real compressor with single and multiple faults. It has been claimed that SVMs have four important advantages over the more traditional ANN. First and most important, is that SVM training uses the powerful mathematical technique of global optimized solutions and so has largely eliminated a major irritant of ANNs: convergence to local maxima and minima (Rychetsky, 2001). Second the simple geometric interpretation available for SVMs has proved very useful in extending its application to new areas and theoretically can give a sparse solution – that is the solution for the lowest number of entries (V. Vapnik, 1999). Third, during training, the SVM uses structural risk minimization which permits the software designer to allow for sparseness of data and which can lead to a better performance for SVMs than ANNs (Ghate, 2009). Fourthly, it has become clear that SVM is relatively very efficient when dealing with large classification problems (very large feature spaces), because the process of linearization means that the number of dimensions is less important with SVMs than with conventional classifiers (A. Widodo, Yang, B-S. , 2007). However, it has also been pointed out that SVMs have a number of less satisfactory features: limited speed both in training and testing, extensive memory requirements, the solutions while geometrically simple can be algebraically complex, and the design of SVMs is not yet anywhere near optimal (Suykens, 2003).

The SVM is a binary classifier it compares only two things at a time (Wei, 2002). This means that if there are N items to be compared there will $N * (N - 1)/2$ comparisons. Thus, in a real situation there will usually be will huge number of comparisons to be made. This is made worse by the parallel necessity to miss nothing of consequence when taking measurements and to ensure all possible useful features are recorded. But not all features are equally informative about the condition of the machine, and to increase the speed and accuracy of the classifier feature selection and extraction should be limited to those features useful for classification (Ghate, 2009; A. Widodo, Yang, B-S. , 2007).

This work makes use of SVM for fault diagnosis of a RC. SVM is initially developed for binary classification. Several techniques have been proposed for the SVM classifier to be extended to multiclass problems. Amongst them, the method: OVA have been used. A major improvement of the SVM performance can be achieved by the proper selection of the classifier parameters. To obtain the optimal parameter values for the SVM classifiers, the GA, as a well-known optimization solver is used. The result indicated improved classification performance.

8.2   SVM and GA-SVM Results and Discussion

8.2.1 Binary Classification

According to the different faults on the RC, these can be divided into seven types which are discharge valve leakage, suction valve leakage, inter cooler leakage, loose belt, discharge valve leakage and suction valve leakage combined fault, Suction valve leakage with Inter Cooler Leakage combined fault and discharge valve leakage with inter cooler leakage combined fault respectively. Because SVM is a binary classifier which differentiates two types of things at time. In this part of the work the SVM classifier was constructed using the one-versus-all method to detect the fault only by the trained model with healthy condition and faulty condition for each iteration with optimal parameters. The performance and accuracy rate of the classifier model is then verified subsequently.

Table 8.1 presents classification results for binary class fault detection obtained with the SVMs using features extracted from the envelope spectrum. There were a number of peaks in the envelope spectrum and each one was a possible feature. In each table there is a column headed "number of features", the 2, 10, 15 or 120. The table includes performance of the SVM classifier with a binary class using features from the envelope spectrum.

No of input Harmonics from the Envelope spectrum	Classification success rate binary class
2	100%
7	100%
15	100%
20	100%
30	100%
45	100%
50	100%
60	100%
75	100%
85	100%
100	100%
120	100%

Table 8.1 Performance of SVM classifier: features from the envelope spectrum, binary classification

The results show in the Figure 8.1 that SVM classifier model has the ability to separate healthy class from other classes with 100% classification with two input features. SVM classifier has accuracy and good generalisation performance with high classification accuracy on all classes with different number of input features.
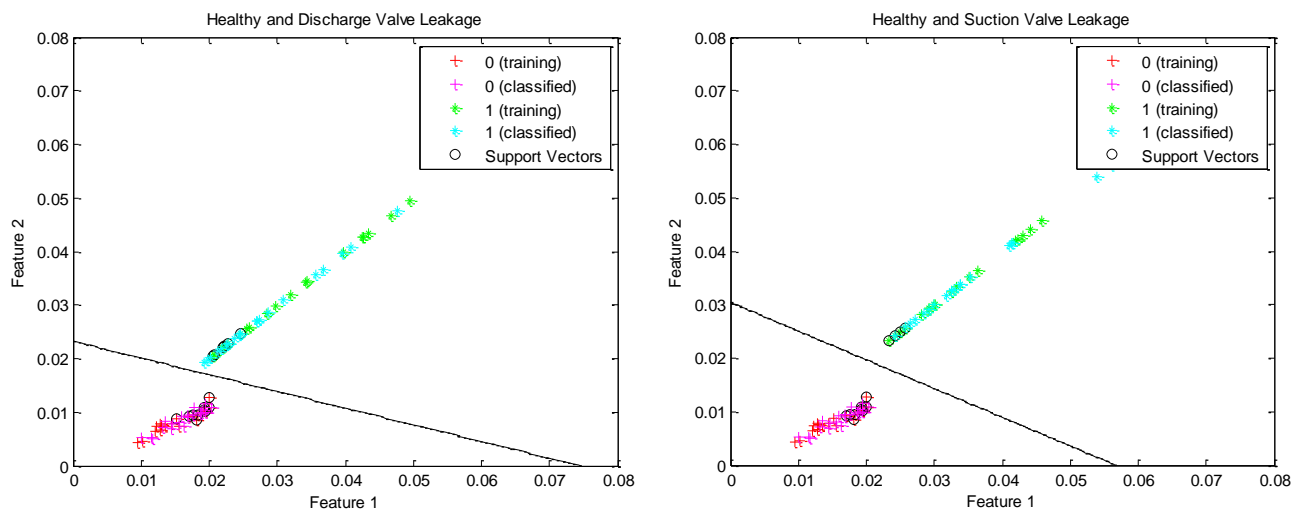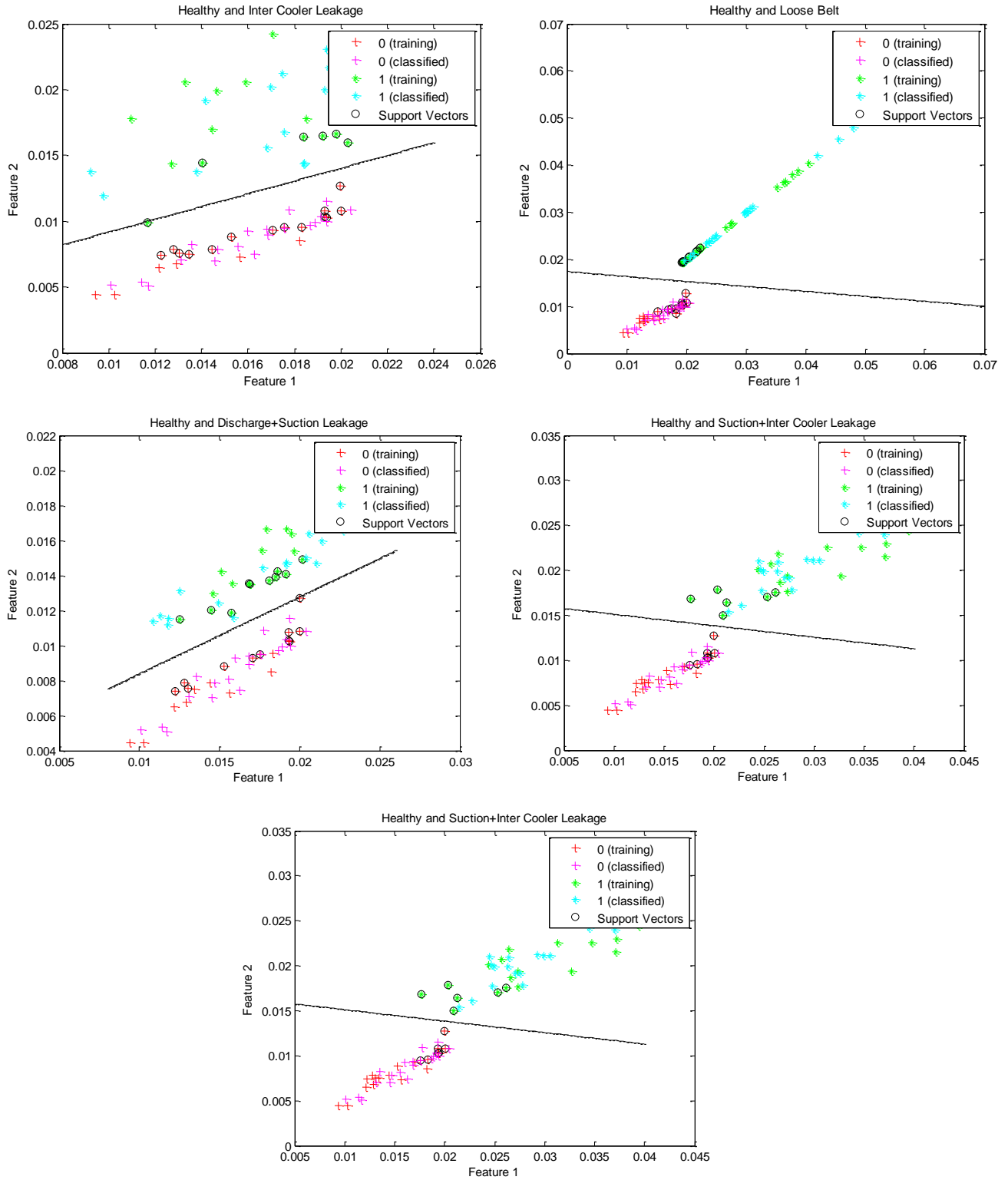
Figure 8.1 SVM binary classifier: Healthy class and other different condition classes with two input features from the envelope spectrum
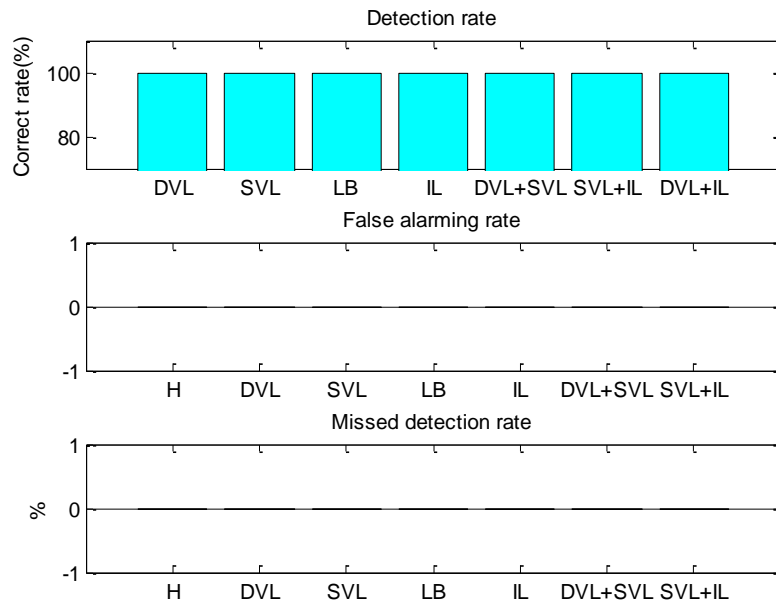
Figure 8.2 Classification Accuracy with Error Criterion for SVM binary classifier

8.2.2 Multi Class Classification

The results of this study which tested the "OAA" strategy with the classical decision making and can be shown in Table 8.2. The table shows results for the success rate of the SVM when using data from the envelope spectrum of the vibration signal when the problem was to identify the presence of eight cases which include the base line and seven faults. The SVM was most successful in providing successful classification with small numbers of input harmonics features, about two, and after reaching a maximum at about two features the success rate then decreases as more features were added. However, it is clear that the success rate of the SVM for this particular problem appears not to reach 100% accuracy rate. The classification error is determined by using ratio of correct classification and on the whole of training as showing in Figure 8.3.

No of input Harmonics from the Envelope spectrum	Classification success rate % Multiple class
2	86.49%
7	82.50%
15	79.69%
20	78.91%
30	78.75%
45	76.96%
50	76.41%
60	75.70%

Table 8.2 Results obtained with the "OAA" strategy on the SVM multi classifier
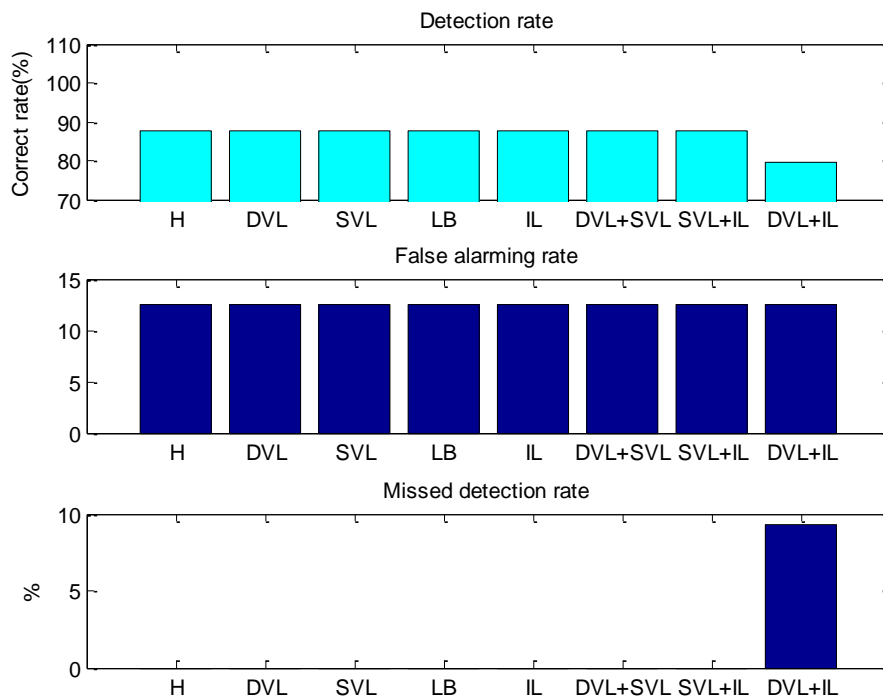
Figure 8.3 Classification Accuracy with Error Criterion for SVM multi classifier with two input features

8.3 GA-SVM Feature Selection

In this section, we describe the hybrid GA-SVM algorithm for carrying out the feature selection and classification. The GA model designed both for discovering the optimal features and for final features selection and classification. The SVM based classifier is used to ensure the fitness evaluation of each candidate feature subset.

Simple operations of the GA optimisation procedure are described in chapter seven at section 8.3. For the envelope spectrum the parameters of the GA were varied. The population size used was 40, and the number of generations was either 30 with mutation value of 0.01, see Table 8.3.

Using features from the envelope spectrum with GA optimization a classification success rate of 100% was not achieved. Interestingly, with a small number of harmonics it appears that can play an important role in successful classification of a fault, e.g. the highest accuracy rate classifications use harmonics between 20 and 30 (440Hz and 506Hz, respectively).

Total no of input features	No of selection features	Harmonic order selected	Classification rate	Fitness Value
18	1	17	88.0%	0.1328
20	1	14	88.0%	0.1250
25	1	22	88.0%	0.1250
30	8	6,13,15,24, 25,26,28,29	83.0%	0.1664
40	10	13,16,17,21,22,28,29, 33,34,38	83.3%	0.1711

Table 8.3 Performance of GA-SVM classifier

It can be seen from Table 8.3, that there are multiple solutions to the given problem. For example it can be seen that a classification rate of 88% is achievable using 18, 20 and 25 harmonics. With these input harmonics the GA selected one feature differently when the numbers of harmonics are different.

The result also shows the SVM is unsuccessful when larger numbers of features are used, but the number of features increase selected with larger numbers of input harmonics such as the classification rate was 83.30% with 40 harmonics and GA selected 10 features as be seen in Table 8.3.
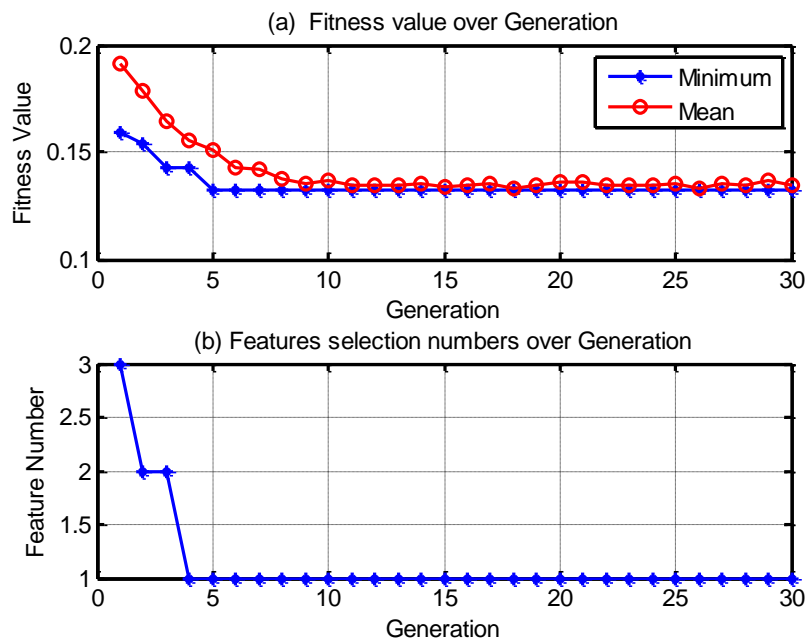
Figure 8.4 Feature selections with 18 input harmonics
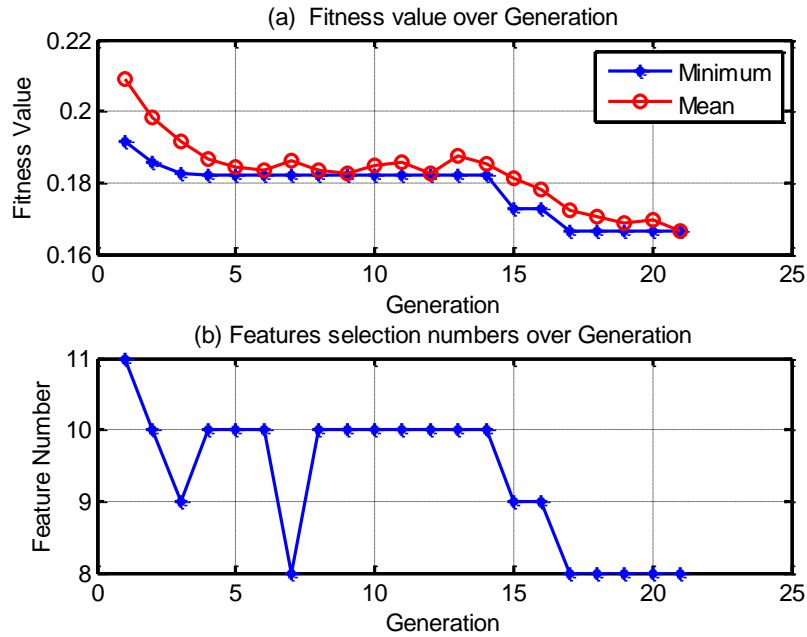
Figure 8.5 Feature selections with 30 input harmonics

Figure 8.4 and Figure 8.5 show the optimization procedure for selected steps of the features procedure. The graphs are a plot of the best and mean fitness values for the population as a function of the number of generation the population has evolved. Figure 8.6 and Figure 8.7 show the features selection with highest fitness function when the procedure is terminated.
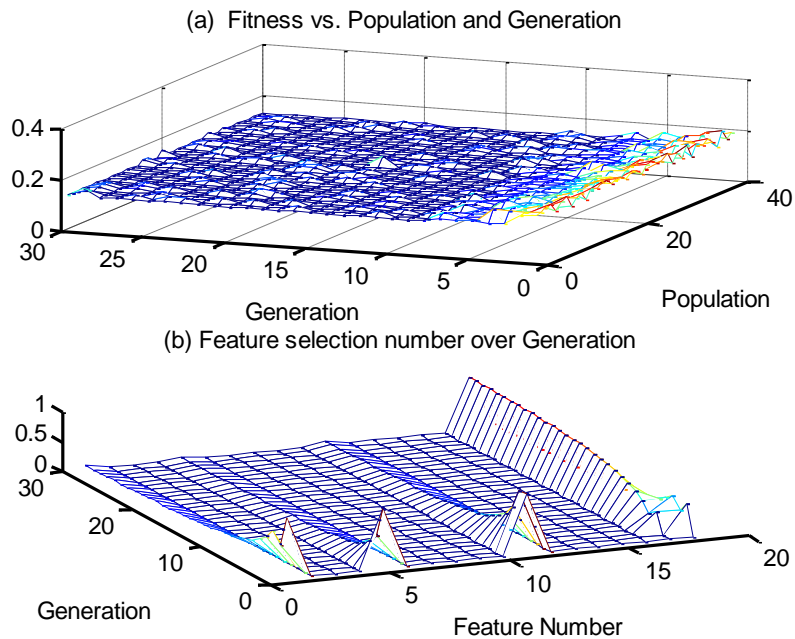
Figure 8.6 Fitness function and feature selection procedure with 18 input harmonics
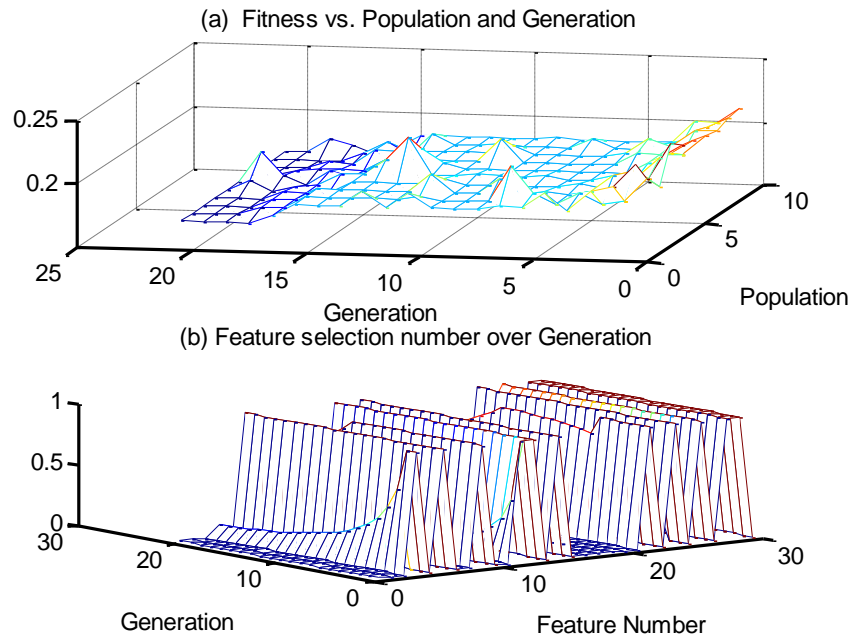
Figure 8.7 Fitness function and feature selection procedure with 30 input harmonics

8.4   Summary

SVM is a popular machine learning technique with a very high generalization capacity. Although SVMs are trained by a small data set, they can scan a very wide region. They are fast algorithms and do not need complex processes. In this study, the SVM classifier, when used with features extracted from the envelope spectrum for binary classes, gave a success rate of 100%. Briefly, the use of SVMs for binary classes as a powerful tool for fault detection is presented. Moreover, the performance of SVM has been found to be substantially better with the OAA strategy. With multiple classes the performance of the SVM classifier was satisfactory. The best success rate (88.00%) was achieved using harmonics between 18 and 25 of the peaks in the envelope spectrum.

The results show the potential application of GAs for the selection of effective features in machine condition detection.

Chapter 9

Fault Diagnosis of Reciprocating Compressors Using Relevance Vector Machines with A Genetic Algorithm Based on Vibration Data

This chapter begins by identifying the problem by comparing the results from SVM and RVM for fault classification. Then, it develops a one-against-one scheme with GA feature optimisation for applying this binary classifier to the compressors data, and examines the harmonics selected for classification to find the accuracy of the classifier in association with the physical supports. Finally, the performance of the multiclass multi-kernel mRVM is also explored for obtaining more efficient fault classification.

9.1   Introduction

ANNs, SVM are the popular classification methods (Shieh & Yang, 2008) . ANNs have the strong ability of non-linear mappings. The ANNs has been the most frequently used in the classification. However, the ANN has low training speed and is easy to produce local extremum. SVM is based on structural risk minimization principle, which is superior to empirical risk minimization principle used by ANNs. RVM is a kind of improved SVM, which incorporates probabilistic output and has a comparable generalization performance (Lima, Coelho, & Chagas, 2009). In order to improve the classification and select the optimal features which extracted from vibration envelope spectrum from RC; the performance of RVM, GA is applied to gain the suitable training parameters of RVM. Thus, a novel classification and feature selection method based on GA-RVM is discussed in the chapter.

9.2   RVM and GA-RVM Results and Discussion

As can be seen from Figure 9.1 and Figure 9.2 they are classifying samples of the same problem, the SVM has selected 24 important samples that lie on the class boundaries while the RVM has selected only 4 samples which are prototypical of the class they represent.
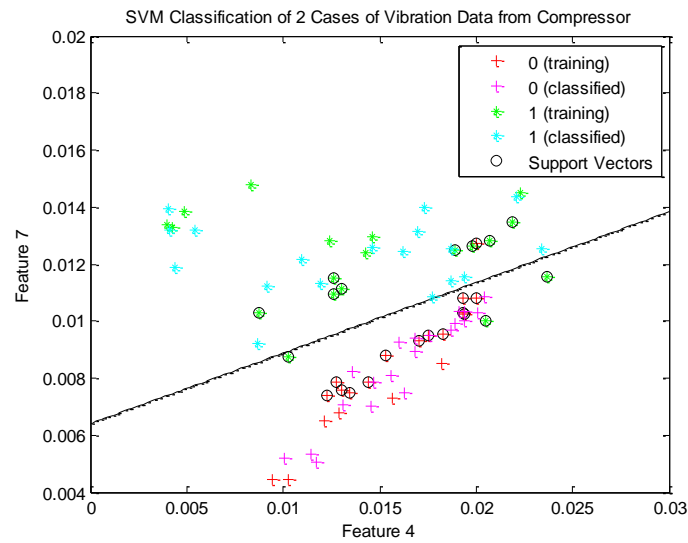
Figure 9.1 SVM binary classifier: Healthy class and DVL with two input features
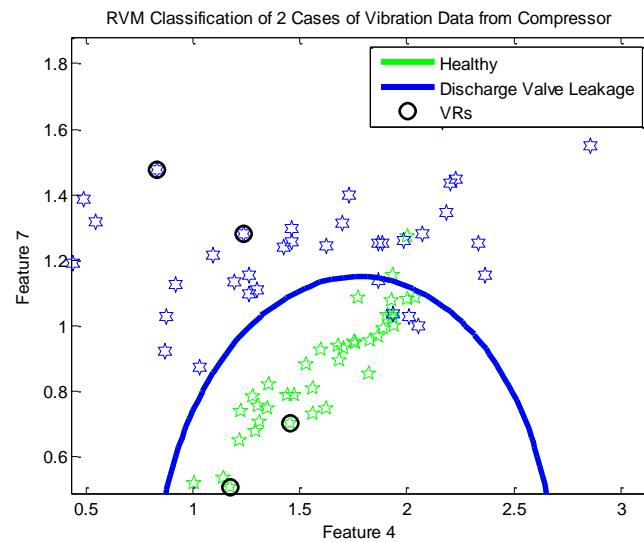
Figure 9.2 RVMs binary classifier: Healthy class and DVL with two input features

The most compelling feature of the RVM is that, it achieves comparable recognition accuracy to the SVM, yet provides a full predictive distribution, and also requires dramatically fewer kernel functions so that they consume much less test time which is the more important consideration in practice, such as on-line fault detection (D. Tzikas, Likas, & Galatsanos, 2007).

As shown in Table 9.1 there are eight classes of data sets corresponding to different healthy and faulty cases of the compressor.

Classes	Class Description
H	Healthy
DVL	Discharge Valve leakage
SVL	Suction Valve leakage
LB	Loose Drive Belt
IL	Intercooler Leakage
DVL+SVL	Discharge Valve leakage with Suction Valve leakage combined fault
SVL+IL	Suction Valve leakage with Intercooler combined fault
DVL+IL	Discharge Valve leakage with Intercooler combined fault

Table 9.1 Fault Situation in the Condition Process

Based on previous studies, the harmonics of compressor operation frequencies obtained from envelope spectra are the most effective features for differentiating between the cases. So the same harmonic components are adopted for RVM based classification.

A preliminary application of the RVM algorithm from M E Tipping (A. Tipping & Faul, 2002; M. E. Tipping, 2001) has found that the algorithms are not stable when the feature size (number of harmonics) is larger than 16. In particular, there are often errors with ill conditioned Hessian matrices due to poor conditioning of numbers in the dataset. In addition, a previous study has identified that the 1st harmonic component performs less well in separating the cases (Ahmed, Smith, Gu, & Ball, 2014). Therefore, only the harmonic components from 2 to 15 are used in a RVM application.

80 samples per case were collected, covering the rated operating pressure range from 70psi to 120psi (4.83 to 6.90 bar). In total the data matrix is 15x80x8. For RVM training a random sample of 40 is selected from each class. The remaining 40 data values being used to validate the trained RVMs.

In addition, OAO binary classifiers are used for the multiclass classification because the training sets are smaller and the problems to be learned are usually easier. Since the classes have less overlap if k is large and we need to evaluate the $k(k-1)/2$ classifiers, then the resulting system may be slower than the corresponding one-against-all RVMs (B.-S. Yang, Han, & Hwang, 2005).

For instance in this study, if $k = 8$, one needs to train 28 binary classifiers rather than 8 classifiers as in the method above. Although training time increases, the individual problems that need to be trained are significantly smaller. Furthermore, if the training algorithm scales are super linearly with the training set size, it is possible to save processing time. This is related to the runtime execution speed. To classify a test pattern all 28 binary classifiers need evaluating and classifying according to the classes which get the highest number of votes. A vote for a given class is defined as a classifier putting the pattern into that class. The individual classifiers, however, are usually smaller in size (they have fewer RVs) than they would be in the one-against-all approach.

9.2.1 RVM without GA Feature Selection

In this section classification results are presented from the straight use of RVMs without feature selection. However, as RVM is virtually a binary classifier, usually two classification schemes are adopted to perform multiclass problems. The first one is OAO and the second one is OAA. According previous research OAO outperforms OAA in that it can be trained more efficiently and can subsequently obtain better classification results (V. Jain, G. Pillai, & I. Gupta, 2011; B.-S. Yang, T. Han, et al., 2005). Therefore, this study only examines the performance of OAO.

The classification performances of the RVM based on OAO scheme (OAO-RVM) is shown in Figure 9.3 to examine the further performance the results is also presented in Table 9.2.

Class no	Cases	Input Harmonics	Classification Rate (%)
Overall	Overall	2-15	95.95
C_1	Healthy	2-15	90.00
C_2	DVL	2-15	100.00
C_3	SVL	2-15	97.50
C_4	LB	2-15	95.00
C_5	IL	2-15	95.00
C_6	DVL+SVL	2-15	100.00
C_7	SVL+IL	2-15	90.00
C_8	DVL+IL	2-15	100.00

Table 9.2 Average classification rates of the trained classifiers
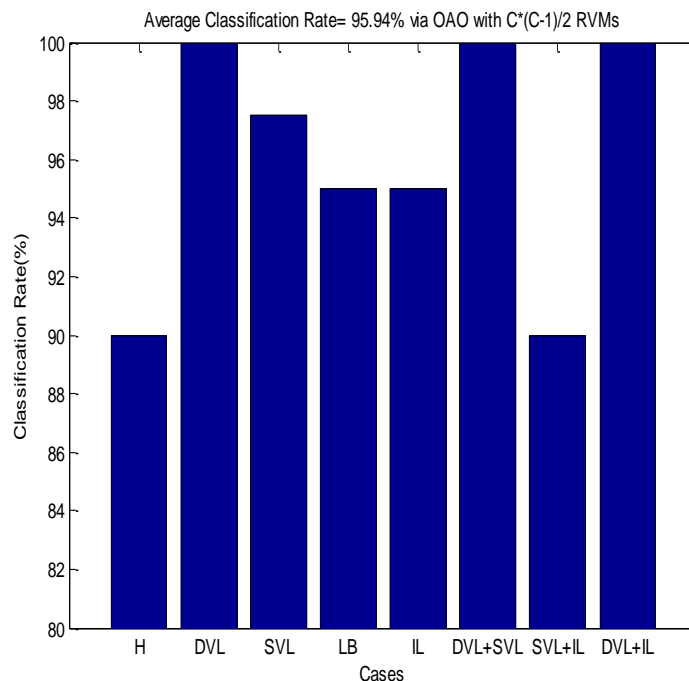
Figure 9.3 RVM–OAO Classification rates of different cases

Figure 9.3 shows the performances of RVM classification rate on extracted features of vibration signal (%) of RVM-OAO using Gaussian kernel for testing datasets for different classes. An overall accuracy of 95.95% was achieved using the technique developed.
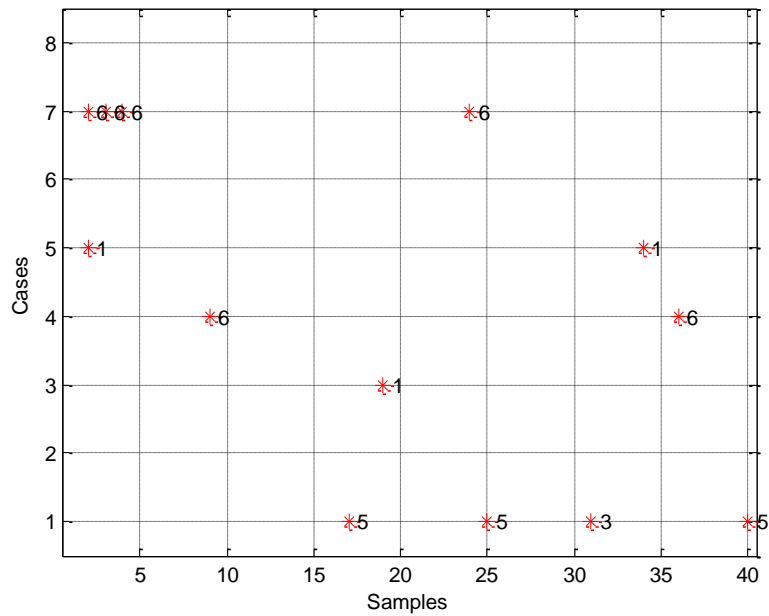
Figure 9.4 Misclassification rate of OAO

The errors occurred in classification of data between Class 1, Class 3, class 4, class 5 and 7 because these classes are close to each other, some of them having the same fault characteristic(s) as shown in Figure 9.4 and demonstrated in Table 9.3 with the 0% to 10% average error range. It is confirmed that the suggested RVM-OAO algorithm has high classification accuracy with small failure in some cases.

Class no	Cases	Misclassification classes	No of misclassification Samples	Misclassification Rate (%)
1	Healthy	IL, SVL	4	10
2	DVL	-	0	0
3	SVL	Healthy	1	2.5
4	LB	DVL+SVL	2	5
5	IL	Healthy	2	5
6	DVL+SVL	-	0	0
7	SVL+IL	DVL+SVL	4	10
8	DVL+IL	-	0	0

Table 9.3 Misclassification between classes

9.2.2 RVM with GA Feature Selection

A GA-based parameter method was used to automatically obtain the optimal features of the RVM classifier. The classification rate of the training data is the fitness function of the GA. Therefore, the optimum features are achieved when a minimum error is detected by the classifier or complete the generation. Also for the GA-RVM-OAO method, the population size is considered to be equal to 20. The initial range is taken to be within [0, 2] for all individuals.

In this section, two kinds of experiments were carried out. The first one had as objective the classification rate of accuracy. And the second had as an objective the output of feature selection subset with a high rate of classification.

Figure 9.5 shows the results of GA classification performance. It is remarkable the influence of the classification using GA-RVM-OAO approach, with an accuracy rate close to 97%.
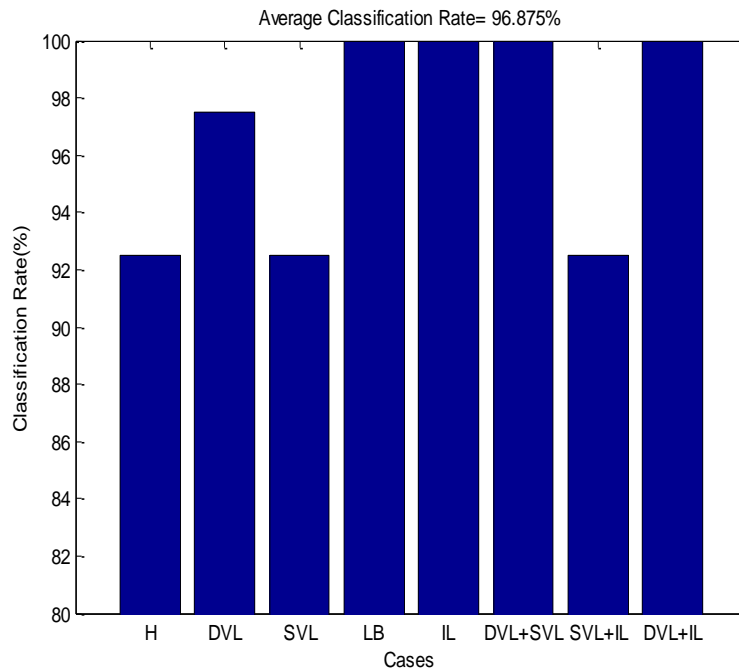
Figure 9.5 RVM–OAO classification rates of different cases based on GA-based parameter selections

Moreover, it should be noted that a significant reduction of the number of features is achieved by GA optimisation. As shown in Figure 9.6, an initial set of 15 variables is reduced to features subsets with no more than 10 features in all RVMs training. This allows the online implementation of RVM to be more efficient.

Furthermore, the relation between the final feature subset selected and the classifier used in the fitness function is also taken into account. Best results are achieved when using the same function in the classifier as in the feature selection and classification stage.
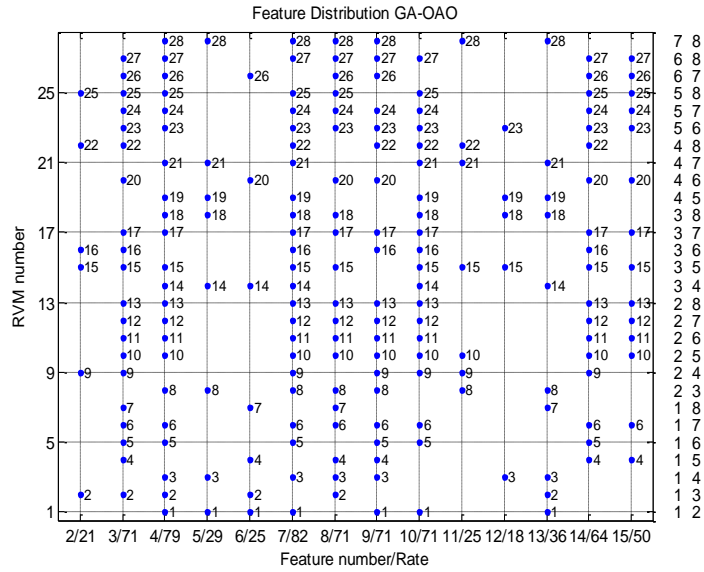
Figure 9.6 Features selection with GA-RVM

As shown in Figure 9.7, the proposed indexing method makes use of a knowledge-base, consisting of relations between features and concept classes. The relationships are not necessarily trivial, thus the key of this method lies in acquiring distinction. This figure demonstrates empirically that the set of relevant features is individual for each class. Indeed whilst there are features shared by all the classes there are other features which are entirely absent from a class. In particular, feature 11 is not selected for class 1, feature 12 is not used in class 2 while features 5, 11 and 13 are not selected in class number 6. On the other hand all classes utilise feature 10.
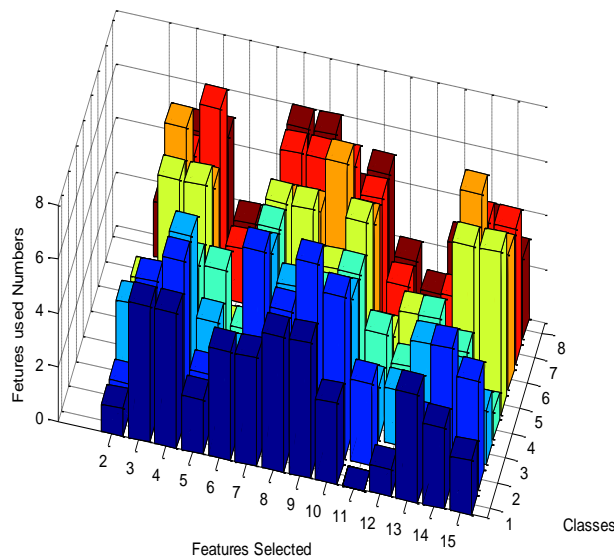
Figure 9.7 Relating features with classes

Figure 9.8 shows the mutual information between each feature and diagnosis. The estimation is based upon all available cases of the dataset. Clearly, features 4, and 7 stand out as the

most informative ones for the diagnosis of the RC, accounting for 85% and 80% of the total selected, respectively. Features 3, 8, 9, 10 and 14 are reasonably predictive of RC diagnosis with 70%. The remaining features number 2, 6, 11, 12 and 15 showed close to chance behaviour and ranged between 20% and 30%.
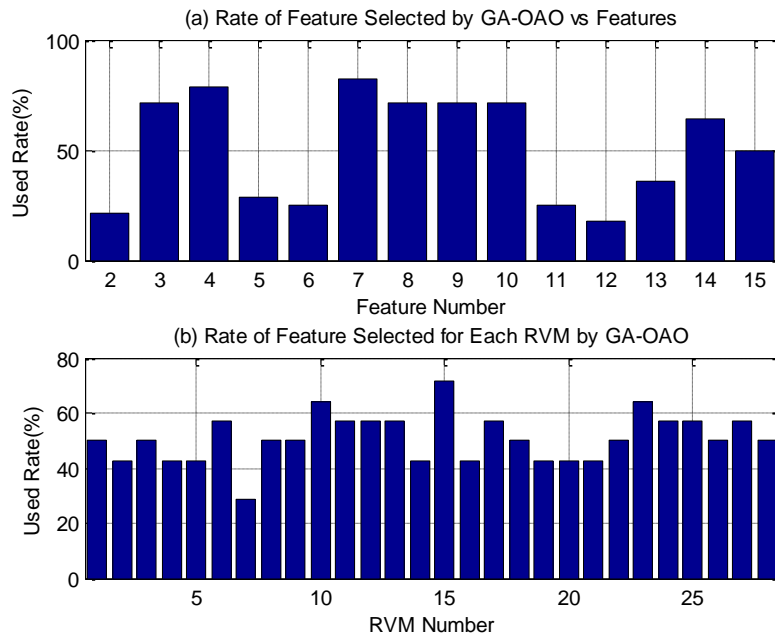
Figure 9.8 Performance of GA- OAO-RVMs for different number of selected features

Class No	Features													
	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_1	1	5	5	2	4	4	5	5	3	0	1	4	3	2
C_2	1	5	6	2	1	7	5	7	6	3	0	2	5	4
C_3	3	4	6	3	2	6	5	3	5	2	2	4	3	2
C_4	2	3	4	4	2	6	2	4	5	3	2	4	3	1
C_5	2	6	6	1	1	6	6	4	6	2	3	1	6	6
C_6	1	7	5	0	2	5	5	7	5	0	1	0	7	5
C_7	0	5	7	2	1	6	6	6	5	2	0	2	5	5
C_8	2	5	5	2	1	6	6	4	5	2	1	3	4	3

Table 9.4 Relating features with classes

From Table 9.4, it is apparent that the RVM-OAO based on GA techniques selected features very different from the ones selected in each case. The different selections affected the diagnosis performance to achieve separation of the classes. The RVM training utilised two important features 4 and 7: selected based on GA for all classes separation as the following subsets as displayed in Table 9.5.

Class	Features uses
C_1	3,4,6,7,8,9,13
C_2	3,4,7,8,9,14
C_3	3,4,7,8,10,13
C_4	4,5,7,9,10,13
C_5	3,4,7,8,9,14,15
C_6	3,4,7,8,8,10,14
C_7	3,4,7,8,9,10,14,15
C_8	3,4,7,8,9,10,14

Table 9.5 The most features using in each class

9.2.3 Multi-class Relevance Vector Machine Classification

To evaluate the performance of mRVM in classifying compressor vibration data, the mRVM model was trained for four classes of samples: healthy, intercooler leakage, discharge valve leakage with intercooler leakage combined fault and suction valve leakage with discharge valve leakage combined fault with two input parameters to the model. The generalisation performance and accuracy rate of the classifier model was then verified. The classification results of mRVM in Figure 9.9 shows that a mRVM classifier model with two input harmonics has good generalisation performance and high classification accuracy with low error. See Table 9.6.

State	Number of Samples	Correct classification	Accuracy	Error Rate
Healthy	40	40	100%	
IL	40	40	100%	0.0125%
SVL+IL	40	39	97.5%	
ICL+DVL	40	39	97.5%	
Overall			98.75%	

Table 9.6 Classification results for four classes with two harmonics

Similarly, this approach has also been applied for the defect diagnostics for more than 4-class cases, considering multiple defects with different numbers of input harmonics that ranged between 2 to 15 as explained in section 9.2.1.
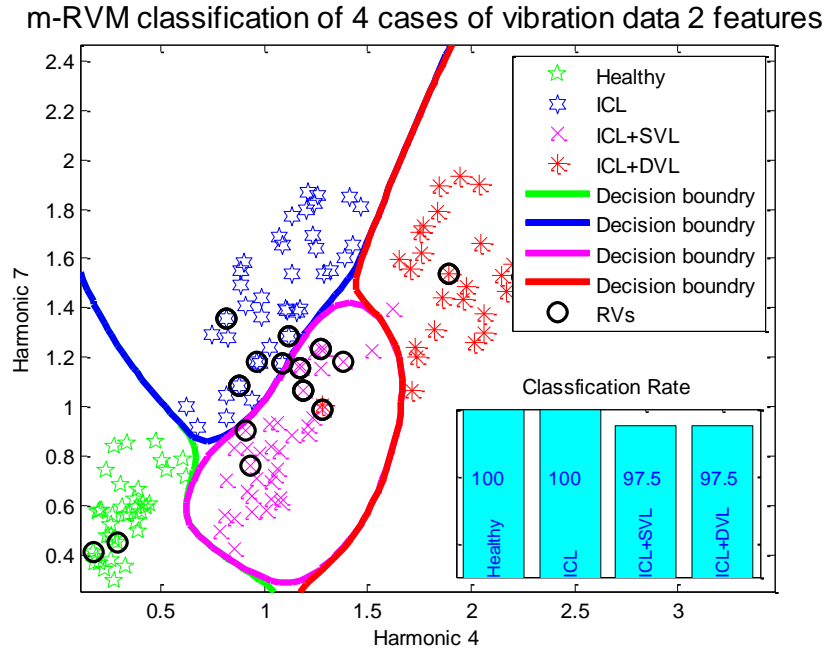
Figure 9.9 Classification results using mRVM four classes with two input harmonics

Classes	Harmonic Orders	Accuracy	Error Rate
1,4,6,7	4-15	98.75%	0.0125
1,4,5,7,8	4-15	86%	0.160
1,2,4,5,7,8	2-15	73.75%	0.2625
1,2,3,4,5	2-15	77.0%	0.23
1,2,3,4,5,6	2-15	50.84%	0.4917
1,2,3,4,5,6,7	2-15	59.29%	0.4071
1-8	4,5,6,10,15	24.16 %	0.3969
1-8	2-15	50%	0.50

Table 9.7 Classification results for different classes with different harmonics order

Table 9.7 shows results for the mRVM. The harmonic peaks in the spectral analysis provided a total of 14 possible features for use in classification of the seeded faults, the number of peaks in the spectrum which are used for classification. The first column refers to the number of input cases. The second column represents the number of features, ranging between 2 to 15. For example, 12 features are tested in the first row (between 4-15) and the peaks which gave the best classification rate were chosen. The correct classification rate was 98.75%, but as the number of features decreased the classification rate dropped quite sharply. The classification rates were 24.16%, obtained with 5 features and 8 cases as shown in row 7. This is due to the bad scaling of the type-II ML procedure with respect to the number of classes (Damoulas, Ying, Girolami, & Campbell, 2008; Ioannis Psorakis, Theodoros Damoulas, & Mark A Girolami, 2010b) and the dimensionality of the Hessian required for the Laplace approximation (Bishop, 2006; Psorakis, et al., 2010b).

Looking at Figure 9.9 it can be seen that some faults are easier to identify than others using the mRVM. For example there are large boundaries for healthy, intercooler leakage, discharge

valve leakage with intercooler leakage combined fault and suction valve leakage with discharge valve leakage combined fault so those four faults could be identified relatively easily. However, as the requirement is that the four faults be correctly identified, a relatively large number of peaks are required for high classification rates. Moreover, the algorithm is very sensitive in that it can only accept a restricted number of harmonics as discussed previously, and needs to improve in future work.

9.3   Summary

A procedure is presented for the detection and diagnosis of RCs using RVMs-OAO classifiers and RVMs-OAO with GA-based feature selection from the envelope spectrum of vibration signals. The selection of input features and the appropriate classifier parameters have been optimised using a GA-based approach.

The characteristics of vibration signals, obtained under normal operation and operation with various defects have been investigated. The classification accuracy of RVMs with GA was better than that of RVMs without GA. For RVMs without GA selection 95.95% classification success was achieved in overall test cases. The classification accuracy was 96.875% with selected features using GA. The results show the potential application of GAs for feature selection. This also opens up the potential use of optimised features and classifier parameters for real-time implementation leading to possible development of an automated machine CM and diagnostic system.

Table 9.4 and 9.5 summarise the importance of features 4 and 7 in explaining behaviour in all cases, each time being used predominantly more than any other features. As previously discussed, mRVMs have the capability to produce reliable classification rates through both reduced sample and computational times. The high importance of the two features 4 and 7 is ably demonstrated in Figure 9.9 where the four classes (C1, C5, C7 and C8) have a combined classification rate of 98.75% using these two features alone. Also, 100% classification success was achieved in the single fault cases.

Other features would seem to have particular relevance to a given fault being used repeatedly in that faults presence. For example, feature 9 has a strong association with the DVL cases, whilst the SVL relies heavily on features 8 and 10. The intercooler leak obviously emulates some of the characteristics of the other faults and in addition to the 'base' features (4 and 7) utilises the SVL set (8 and 10) along with features 3, 14 and 15.

Chapter 10

Fault Detection and Diagnosis Using Principal Component Analysis of Vibration Data From A Reciprocating Compressor

This chapter begins by developed the PCA model from the baseline. Then, it discusses how compressor faults could be detected using Q and T^2 Statistics. Finally, it is describing analytical techniques for fault diagnosis by using contributions plots.

10.1 Introduction

In CM, one problem is how to use huge databases of process measurements containing information about the state of the process. These real time databases are multivariate in nature i.e., many different variables are measured and recorded on a frequent basis, but with no further processing can be characterized as data rich but information poor.

PCA is a basic method in the framework of the multivariate analysis techniques. It has been successfully used in numerous areas including data compression, feature extraction, image processing, pattern recognition, signal analysis, and process monitoring.

The novelty of this work is used the PCA as an approach for feature space dimensionality reduction from vibration time domains statistical parameters and for detection the faults by using the T^2 or SPE from the PCs model using the baseline. Moreover, the contributions of the Q-plot and D-plot were applied in a way which allows it to be used with any latent variable component or regression model to find effective features for fault classification.

10.2 PCA Results and Discussion

10.2.1 PCA Model Development

Figure 10.1 shows the relative variance of the fourteen variables selected for PCA. It also shows that seven of these accounts for 99% of the variance, and this means that the subspace composed of those seven PCs contain enough information on the variation of the original features for it to be sufficient to detect the faults in the RC. The obtained PCA model could be used to check for new measurement data and fault detection. In order to detect the fault, confidence limits must be determined according to equations: 6.31 and 6.33.

T^2 Confidence limit=12.4106

Q Confidence limit= 0.1976
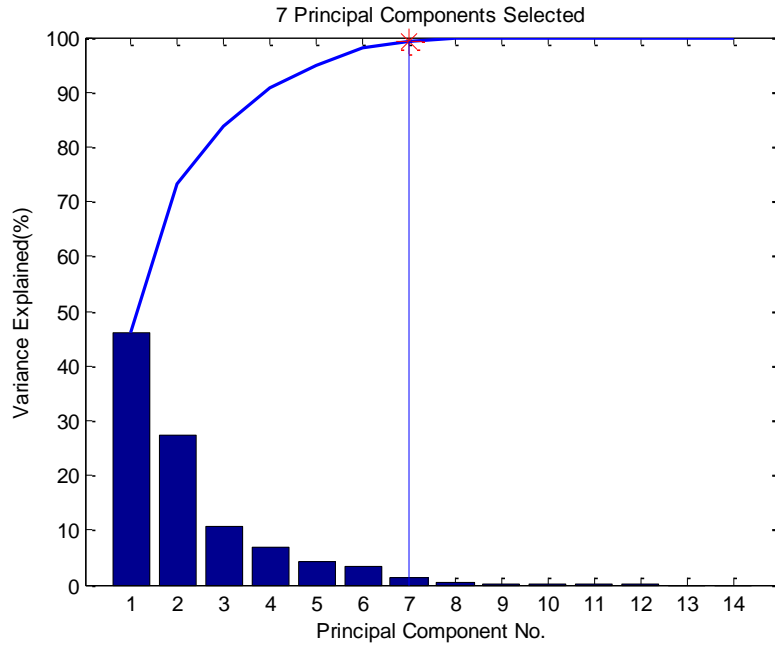
Figure 10.1 PCA selections

10.2.2   PCA Model Based Detection

From Figure 10.2 it can be seen that most of both T^2 and SPE are within the thresholds but there are three occasions, at samples 2, 30, 45, at which the threshold is exceeded. This can be due to the non-stationary behaviour of the vibration signal and the ability of PCA to detect the changes which are acceptable from statistical analysis, but means the confidence level has to be selected appropriately.
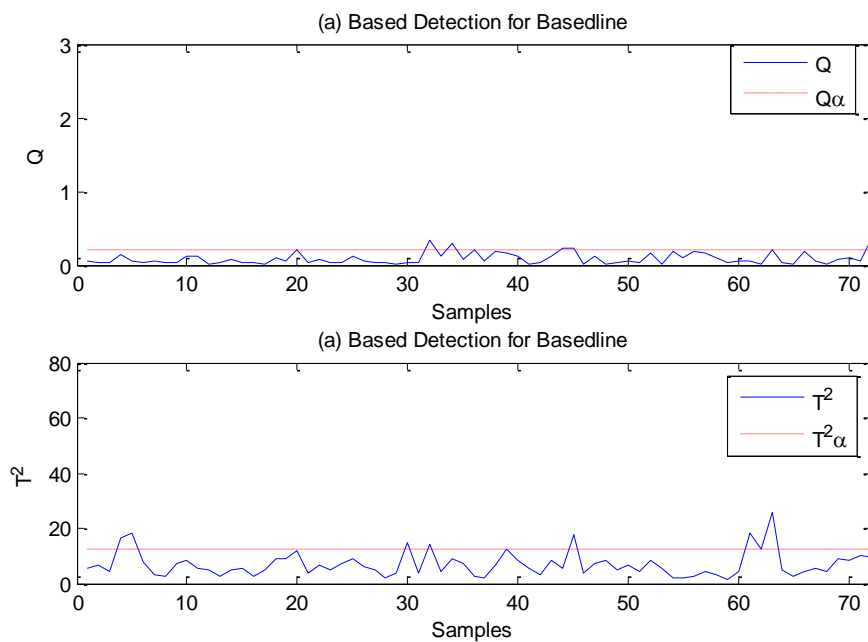
Figure 10.2 PCA model evaluation

On the Q-chart with the leaky discharge valve in Figure 10.3, there are some points at which the control limit is exceeded, and these indicate false alarms. However, the T^2 statistics detected a fault at the same points as shown in Figure 10.3, which shows too many contents reflected by the latent PCs and indicate the presence of a fault.
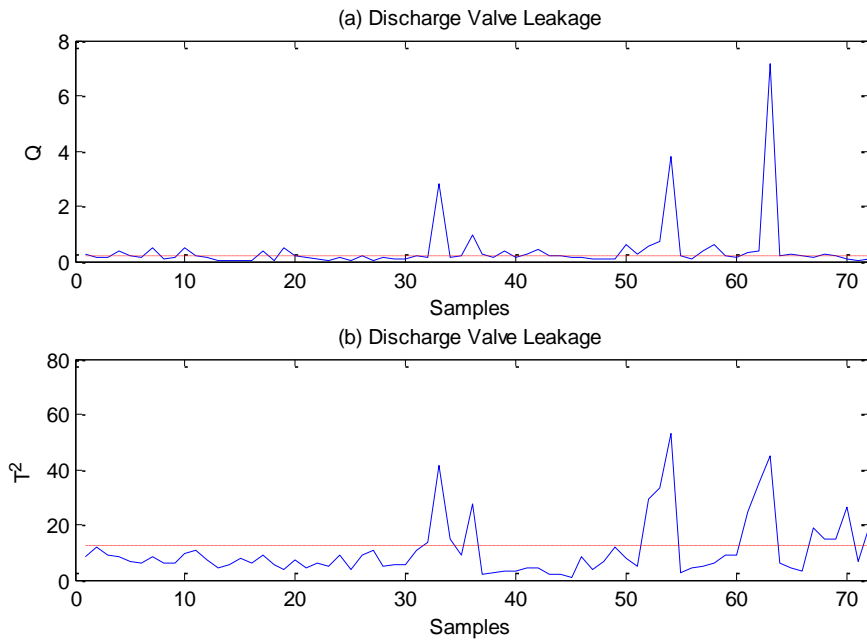
Figure 10.3 Discharge valve leakage detection by T^2 and Q statistics

The performance of the Q method with the leaky suction valve is shown in Figure 10.4. It can be seen that the SPE value exceeds the threshold value many times which indicates the occurrence of major faults while the T^2 method crossed the control limits fewer times as can be seen from the Figure 10.4.
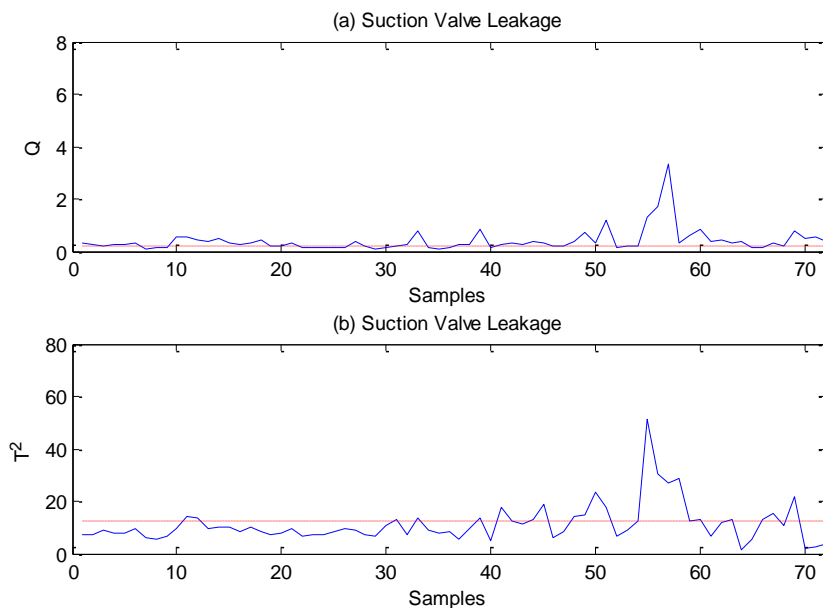
Figure 10.4 Suction valve leakage detection by T^2 and Q statistics

The result in Figure 10.5 shows the Q and T^2 statistics for the intercooler leakage fault, and the values at which the threshold is crossed can be clearly seen in both plots but with larger deviation amplitude in the T^2 method.
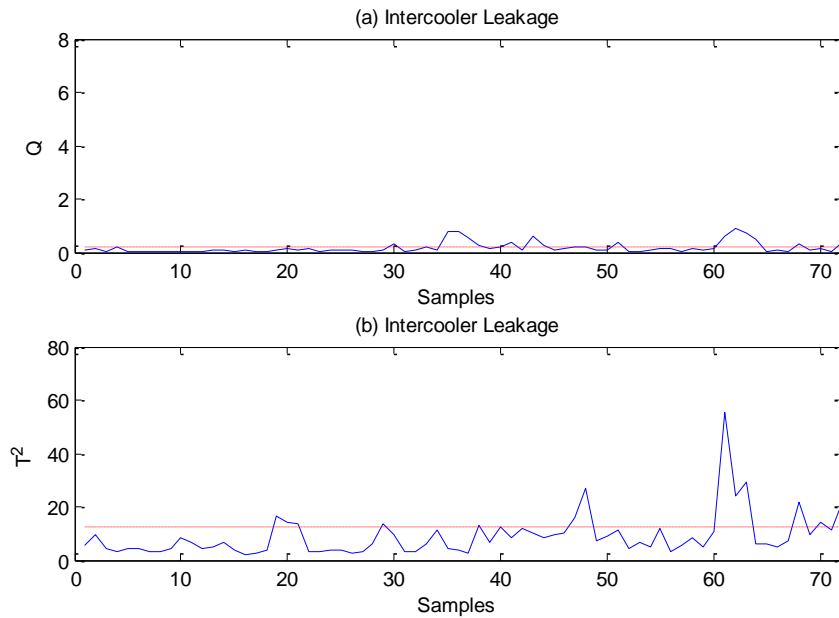
Figure 10.5 Intercooler detection by T^2 and Q statistics

Figure 10.6 depicts the performance of the T^2 and Q methods on the loose belt fault. From the results obtained, it can be seen that the T^2-statistic values cross the threshold many times with high amplitudes, which indicates the occurrence of the major faults whereas the SPE statistic has crossed the threshold in less points with varying amplitude.
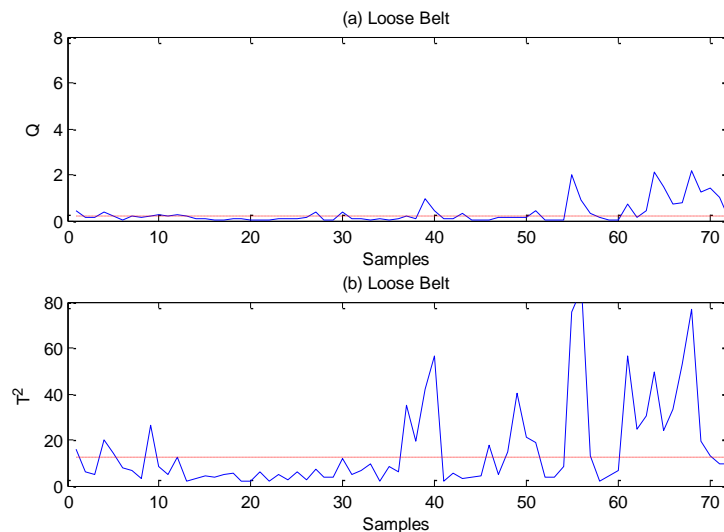
Figure 10.6 Loose belt detection by T^2 and Q statistics

The performance of T^2 and Q statistics models with combined faults was also investigated. Figure 10.7 shows the results for combined discharge valve leakage and suction valve leakage.

It can be seen that with the T^2 method there are a number of occasions which exceed the threshold. Similarly the Q statistics clearly shown the SPE plot crossed the threshold a large number of times, indicating the occurrence of major faults. This confirms the ability of the T^2 method to detect combined faults.
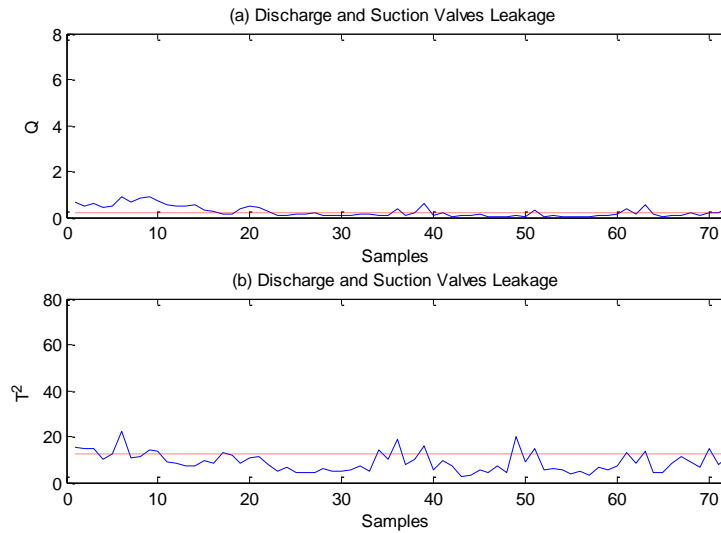
Figure 10.7 Combined discharge valve leakage and suction valve leakage detection by T^2 and Q statistics

For combined suction valve leakage and intercooler leakage, both T^2 and Q statistics detected the faults as shown in Figure 10.8, where it can be clearly seen that many data points exceed the threshold. Both models exhibited similar performance for detection of this fault with particularly high deviation amplitudes in the Q statistics.
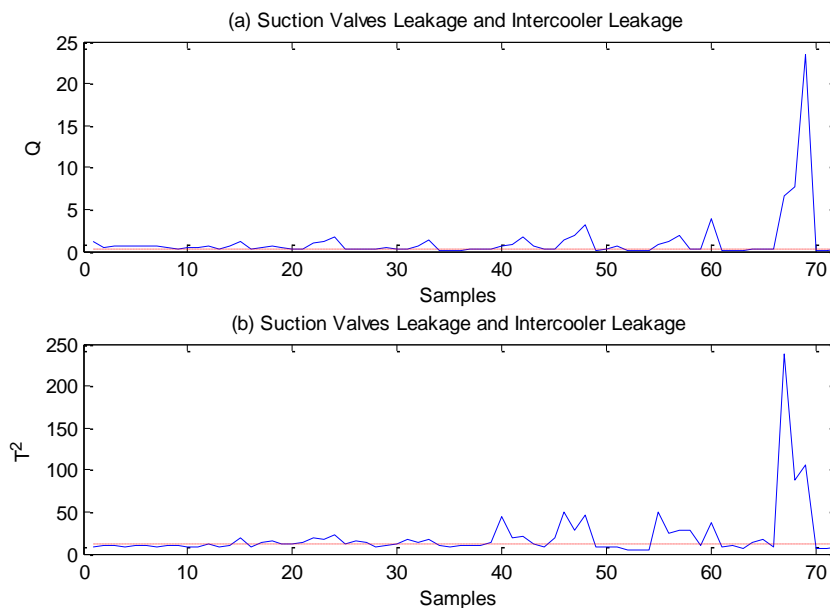
Figure 10.8 Combined suction valve leakage and intercooler leakage detection by T^2 and Q statistics

From the Figure 10.9, the combined discharge valve leakage and intercooler leakage provide many data points that exceed the threshold for the both T^2 and Q statistics and hence indicate the presence of severe faults.
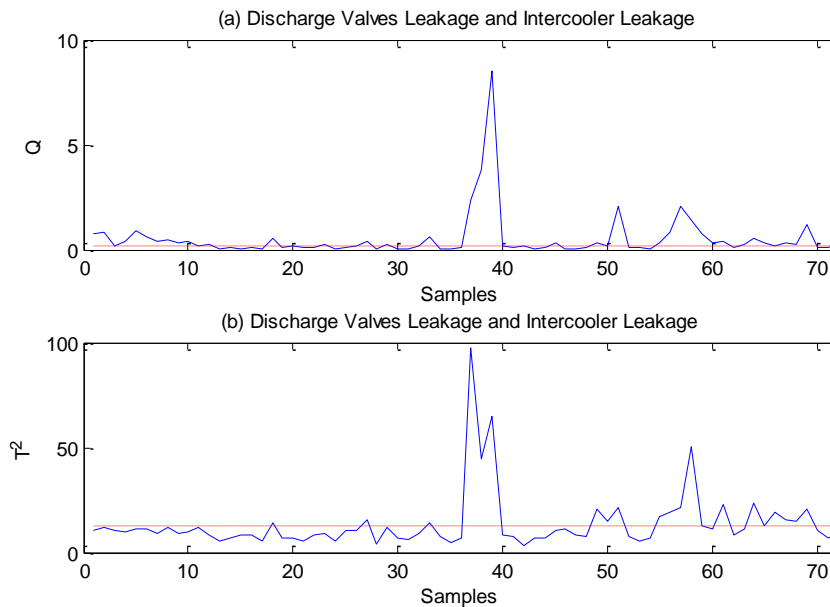
Figure 10.9 Combined discharge valve leakage and intercooler leakage detection by T^2 and Q statistics

10.2.3 PCA Model Based Diagnoses

Once a fault has been detected, it is important to identify an assignable cause. Identification of the source of the fault is facilitated by inspecting the plots showing the contributions of the various measurement variables to the deviations observed in the monitored metric. Such contribution or diagnostic charts can be immediately displayed on line by the system, as soon as the special event is detected. Although they may not provide an unequivocal diagnosis, they should at least clearly indicate the group of variables that are primarily responsible for the detected fault.

10.2.3.1 Contribution Plots Q Statistic

The contribution plots obtained from the data in different cases is shown in Figure 10.10, and the contribution of each variable is different. The major variables contributing in these deviations were mostly variables Amplitude Mean, Range Clearance Factor, Histogram Lower Bound, Histogram Higher Bound, Normal Negative log-liklihood value and Weibull Negative log-likelihood value. Also, the variables contributing significantly to the dissimilarity are variables Max value, RMS, Peak Factor and Skewness. On the other hand, the figure shows that the rest of variables have little variation over different faults and can exclude such as features Entropy, Range and Inter Range.

Number	Feature Name	Case Symbol	Case Name
1	Max	H	Healthy
2	Amplitude Mean	DVL	Discharge Valve leakage
3	RMS	SVL	Suction Valve leakage
4	Peak Factor	BL	Loose Drive Belt
5	Skewness	IL	Intercooler Leakage
6	Kurtosis	DVL+SVL	Discharge Valve leakage with Suction Valve leakage combined fault
7	Entropy	SVL+IL	Suction Valve leakage with Intercooler combined fault
8	Range	DVL+IL	Discharge Valve leakage with Intercooler combined fault
9	Inter Range		
10	Clearance Factor		
11	Histogram Lower Bound		
12	Histogram Higher Bound		
13	Normal Negative log-liklihood value		
14	Weibull Negative log-likilihood value		

Table 10.1 Key to Symbols

Figure 10.10 Q contribution charts for 8 cases based on PCA model

Table 10.2 present the Euclidian distances results obtained for straight Q contributions from the features that were extracted from the vibration time domain. The table shows the results of a total of fourteen features: Max, Abs, Rms, Crest Factor, Skewness, Kurtosis, Spectral Entropy, Range, Interquartile Range, Clearance, Histogram Lower Bound, Histogram Upper

Bound, Weibull Negative Log-likelihood Value and Normal Log-likelihood Value. The first test used the two features one at a time, and the feature that gave the largest minimum Euclidean distance value between the two classes were Clearance Factor and Histogram Lower Bound, with the distance 0.038706 between the suction valve leakage and the discharge valve leakage with the suction valve leakage combined fault. It also recorded that the largest maximum Euclidean distance value with this case is 0.36295 between the loose drive belt fault and suction valve leakage with the intercooler leakage combined fault. The second test combined the three features at a time and here the features Amplitude Mean, Range and Clearance Factor gave the largest minimum Euclidean distance value between the two classes.

By selecting four features from fourteen, the best results (0.075514) were obtained from the combination features of Amplitude Mean, Range, Clearance Factor and Normal Negative log-likelihood value. The minimum of Euclidean distance value between the two classes increased when additional features were added until when using all fourteen features the largest minimum Euclidean distance value between the two classes is 0.090919.

No of input features	Feature number	largest minimum Euclidean distance value between two classes	Classes names	largest maximum Euclidean distance value between Two classes	Classes names
2	10,11	0.038706	SVL,DVL+SVL	0.36295	LB,SVL+IL
3	2,8,10	0.049672	LB,IC	0.336	H,SVL+IL
4	2,8,10,13	0.075514	H,IL	0.34664	H,SVL+IL
5	2,8,10,11,13	0.085061	SVL+IL,DVL+IL	0.37783	LB,SVL+IL
6	2,8,10,11,12,13	0.086911	H,IL	0.37819	LB,SVL+IL
14	1-14	0.090919	H,IL	0.38552	LB,SVL+IL

Table 10.2 Performance of the best Euclidian distances with different feature combinations

As can be seen from the previous table, the variables contributing most significantly to the Q-statistic with different combinations are Clearance Factor and Histogram Lower Bound because they are have high amplitude and these features were repeated for different combinations of features and achieved the largest minimum Euclidean distance value between the classes. This also shows that these features are sensitive to change in each case. See Figure 10.11.
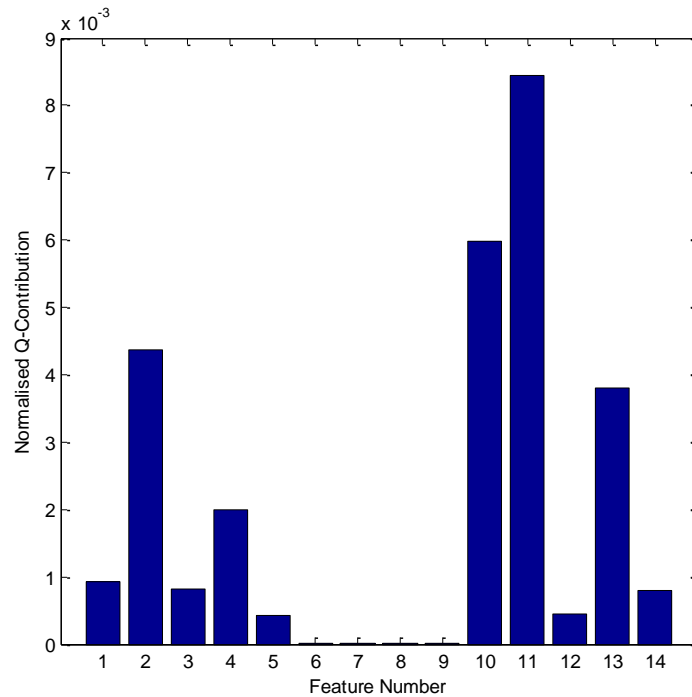
Figure 10.11 Overall Q contribution charts for 8 cases based on PCA model

Figure 10.12 shows the Features Clearance Factor and Histogram Lower Bound have a large difference between cases. The figure also shows that variable Clearance Factor contributes most to the combined suction valve with an intercooler leakage fault and combined discharge valve leakage with an intercooler leakage fault. Variable Histogram Lower Bound also recorded the highest contribution for a loose drive belt fault and inter cooler leakage fault. Thus, the information obtained from the contribution plots is useful for investigating the cause of the fault.

Figure 10.12 Q contribution charts for fault classification based on features Clearance Factor and Histogram Lower Bound

We can therefore represent the faults as combinations of variables. Figure 10.13 and Figure 10.14 present a way to achieve separation between the normal operation and operation with any of the given faults with the optimal distances between classes. It provides the best combination of variables, with which to detect faults most effectively.
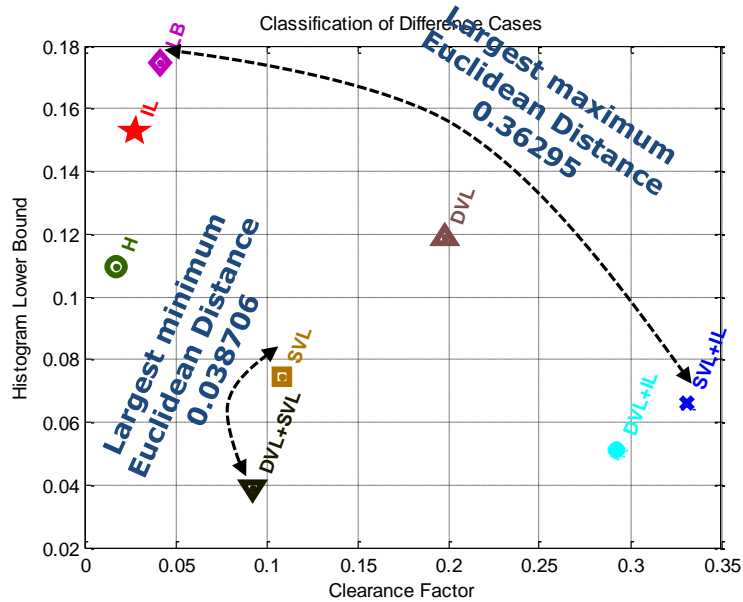
Figure 10.13 Q contribution Fault classifications based on feature clearance factor and histogram lower bound combination
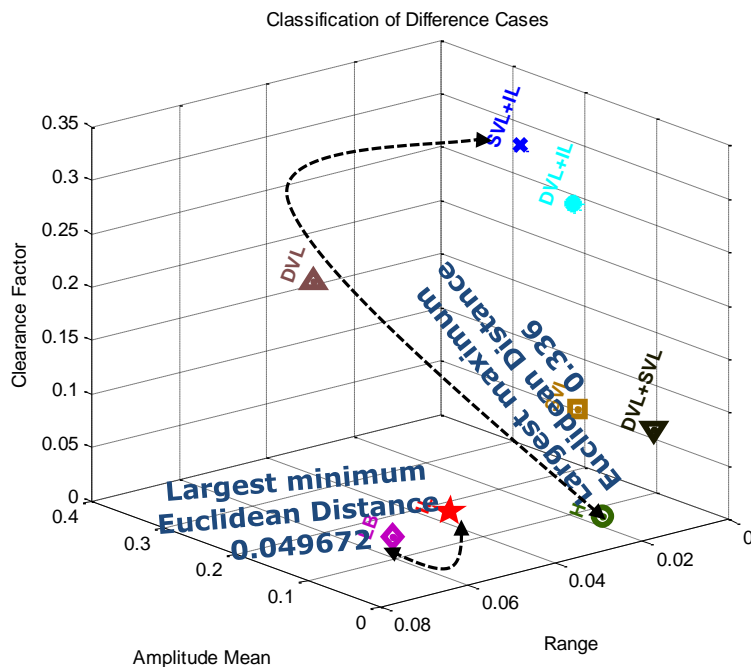
Figure 10.14 Q contribution Fault classifications based on features mean, range and clearance factor combinations

10.2.3.2 *Contribution Plots D Statistic*

The contribution plots for T^2 can be used to help identify a signature of a particular fault as can be shown in Figure 10.15. One can see that certain variables have large contribution, i.e. large deviation from normal process behaviour, and can be associated with a particular fault.

The figure shows the D contributions plots for the fourteen features of eight conditions cases. We can conclude that all the variables are contributing. The variables that had high deviations were mostly variables Amplitude Mean, Peak Factor, Kurtosis, Entropy, Inter Range, Clearance Factor, Normal Negative log-likelihood value and Weibull Negative log-likelihood value. On the other hand, the reset of variables contributed significantly to the dissimilarity.
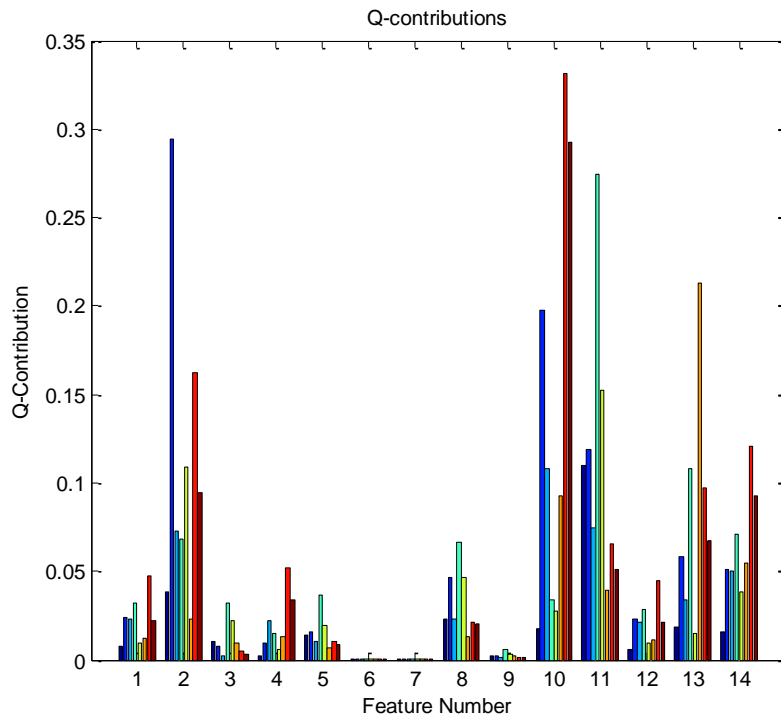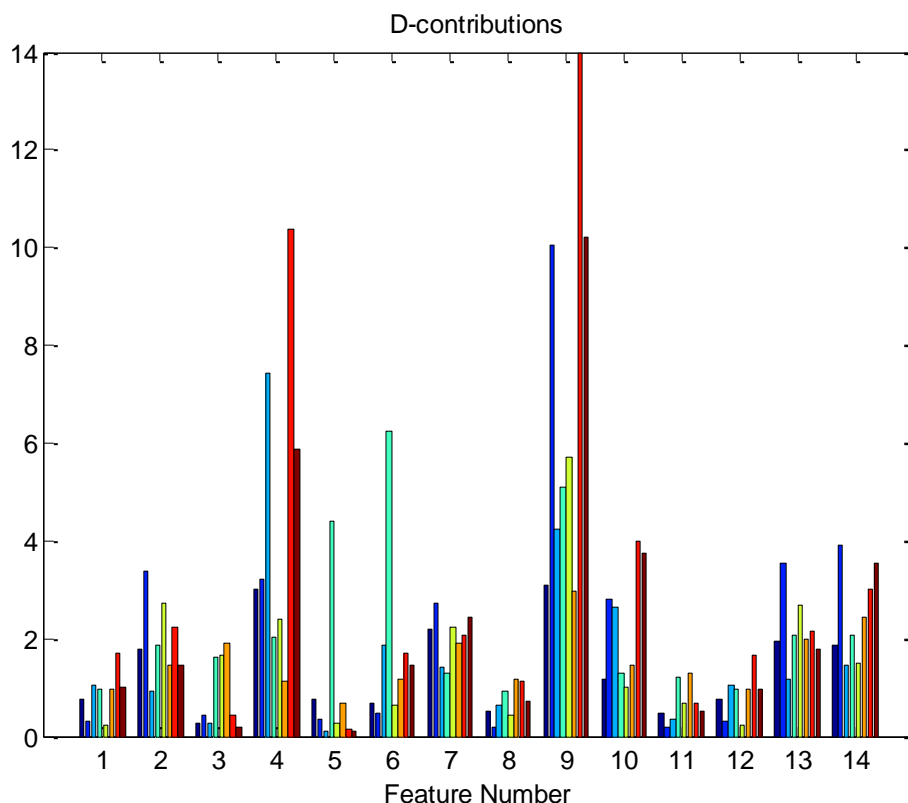

Figure 10.15 D contribution charts for 8 cases based on PCA model

Table 10.3 shows the results for Euclidian distances results obtained for straight D contributions from the features extracted from the vibration time domain. The number of input features from 2 combinations to 14 combinations was set and the Euclidian distances function selected the particular features which it calculated and gave the optimal maximum distance value. When n =2 the optimal maximum distance value calculated that the best two features to use were Amplitude Mean and Peak Factor with the largest minimum distance value is 0. 0.89144 between the loose drive belt fault and discharge valve leakage with suction valve leakage combined fault. Whereas, it recorded the largest maximum Euclidean distance value with this case as 14. 14.8228 between loose drive belt fault suction valve leakage with intercooler leakage fault combined fault.

No of input features	Feature number	largest minimum Euclidean distance value between two classes	Classes names	largest maximum Euclidean distance value between Two classes	Classes names
2	2,4	0.89144	LB, DVL+SVL	14.14.8228	LB, SVL+IL
3	3,4,9	0.79031	LB, DVL+SVL	14.8675	LB, SVL+IL
4	4,8,9,11	0.89755	LB, DVL+SVL	14.8288	LB, SVL+IL
5	1,2,3,4,9	0.96775	LB, DVL+SVL	14.8997	LB, SVL+IL
6	1,2,3,4,9,11	1.0224	LB, DVL+SVL	14.9025	LB, SVL+IL
7	1,2,3,4,5,9, 11	2.2865	LB, DVL+SVL	15.1242	LB, SVL+IL
8	1,2,3,4,5,6, 9,11	2.6909	H, DVL+SVL	15.3336	LB, SVL+IL
9	1,2,3,4,5,6, 7,9,11	2.7044	H, DVL+SVL	15.3367	LB, SVL+IL
10	1,2,3,4,5,6, 7,8,9,11	2.7853	H, DVL+SVL	15.3415	LB, SVL+IL
11	1-11	2.8025	H, DVL+SVL	15.6214	LB, SVL+IL
12	1-12	2.8101	H, DVL,SVL	15.649	LB, SVL+IL
13	1-13	2.8104	H, DVL+SVL	15.6491	LB, SVL+IL
14	1-14	2.8672	H, DVL+SVL	15.7084	LB, SVL+IL

Table 10.3 Performance of the best Euclidian distances with different feature combinations

As can be seen from the previous table that the variables contributing most significantly to the D-contribution with two features combinations are Amplitude Mean and Peak Factor because they have high amplitude and these combined features were repeated for different combinations of features and achieved the largest minimum Euclidean distance value from different two features combinations between the classes also showing that these features are sensitive for change for each cases. See Figure 10.16.
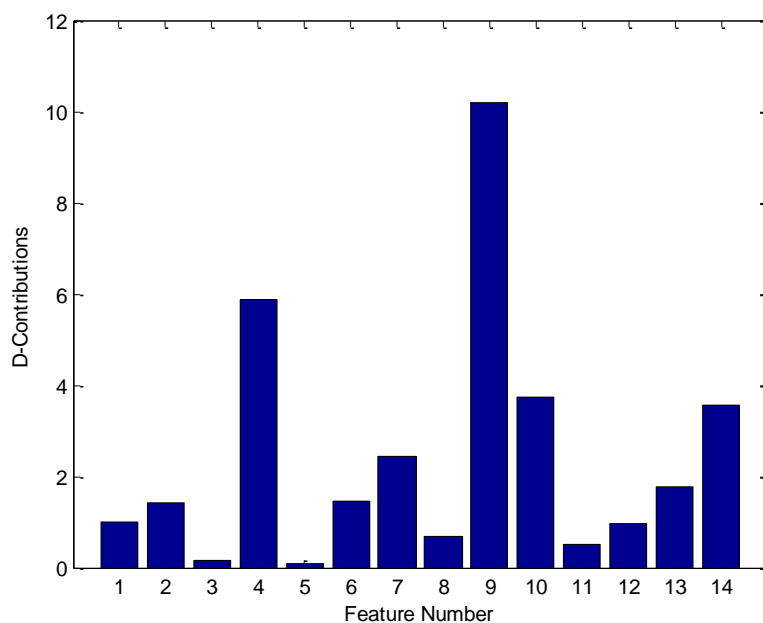
Figure 10.16 Overall D contribution charts for 8 cases based on PCA model

As can be seen from the Figure 10.17 the features Amplitude Mean and Peak Factor show large differences between cases. The figure shows that variable Amplitude Mean contributes most to the discharge valve leakage fault and the loose drive belt fault. Variable Peak Factor also recorded the highest contribution for the combined suction valve leakage with discharge valve leakage fault and suction valve leakage fault. Thus, the information obtained from the contribution plots is useful for investigating the cause of the fault.
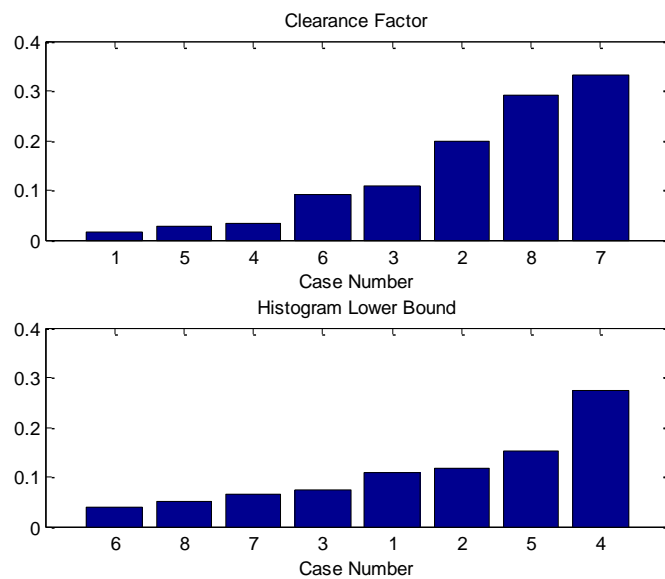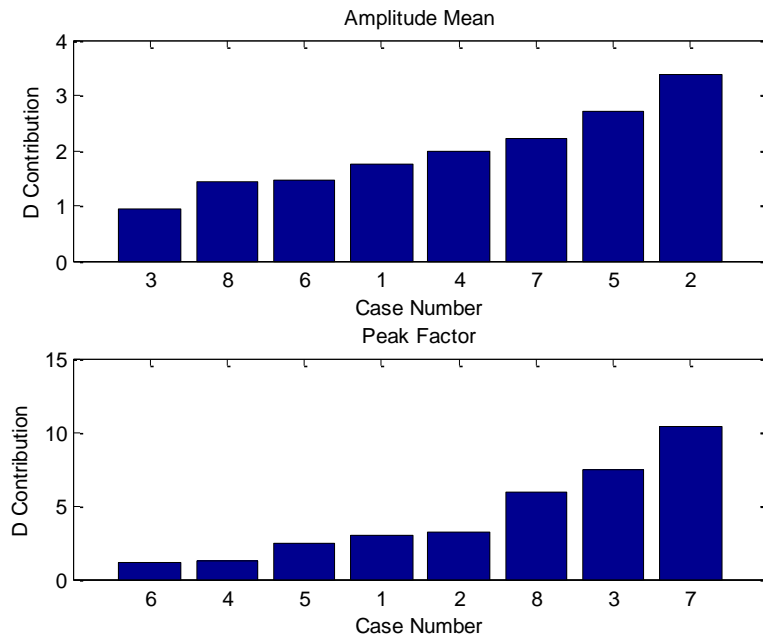

Figure 10.17 *D* **contribution charts for fault classification based on features Amplitude Mean and Peak Factor**

We can therefore represent the faults as combinations of variables. Figure 10.18 and Figure 10.19 present a way to achieve separation between the normal operation and operation with any of the given faults. It provides the best combination of variables with which to detect faults most effectively depends on Euclidian distance.

Figure 10.18 shown that the eight classes can be separated by two combinations of variables that include features Amplitude Mean and Peak Factor which, given by the D-plots with the largest minimum Euclidean distance, the value between the classes is 0.89144. Whereas, the scatter plot is shown in Figure 10.19. It can be seen in the representation of a three-dimensional figure that the contribution plot based on PCA D contribution shows that the three combinations features RMS, Peak Factor and Inter Range, as this combination also gives a full separation with optimal distance between two classes being 0.79031. However, the optimal values of the Euclidean distance between the two classes increase with the increase in the number of features combinations. These combinations give a direction in the multivariate tool-state variable space, onto which the data can be projected, which can be used for detecting a specific class of fault.
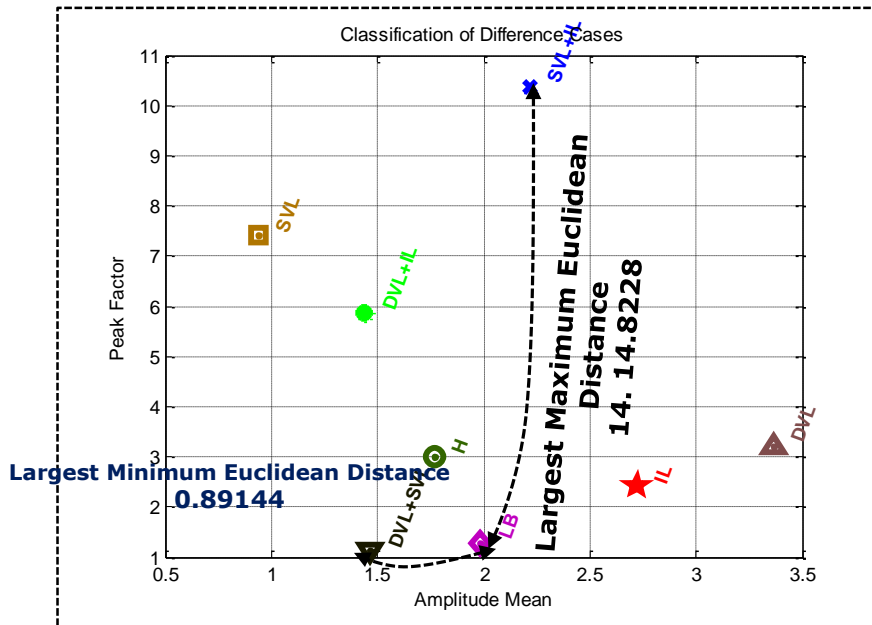
Figure 10.18 D contribution Fault classifications based on feature Mean and Peak factor combination
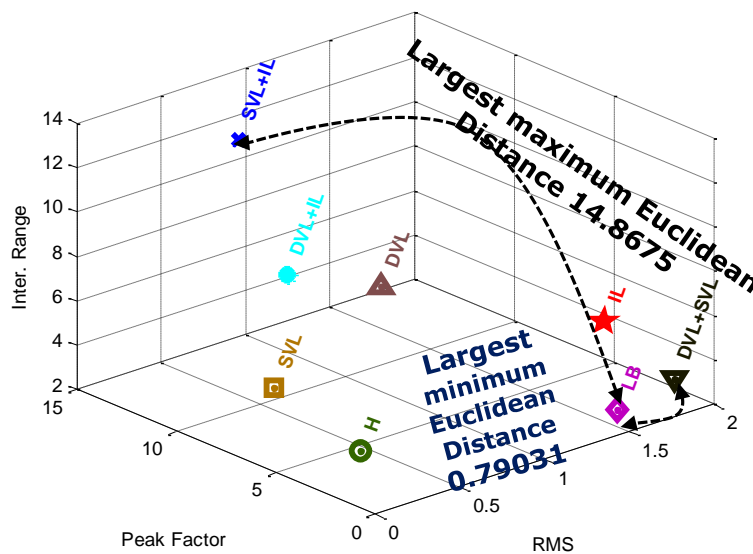
Figure 10.19 D contribution Fault classifications based on features RMS, Peak factor and inter range combinations

10.3   Summary

It has been demonstrated in this study that PCA based approaches allow the detection of single and multiple faults in a RC. The model developed from baseline consists of the seven most important PCs which explain nearly 99% of the variances from 14 original vibration features. The presence of faults can be detected by comparing the feature values from the time domain of the vibration signal with the T^2 and Q statistics. However the Q-statistic had better detection ability for all faults investigated. Following successful detection of faults, the

contributions of the Q-plot and D-plot, were presented in a way which allows it to be used with any latent variable component or regression model to detect a specific progress variable, the variables related to the faults are identified. Intention is to find effective features for fault classification. In the same time the Euclidean distances were used to calculate the optimal distances between the classes.

The Q-contributions show that two particular variables - Clearance Factor and Normal Negative log-likelihood value - gave the largest values of the minimum difference between different cases, whereas the variables Amplitude Mean and Peak Factor were higher contributions for D-plots, thus these were used to detect and differentiate the given faults.

The performance of the PCA improved considerably when using features extracted from the time domain. It did not outperform the PCA using features extracted from envelope spectrum in the fault diagnosis due to noise and the PCA procedure needs to be improved.

The previous method is useful in fault location and understanding the fault impacts through the identified variables for process time. This proposed procedure will be applied with data measured from a real batch process involving multi-stages and phases for RC.

Chapter 11
Conclusion and Plan for Future Work

In this chapter, the author presents his conclusions on the CM of RCs using AI approaches and PCA method with vibration data. Finally, it describes the contributions to knowledge made by this research and makes recommendations for future work.

11.1   Conclusions

Although several techniques have been reported in the literature for RCs fault detection and diagnosis, it is still challenging to implement a reliable CM system for real-world industrial applications due to the complex structure and the often poor operating conditions of these compressors. Failures often occur which can have severe consequences. To find a technique which can accurately monitor and diagnosis the condition of the compressor is becoming increasingly important to avoid system and machine failure, and so improve general reliability and uptime. Although, the vibration signal of a RC in a single service cycle are contains non-linear characteristics (e.g. due to the impacts resulting from the movement of the suction and discharge valves). The theme of this thesis is to develop a novel intelligent system to tackle these related challenges. The strategy is to develop more robust techniques at each processing stage to improve the CM reliability.

This work presented an approach for RC fault detection and diagnosis in time domain, frequency domain, envelope spectrum vibration analysis, and AI field using NNs, SVM, RVM, GAs techniques and PCA method see Figure 11.1.
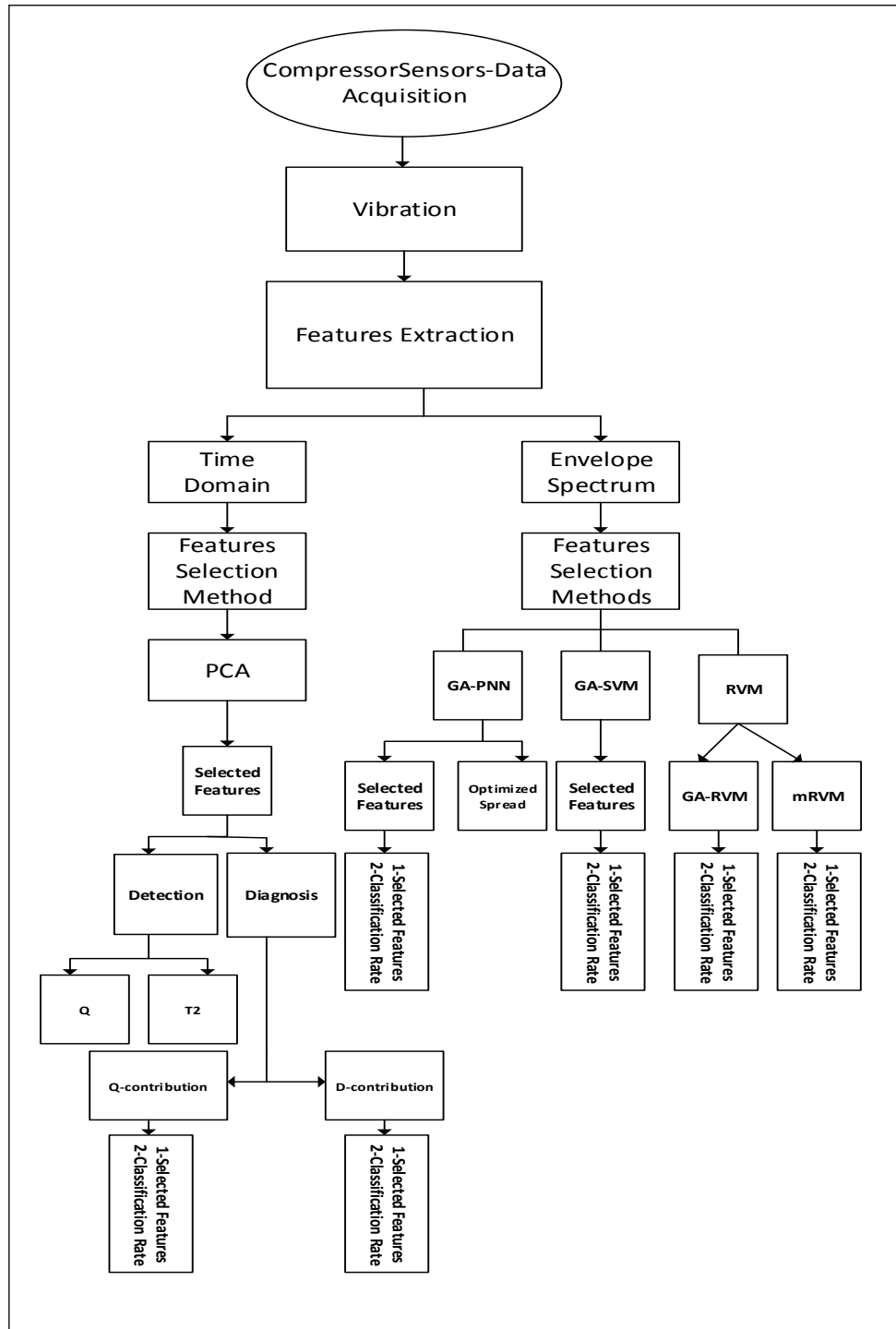
Figure 11.1 Conclusion of thesis work

The results obtained in this project have revealed that it can be concluded that:

Conclusion 1: An available RC was integrated with additional sensors and used as a test facility for experimental measurements required for the development of the key parameters for

compressor condition evaluation. A hardware signal conditioning and data acquisition program was designed and implemented.

Conclusion 2: Seven typical compressor common faults types are investigated. Discharge valve leakage causes a small increase of modulation due to a higher cylinder pressure in the second stage whereas the intercooler leakage causes a small decrease in the modulation due to a lower cylinder pressure in both stages. However, belt looseness results in significant reduction of the modulation because of belt slippage and damping effects. In addition, three more combined faults were created on the test rig: discharge valve leakage with intercooler leakage combined fault, suction valve leakage with intercooler leakage combined fault and discharge valve leakage with suction valve leakage combined fault. Correspondingly, these different faults introduce local increase, local decrease and global changes to the vibration valve motion and hence are representative to different compressor faults.

Conclusion 3: The mathematical model of the two-stage RC developed as part of this project has been tested against the experimental rig for a number of operating conditions. As seen in Section 4.9.2 the 1st stage and 2nd stage healthy valve operation shows that is good qualitative and reasonable quantitative between the model predictions of the crank angle at which the suction and discharge valves open and close, and corresponding measured vibration signals from the head of the 2nd stage cylinder.

Conclusion 4: Examination of vibration signals under different operating conditions has shown that the signals are very complex and there is an enormous amount of unknown information associated with the waveform of the signals. The difference in amplitudes of the signals was due to different operating conditions and limited information could be extracted from the vibration waveform. Results from the vibration waveform showed that faults in the RC are difficult to detect using this approach.

Conclusion 5: Most statistical parameters were applied in the RC with different conditions. The statistical analysis based on RMS has the ability to detect the compressor defects at intercooler leakage and the combination faults that include the intercooler faults for different discharge pressure. Comparison of the peak factor, kurtosis and skewness values of a given vibration signal with values determined using vibrations from a healthy compressor and faulty compressor could not detect the presence defect with different discharge pressure.

Conclusion 6: The results obtained demonstrate the capacity of vibration analysis for the detection of RC faults. It is concluded that vibration analysis can be used as a reliable technique to identify the health status of the RC driven by the valve motions.

Conclusion 7: The Envelope spectrum method can be used more reliably in fault detections than time domain analysis and frequency domains are more sensitive to fault detection. The

results in Figures 5.15 to 5.18 showed that the use of such analyses can enhance the ability to differentiate between healthy and faulty RC patterns in envelope spectra under certain RC operating conditions.

Conclusion 8: Features from the time domain and envelope spectrum were extracted from vibration signals, for separating different healthy conditions of the RC. It was shown that PNNs together with the envelope spectrum features are suited for CM of a RC. One major drawback of the use of single valued statistical and spectral measures is that they produce a large number of features and in general only a subset of those components generated are used for classification.

The PNN using the combination of root mean square, peak factor, lower bound, upper bound, entropy, variance, skewness, kurtosis, maximum value and range from the time domain gives a classification rate of 43.0%. With envelope spectrum analysis a successful classification rate of 100% has been achieved, which is superior but it requires 30 harmonic peaks in the spectrum, which indicates that the envelope spectrum feature is more useful for classification.

The GAs approach based on the concept of dominance was developed for optimal feature selection. It is shown that GA is well suited to feature selection as it can produce a diverse set of solutions with differing performance versus complexity trade-off characteristics in a single population. The spread parameter of the PNN was also optimised via GA. With GA-PNN approach, the success rate of using the time domain feature is very low, only 48% correct classification. However, in the envelope spectrum the success rate is uniformly high. Especially, with 30 input harmonics it reaches 100% successful classification.

It thus can be stated that the application of GA with features from the time domain optimised the number of input features with high correct classification than PNN without GA. On the other hand, the PNN approach with envelope spectrum was tested with a large number of inputs and very good results were achieved. It was demonstrated that a GA is capable of selecting a subset of less than half the inputs from a set of 120 features that allow the ANN to perform with 99% accuracy. Moreover, it is concluded that features from the envelope spectrum are the optimal features for classifying the healthy and faulty conditions of RCs.

Conclusion 9: The SVM was developed and shown capable of detecting the designated faults and distinguishing different types of faults. In this study, the SVM classifier when used with features extracted from the envelope spectrum for binary classes gave a success rate of 100%. Briefly, the use of SVMs for binary classes as a powerful tool for fault detection is presented. Moreover, the performance of SVM has been found to be substantially better with the OAA strategy. With multiple classes the performance of the SVM classifier was satisfied. The best success rate (88.00%) was achieved using harmonics between 18 and 25 of the

peaks in the envelope spectrum. The results show the potential application of GAs for selection of features in machine condition detection.

Conclusion 10: Two procedures of new RC fault detection and diagnosis methods based on RVM of vibration data are proposed in this research. The first procedure is presented for detection of RC condition using RVMs-OAO classifiers from envelope spectrum vibration signals and the selection of input features and the appropriate classifier parameters have been optimised using RVMs-OAO based on GAs approach.

The second procedure is to implement the direct multi-class mRVMs algorithms of RC fault detection and diagnosis from envelope spectrum vibration data. Both procedures of RVMs are viewed as classifier and features selection respectively.

The RVM develops an OAO scheme with GA feature optimisation for applying this binary classifier to the compressor's data, and examines the harmonics selected for classification to find the insight of the classifier in association with the physical supports. In addition, the performance of the multiclass multi-kernel mRVM is also explored for obtaining more efficient fault classification. The accuracy of both techniques is discussed individually to determine the optimum fault classifier. The results show that the models behave well, and the classification accuracy rate is up to 97% for both algorithms.

Conclusion 11: The PCA model approaches demonstrated allow the detection of single and hybrid faults in a two stage RC. The model developed from baseline consists of the seven most important PCs which explain nearly 99% of the variances from 14 original vibration features extracted from the statistical time domain.

The presence of faults can be detected by comparing the T^2 and Q values from fault features from the time domain of vibration signals with the corresponding thresholds developed based on baseline data. However, the Q statistic produces a better detection for the most faults cases investigated, showing it more suitable for fault detection. Moreover, the contributions of the Q-plot and D- plot, were presented in a way which allows it to be used with any latent variable component or regression model to detect a specific progress variable, the variables related to the faults are identified. The intention is to find effective features for fault classification. In the same time the Euclidean distances were used to calculate the optimal distances between the classes.

The Q-contributions show that two particular variables Clearance Factor and Normal Negative log-liklihood value gave the largest values of the minimum difference between different cases, whereas the variables Amplitude Mean and Peak Factor were higher contributions for D-plots, thus these were used to detect and differentiate the given faults.

Conclusion 12: The PCA clearly performed better than the AI approaches when diagnosing the faults using features extracted from the time-domain.

11.2 Summary of Contributions

The work conducted by the author and described by this thesis included several aspects that were novel and not previously implemented by other researchers or practitioners. A summary of these contributions is given below:

Contribution 1: The author believes that the application of vibration for detection and diagnosis of RC with separate and combination faults (e.g. discharge valve leakage, suction of valve leakage, loose drive belt, intercooler leakage and different combination faults such as discharge valve leakage combined with suction valve leakage, discharge valve leakage combined with intercooler leakage and suction valve leakage combined with intercooler leakage) is novel. No work has been found in the literature that describes the use of multiple faults for CM of multi-stage RCs using vibration data, either experimentally or using a mathematical model.

Contribution 2: The application of a mathematical model for a two-stage RC has not previously been used for RC faults CM with different faults.

Contribution 3: The author believes that the use of envelope spectrum techniques for the analysis of the vibration for a RC monitoring for CM is novel as no reports in the literature have been found of using envelope for fault detection and fault diagnoses for RCs.

Contribution 4: The author introduced an effective features selection using PNN technique and GAs and also optimised spread parameter of the PNN features, which has not formerly been employed for RC fault monitoring using vibration.

Contribution 5: The author of this research believes that the research work in this thesis is the first work to find the optimal features and high accuracy rate for classification with eight different conditions classes for the RC from vibration data using hybrid techniques, which optimise the SVM with GAs for training process.

Contribution 6: There are no previous reports in the literature of the use of the RVM with binary and multi classes via GAs to find the optimal features and the high accuracy rate for classification from the vibration envelope features extraction of a RC. In addition, the mRVM direct was applied for first time to find the optimal features and calculate the classification rate for the vibration data from RC.

Contribution 7: The application of the PCA model has investigated using the time domain vibration features for detection and diagnosis from a multi stage RC with base line data sets and seven different faults can be detected by comparing the values from the features of fault

vibration signals with corresponding thresholds developed from baseline data. However, the Q-statistic T^2-statistic procedure produces a better detection as it can separate the seven faults completely. Furthermore, a study of Q-contributions and D-contribution have found and determined the original features which allow full classification of different simulated faults. This study has not previously been used for RC fault CM.

11.3 Recommendations for Future Work

This section lists recommendations for further work on the research:

Recommendation 1: It is recommended that further academic research should be conducted into the effective utilisation of envelope features to detect vibration faults in RCs.

Recommendation 2: Further investigations using the vibration in a RC could be extended as a fault detection capability for other faults including the piston system, transmission system and the driving electrical motor. These faults could also include both first and second stages of the RC.

Recommendation 3: Data sets could be extracted from more than one sensor which would be mounted in different positions: these could be two accelerometers mounted at 90^0 to each other, or an accelerometer with e.g. a microphone for airborne sound or an encoder for instantaneous angular speed.

Recommendation 4: To develop equations for the determination of suction and discharge conditions when the compressor is operated with different combined faults.

Recommendation 5: It is recommended that further academic research to extract features from the time-frequency domain can be used as inputs into computational intelligence, as studied in this thesis.

Recommendation 6: The proposed RC fault detection techniques and decision-making schemes will be applied to other mechanical systems such as bearing, gearboxes and engines.

Recommendation 7: Implement the developed monitoring tools for real-world industrial monitoring applications, so as to improve production quality and to reduce costs.

Recommendation 8: Development of fault diagnoses depending on AI technologies. There are a range of different types of AI techniques such as fuzzy logic; fuzzy neural networks and that could have a role in developing better-quality vibration diagnostic systems with highly developed sensitivity, dependability and automation.

Recommendation 9: It is recommended that further academic research to select the kernel width with GAs to ensure the ability to obtain the optimal value.

References

Ahmed, M., Gu, F., & Ball, A. (2011). Feature Selection and Fault Classification of Reciprocating Compressors using a Genetic Algorithm and a Probabilistic Neural Network. Paper presented at the 9th International Conferenceon Damage Assessment of Structures (DAMAS2011), Oxford, UK.

Ahmed, M., Smith, A., Gu, F., & Ball, A. D. (2014). Fault Diagnosis of Reciprocating Compressors Using Revelance Vector Machines with A Genetic Algorithm Based on Vibration Data. Proceedings of the 20th International Conference on Automation & Computing, Cranfield University, Bedfordshire, UK, IEEE. ICAC.

Al-Arbi, S. (2012). Condition Monitoring of Gear Systems using Vibration Analysis. University of Huddersfield.

Albert, M. B., Avery, D., Narin, F., & McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. Research policy, 20(3), 251-259.

Amer, F. Z., El-Garhy, A. M., Awadalla, M. H., Rashad, S. M., & Abdien, A. K. (2011). A real-valued genetic algorithm to optimize the parameters of support vector machine for classification of multiple faults in NPP. NUKLEONIKA, 56(4), 323-332.

Asakura, T., Kobayasi, T., Hayashi, S., & Xu, B. (1997). Fault Diagnosis System of Machine Using Neural Networks and Its Application: University of Fukui.

Awadalla, M., Abdien, A., Rashad, S., Ahmed, A., & Al Abri, D. (2014). Classification of Faults in Nuclear Power Plant. WSEAS Transactions on Systems, 13.

Awadallah, M., & Morcos, M. M. (2003). Application of AI tools in fault diagnosis of electrical machines and drives-an overview. Energy Conversion, IEEE Transactions on, 18(2), 245-251.

Ay, H., Jang, J. Y., & Yeh, J. N. (2002). Int J Heat Transfer. 45(4069).

Ball, A. D. (2000). Reciprocating Engines and Compressors. Maintenance Engineering M12, section 11, Course Notes, University of Manchester, Manchester England.

Benaicha, A., Guerfel, M., Bouguila, N., & Benothman, K. (2010). New PCA based methodology for sensor fault detection and localization. 8th International Conference of Modeling and Simulation. Ilammamet, Tunisia.

BenSasi, Y. (2005). The Exploitation of Instantaneous Angular Speed for Machinery Condition Monitoring. PhD thesis, school of mechanical, aerospace and civil engineering, the University of Manchester, UK.

Bishop, C. M. (2006). Pattern recognition and machine learning (Vol. 1): springer New York.

Blickle, T., & Thiele, L. (1995). A Comparison of Selection Schemes used in Genetic Algorithms. TIK-Report, Zurich.

Bloch, H. P., & Hoefner, J. J. (1996). Reciprocating Compressors:: Operation and Maintenance: Gulf Professional Publishing.

Bolat, B., & Yildirim, T. (2003). Performance Increasing Methods for Probabilistic Neural Networks. Pakistan Journal of Information and Technology, 2(3), 250-255.

Boser, B., Guyon, I., & Vapnik, V. N. (1992). A training algorithm for optimal margin cla*ssifiers.* Paper presented at the Proceedings of the Fifth Annual Workshop on Computational Learning Theory, New York.

Chapelle, O., Haffner, P., & Vapnik, V. N. (1999). Support vector machines for histogram-based image classification. Neural Networks, IEEE Transactions on, 10(5), 1055-1064.

Chaykosky, S. (2002). Resolution of a Compressor Valve Failure: A Case Study. Dresser-Rand Literature.

Chen, Z., & Lian, X. (2010). Fault diagnosis for valves of compressors based on support vector machine. IEEE Chinese Control and Decision Conference, 1235 – 1238.

Chen, Z., & Lian, X. (2010). Fault diagnosis for valves of compressors based on support vector machine. IEEE Chinese Control and Decision Conference 1235 – 1238.

Chen, Z., & Lian, X. (2010). Fault Diagnosis for Valves of Compressors Based on Support Vector Machine. Paper presented at the 2010 Chinese Control and Decision Conference, Xuzhou.

Cho, J., & Moon, S. (2005). A numerical analysis of the interaction between the piston oil film and the component deformation in a reciprocating compressor. Tribology International, 38(5), 459-468.

Chou, P.-H., Wu, M.-J., & Chen, K.-K. (2010). Integrating support vector machine and genetic algorithm to implement dynamic wafer quality prediction system. Expert Systems with Applications, 37(6), 4413-4424.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297.

Cristianini, N., & Shawe-Taylor, N. J. (2000). An Introduction to Support Vector Machines. Cambridge: Cambridge University Press.

Cui, H., Zhang, L., Kang, R., & Lan, X. (2009). Research on fault diagnosis for reciprocating compressor valve using information entropy and SVM method. Journal of Loss Prevention in the Process Industries, 22, 864-867.

Damoulas, T., Ying, Y., Girolami, M. A., & Campbell, C. (2008). Inferring sparse kernel combinations and relevance vectors: an application to subcellular localization of proteins. Paper presented at the Machine Learning and Applications, 2008. ICMLA'08. Seventh International Conference on.

Daniel , J. (1994). Vibration-Based Diagnostics of Reciprocating Machinery. PhD Thesis, Massachusetts Institute of Technology.

Dhar, S., Tamma, B., Bhakta, A., & Krishna, M. (2014). An Approach Towards Reed Valve Geometry Design.

Dresser-Rand. (2005). Magnum Valves for Reciprocating Compressors. Olean, UAS.

Elhaj, M. (2005). Condition monitoring of reciprocating compressor valves. the University of Manchester, UK, Manchester.

Elhaj, M., Gu, F., & Ball, A. (2004). Numerical Simulation and Experimental Study of a Two Stage Reciprocating Compressor for Condition Monitoring. ASME Eng. App. of Artificial Intelligence, , 2(2), 115-131.

Elhaj, M., Gu, F., Ball, A., Albarbar, A., Al-Qattan, M., & Naid, A. (2008). Numerical simulation and experimental study of a two-stage reciprocating compressor for condition monitoring. Mechanical Systems and Signal Processing, 22(2), 374-389.

Erdelyi, E., & Erie, P. A. (1956). Vibration Modes of States of Induction Motors. Trans. of ASME, A-28(39-45).

Feng, G., Mustafa, A., Gu, J., Zhen, D., Gu, F., & Ball, A. (2013). The real-time implementation of envelope analysis for bearing fault diagnosis based on wireless sensor network. Paper presented at the Automation and Computing (ICAC), 2013 19th International Conference on.

Filippetti, F., Franceschini, G., Tassoni, C., & Vas, P. (2000). Recent developments of induction motor drives fault diagnosis using AI techniques. Industrial Electronics, IEEE Transactions on, 47(5), 994-1004.

Fleming, J. (1989). A Theoretical and Experimental Investigation of the Flow of Gas Through Reciprocating Compressor Valves. Roc. European Conference, IMechE, C390/011, 117-119.

Foreman, S. (2002). Compressor Valves and Unloaders for Reciprocating Compressors–An OEM's Perspective. Dresser-Rand Literature.

Fuqing, Y. (2011). Failure diagnostic using Support Vector Machine: Lulea University of Technology.

Ghate, V., Dudel, S. (2009). Induction machine fault detection using support vector machine based classifier. WSEAS Transactions on Systems, 8(5), 591-603.

Girondin, V., Loudahi, M., Pekpe, K. M., & Cassar, J. P. (2012). Vibration-based fault detection of accelerometers in helicopters. Paper presented at the Proceedings of Fault Detection, Supervision and Safety of Technical Processes conference.

Goh, T. C. (2002). Probabilistic neural network for evaluating seismic liquefaction potential. Canadian Geotechnical Journal, 39, 219-232.

Gorunescu, F., Gorunescu, M., El-Darzi, E., Gorunescu, S., & K., R. (2005). A cancer diagnosis system based on rough sets and probabilistic neural networks. Paper presented at the Proceedings of the first European conference on health care modelling and computation.

Guarnaccia, C., & Neri, F. (2013). An introduction to the special issue on recent methods on physical polluting agents and environment modeling and simulation. WSEAS Transactions on Systems, 12(2), 53-54.

Gunn, S. (1998). Support vector machines for classification and regression. Technical Report Faculty of Engineering, Science and Mathematics, School of Electronics and Computer Science, Southampton University.

Guyon, I., Gunn, S., Nikravesh, M., & Zadeh, L. A. (2008). Feature Extraction: Foundations and Applications

He, B., Yang, X., Chen, T., & Zhang, J. (2012). Reconstruction-based multivariate contribution analysis for fault isolation: A branch and bound approach. Journal of Process Control, 22(7), 1228-1236. doi: http://dx.doi.org/10.1016/j.jprocont.2012.05.010

He, C., Li, Y., Huang, Y., Liu, C., & Fei, S. (2009). Relevance vector machine based gear fault detection. Paper presented at the Chinese Conference on Pattern Recognition, Nanjing.

He, X., Hua, E., Lin, Y., & Liu, X. (2011). Computer, Informatics, Cybernetics and Applications. New York: Springer.

Heinz, B., & John, J. (1996). Reciprocating Compressors, Operation & Maintenance: Butterworth-Heinemann, Gulf Company, Houston.

Hu, Q., He, Z., Zhang, Z., & Zi, Y. (2007). Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble. Mechanical Systems and Signal Processing, 21(2), 688-705.

Hu, Y., & Luob, P. (2013). Performance Data Prognostics Based on Relevance Vector Machine and Particle Filter. CHEMICAL ENGINEERING, 33, 349-354.

Hung-Yi, L., Chao, S. M., & Tsai, G. L. (2005). INT J Heat Mass Transfer. 48(5386).

Iovine, J. (2012). Understanding Neural Networks The Experimenter's Guide: Images.

ISO/16063-21. (2003). Methods for the calibration of vibration and shock transducers Vibration calibration by comparison to a reference transducer. Geneva: ISO.

Iturriaga-Notario, L. (1978). U.S. Patent No. 4,078,439. Washington, DC: U.S. Patent and Trademark Office.

Jack, L. B., & Nandi, A. K. (2000). Genetic Algorithms for Feature Selection in Machine Condition Monitoring With vibration Signals. IEEE Proceedings on Image Signal Processing, 147(3), 205-212.

Jack, L. B., Nandi., A. K., & McCormick., A. C. (2000). Genetic algorithms for feature extraction in machine condition monitoring with vibration signals. IEE Proceedings Vision and Image Signal Processing, 147, 205-212.

Jain, V., Pillai, G., & Gupta, I. (2011). Fault diagnosis in analog circuits using multiclass relevance vector machine. Paper presented at the Emerging Trends in Electrical and Computer Technology (ICETECT), 2011 International Conference on.

Jain, V., Pillai, G. N., & Gupta, I. (2011). Fault diagnosis in analog circuits using multiclass relevance vector machines. Paper presented at the Proceedings of ICETECT.

Jie, Z., Hong, J., & Zhang, J. (2010). Progressive PCA modeling for enhanced fault diagnosis in a batch process. nternational Conference on Control, Automation and Systems (ICCAS), Gyeonggi do, South Korea, 713-718.

Kano, M., Hasebe, S., & Hashimoto, I. (2000). Contribution Plots for Fault Identication Based on the Dissimilarity of Process Data. AIChE.

Kano, M., Hasebe, S., Hashimoto, I., & Ohno, H. (2001). Fault detection and identification based on dissimilarity of process data. Preprints of European Control Conference (ECC). Porto, Portugal, 1888-1893.

Kjaer, B. (1997). Transducer & Conditioning, Company Catalogue. Denmark.

Kourti, T., & MacGregor, J. F. (1996). Multivariate SPC Methods for Process and Product Monitoring. J.Qual. Technol, 28, 409.

Kowalski, C. T., & Orlowska-Kowalska, T. (2003). Neural networks application for induction motor faults diagnosis. Math. Comput. Simul., 63(3-5), 435-448. doi: http://dx.doi.org/10.1016/S0378-4754(03)00087-9

Kryter, C., & Haynes, H. Condition Monitoring of Machinery Using Motor Current Signature Analysis. Conference: 7. power plant dynamics, control and testing symposium,Knoxville, TN, USA.

Kwon, Tae-Suk, Lee, Dong-Hoon, & Sul, S.-K. (2005). Reduction of Engine Torque Ripple at Starting with Belt Driven Integrated Starter Generator. School of Electrical Engineering & Computer Science, 024, ENG420, Seoul National University.

Lee, J.-M., Yoo, C. K., & Lee, I.-B. (2004). Enhanced process monitoring of fed-batch penicillin cultivation using time-varying and multivariate statistical analysis. Journal of Biotechnology, 110(2), 119-136. doi: http://dx.doi.org/10.1016/j.jbiotec.2004.01.016

Lei, Y., He, Z., Zi, Y., & Hu, Q. (2007). Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. Mechanical Systems and Signal Processing, 21(5), 2280-2294.

Li, N., Liu, C., He, C., Li, Y., & Zha, X. F. (2011). Gear fault detection based on adaptive wavelet packet feature extraction and relevance vector machine. Paper presented at the Proceedings of the Institution of Mechanical Engineers, Part C, Journal of Mechnical Engineering Science.

Li, W. (2000). A Study of Diesel Engine Acoustic Characteristics. PhD Thesis, Manchester School of Engineering, Department of Mechanical Engineering, Manchester, UK.

Li, Y., Gu, F., Harris, G., Ball, A., Bennett, N., & Travis, K. (2005). The measurement of instantaneous angular speed. Mechanical Systems and Signal Processing, 19(4), 786-805.

Li, Y., & Liu, T. (2010). Study on classification model based on relevance vector machine with genetic algorithm. Paper presented at the 2nd IEEE International Conference on Information and Financial Engineering, Chongqing.

Liang, B. (2000). Condition Monitoring and Fault Diagnosis of Induction Motors. PhD Thesis, Manchester School of Engineering, Department of Mechanical Engineering, Manchester, UK.

Lima, C. A., Coelho, A. L., & Chagas, S. (2009). Automatic EEG signal classification for epilepsy diagnosis with Relevance Vector Machines. Expert Systems with Applications, 36(6), 10054-10059.

Liu, J. (2008). An intelligent system for bearing condition monitoring. Waterloo: University of Waterloo.

Liu, J., Wang, W., Golnaraghi, F., & Liu, K. (2008). Wavelet spectrum analysis for bearing fault diagnostics. Measurement Science and Technology, 19(1), 015105.

Ltd, C. (1991). The CED 1401 and 1401 Plus, Intelligent Interface Programmer's Handbook. Cambridge Electronic Design(Cambridge, U.K).

MacGregor, J. F., & Kourti, T. (1995). Statistical Process Control of Multivariate Processes. Control Engineering Practice, 3(3), 403-414.

MacGregor, J. F., & Kourti, T. (1995). Statistical process control of multivariate processes. Control Engineering Practice, 3, 403-414.

Mahamad, A. K. (2010). Diagnosis, classification and prognosis of rotating machine using artificial intelligence. Kumamoto University.

Mahamad, A. K., & Hiyama, T. (2008). Development of Artificial Neural Network Based Fault Diagnosis of Induction Motor Bearing. 2nd IEEE International Conference on Power and Energy (PECon 08).

Mahmud, A., Ann, S., Feng, G., & Andrew, D. B. (2014). Fault Diagnosis of Reciprocating Compressors Using Revelance Vector Machines with A Genetic Algorithm Based on Vibration Data. Proceedings of the 20th International Conference on Automation & Computing, Cranfield University, Bedfordshire, UK, 12-13 September 2014 IEEE. ICAC.

Martino, S. D., Ferrucci, F., Gravino, C., & Sarro, F. (2011). A Genetic Algorithm to Configure Support Vector Machines for Predicting Fault-Prone Components. Product-Focused Software Process Improvement Lecture Notes in Computer Science, 6759, 247-261.

Masters, T. (1995). Advanced Algorithms for Neural Networks A C++ Sourcebook. New York: Academic Press.

Matsuura, T. (2004). An application of neural network for selecting feature parameters in machinery diagnosis. Journal of materials processing technology, 157, 203-207.

Mayes, I. W., Steer, A. G., & Thomas, G. B. (1981). The Application of Vibration Monitoring for Fault Diagnosis in Large Turbo-generators. 6th Thermal Generation Specialists Meeting, Madrid, 567-575.

McCarthy, D. J. (1997). Prognostics for a reciprocating air compressor. The Journal of the Acoustical Society of America, 101, 3029.

Mdlazi, L., Marwala, T., Stander, C. J., Scheffer, C., & Heyns, P. S. (2007). Principal component analysis and automatic relevance determination in damage identification. Pretoria: Department of Mechanical and Aeronautical Engineering University of Pretoria.

Miller, C. M. (2012). Evolutionary Artificial Neural Network Weight Tuning to Optimize Decision Making for an Abstract Game: BiblioScholar.

Montavon, G., Orr, G., & Muller, K.-R. (2012). Neural Networks: Tricks of the Trade (Lecture Notes in Computer Science / Theoretical Computer Science and General Issues). New York: Springer.

Motriuk, R. W. (1996). Reciprocating compressor valve failure - digital mod-elling and analysis. In Proceedings of the 1st International Pipeline Conference, IPC. Part 2 (of 2), ASME(2), 993–1002.

Naid, A. (2009). Fault detection and diagnosis of reciprocating compressors using motor current signature analysis. PhD thesis, University of Huddersfield.

National Instrument Company. (2000). LabWindows/CVI User Manual, Version 5.5. Internet, London, UK.

Patki, P. S., & V.Kelkar. (2013). Classification using Different Normalization Techniques in Support Vector Machine International Journal of Computer Applications 0975 (8887).

Peters, R. (2002). Machine and Systems Condition Monitoring Series.

Psorakis, I., Damoulas, T., & Girolami, M. A. (2010a). Multiclass relevance vector machines: sparsity and accuracy. IEEE Transactions on Neural Networks, 21(10), 1588-1598.

Psorakis, I., Damoulas, T., & Girolami, M. A. (2010b). Multiclass relevance vector machines: sparsity and accuracy. Neural Networks, IEEE Transactions on, 21(10), 1588-1598.

Qin, S. J. (2003). Statistical Process Monitoring: Basics and Beyond. Journal of Chemometrics, 17, 480.

Quillen, C. D., & Webster, O. H. (2001). Continuing patent applications and performance of the US patent and trademark office. Fed. Cir. BJ, 11, 1.

Rafiee, J., Arvani, F., Harifi, A., & Sadeghi, M. H. (2007). Intelligent condition monitoring of a gearbox using artificial neural network. Mechanical Systems and Signal Processing, 21, 1746-1754.

Razali, N. M., & Geraghty, J. Genetic Algorithm Performance with Different Selection Strategies in Solving TSP. Procedings of the World Congress on Engineering II WCE.

Ren, Q., Ma, X., & Miao, G. (2005). Application of support vector machine in reciprocating compressor valve fault diagnosis. Lecture Notes in Computer Science, 12, 81–84.

Rychetsky, M. (2001). Algorithms and architecture for machine learning based on regularised neural networks and support vector approaches. Sheker-Verlag.

Samanta, B. (2004). Gear Fault Detection Using Artificial Neural Networks and Support Vector Machines with Genetic Algorithms. Mechanical Systems and Signal Processing, 18, 625-644.

Samanta, B., Al-Balushi, K. R., & Al-Araimi, S. A. (2004). Bearing fault detection using artificial neural networks and genetic algorithm. EURASIP J. Appl. Signal Process., 2004, 366-377. doi: http://dx.doi.org/10.1155/S1110865704310085

Samanta, B., Al-Balushi, K. R., & Al-Araimi, S. A. (2006). Artificial neural networks and genetic algorithm for bearing fault detection. Soft Computing - A Fusion of Foundations, Methodologies and Applications, 10(3), 264-271. doi: 10.1007/s00500-005-0481-0

Saxena, A., & Saad, A. (2006). Genetic Algorithms for Artificial Neural Net-based Condition Monitoring System Design for Rotating Mechanical Systems. Applied Soft Computing Technologies: The Challenge of Complexity Advances in Soft Computing, 34, 135-149.

Saxena, A., & Saad, A. (2007). Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems. [doi: DOI: 10.1016/j.asoc.2005.10.001]. Applied Soft Computing, 7(1), 441-454.

Sen, Z. (2001). Genetic algorithms for the classification and prediction of precipitation occurrence. Hydrological Sciences-Journal, 46(2), 255-267.

Shaffer, R. E., Rose-Pehrsson, S. L., & McGill, R. A. (1999). A Comparison Study of Chemical Sensor Array Pattern Recognition Algorithms. Analytica Chimica Acta, 384, 305-317.

Sheng, F., Jing, L., & Yabin, Z. (2009). Fault Diagnosis System for Reciprocating Air Compressor Based on Support Vector Machine. Paper presented at the Proceedings of the 2009 International Workshop on Information Security and Application, Qingdao.

Shi, Z. (2004). Hybrid Model-based Approach to Condition Monitoring. Manchester: University of Manchester.

Shieh, M.-D., & Yang, C.-C. (2008). Classification model for product form design using fuzzy support vector machines. Computers & Industrial Engineering, 55(1), 150-164.

Singh, S., & Singh, A. (2011). Condition monitoring of reciprocating air compressor using IR thermography. Indian Journal of Pure & Applied Physics, 49, 35-41.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and computing, 14(3), 199-222.

Specht, D. F. (1988). *Probabilistic neural network for classification, map, or associative memory.* Paper presented at the Proceeding of the IEEE International Conference on Neural Network.

Stronach, A., Johnston, A., & Cudworth, C. (1984). Detection of Valve Faults on Reciprocating Compressors. Pineridge Press Ltd, 293-304.

Suykens, J. e. a. (2003). Least squares support vector machines. World Scientific Publishing Company, London.

Tafazzoli, E., & Saif, M. (2009). Application of combined support vector machines in process fault diagnosis. Paper presented at the 2009 American Control Conference, St. Louis.

Tang, K. S., Man, K. F., Kwong, S., & He, Q. (1996). Genetic Algorithms and Their Applications. IEEE Signal Processing Magazine, 21-37.

Tipping, A., & Faul, A. (2002). Analysis of sparse Bayesian learning. Advances in neural information processing systems, 14, 383-389.

Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. The Journal of Machine Learning Research, 1, 211-244.

Tiwaria, A., & Yadav, P. (2005). Application of ANN in condition monitoring of a deefective reciprocating air compressor. J. Instrum. Soc. India, 38(I), 7.

To, C. (1984). The acoustic simulation and analysis of complicated reciprocating compressor piping systems, I: Analysis technique and parameter matrices of acoustic elements. Journal of Sound and Vibration, 96(2), 175-194.

Tran, V. T., Yang, B.-S., Gu, F., & Ball, A. (2013). Thermal image enhancement using bi-dimensional empirical mode decomposition in combination with relevance vector machine for rotating machinery fault diagnosis. Mechanical Systems and Signal Processing, 38(2), 601-614.

Tripathy, M., Maheshwar, R. P., & Verma, H. K. (2010). Improved transformer protection using probabilistic neural network power differential method. International Journal of Engineering, Science and Technology, 2(3), 29-44.

Tumer, I. Y., & Huff, E. M. (2002). Principal component analysis of tri-axial vibration data from helicopter transmission. Paper presented at the 56th Meeting of the Society of Machine Failure Prevention Technology.

Tzikas, D., Likas, A., & Galatsanos, N. (2007). Large scale multikernel relevance vector machine for object detection. International Journal on Artificial Intelligence Tools, 16(06), 967-979.

Tzikas, D. G., Wei, L., Likas, A., Yang, Y., & Galatsanos, N. P. (2006). A tutorial on relevance vector machines for regression and classification with applications. EURASIP News Letter, 17, 2, 4, 23.

Vachtsevanos, G., Lewis, F., Roemer, M., Hess, A., & Wu, B. Intelligent fault diagnosis and prognosis for engineering systems, 2006. Usa 454p Isbn, 13, 978-970.

Vapnik, V. (1999). The nature of statistical learning theory. Springer, New York.

Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. New York: Springer.

Villegas, T., Fuente, M. J., & Rodríguez, M. (2010). Principal component analysis for fault detection and diagnosis. experience with a pilot plant. CIMMACS '10 Proceedings of the 9th WSEAS international conference on computational intelligence, man-machine systems and cybernetics, 147-152

Vlachos, M., Domeniconi, C., Gunopulos, G., & Kollios, G. (2002). Non-Linear Dimensionality Reduction Techniques for Classification and Visualization. Paper presented at the Proceedings of 8th SIGKDD, Edmonton, Canada.

Volos, C., & Neri, F. (2012). An introduction to the special issue: Recent advances in defense systems: Applications, methodology, technology. WSEAS Transactions on Systems, 11(9), 477-478.

Wade, B. (1964). Operating Maintenance and Parts Manual. Bucks. England.

Wang, K. (2003). Intelligent condition monitoring and diagnosis systems: a computational intelligence approach (Vol. 93): IOS press.

Wei, C., Chih, H., Lin, J. . (2002). A comparison of methods for multi-class support vector machines. IEEE Trans. On Neural Networks, 13(2), 415-425.

Weston, J., Watkins, C., & (1999). Multi-class support vector machines. Proceedings of ESANN99, M.Verleysen, Ed., Brussels, Belgium.

Whitley, D. (1989). The genitor algorithm and selection pressure: Why rank based allocation of reproductive trials is the best In Proceedings of the 3rd International Conference on Genetic Algorithms.

Widodo, A., Kim, E. Y., Son, J.-D., Yang, B.-S., Tan, A. C. C., Gu, D.-S., . . . Mathew, J. (2009). Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine. Expert Systems with Applications, 36, 7252-7261.

Widodo, A., & Yang, B.-S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing, 21, 2560-2574.

Widodo, A., Yang, B-S. . (2007). Support vector machine in condition monitoring and fault diagnosis. Science Direct, Mechanical Systems and Signal Processing, 21(2560–2574).

Winandy, E., Saavedra O, C., & Lebrun, J. (2002). Simplified modelling of an open-type reciprocating compressor. International journal of thermal sciences, 41(2), 183-192.

Wise, B. M., & Gallagher, N. B. (1996). The Process Chemometrics Approached to process and fault detection. J. of Process Control, 6(6), 329-348.

Worden K, M., G., Hilson G., Pierce S.G. (January 2008). Genetic optimisation of a neural damage locator Original. Journal of Sound and Vibration, 309(3-5, 22), 213-226.

Xu, Z., Xuan, J., Shi, T., Wu, B., & Hu, Y. (2009a). Application of a modified fuzzy ARTMAP with feature-weight learning for the fault diagnosis of bearing. Expert Systems with Applications, 36(6), 9961-9968.

Xu, Z., Xuan, J., Shi, T., Wu, B., & Hu, Y. (2009b). A novel fault diagnosis method of bearing based on improved fuzzy ARTMAP and modified distance discriminant technique. Expert Systems with Applications, 36(9), 11801-11807.

Yadav, M., & Wadhwani, D. S. (2011). Vabration analysis of bearing for fault detection using time domain features and neural network. International journal of Applied Reserch in Mechanical Engineering, 1(1), 6.

Yang, B.-S., Han, T., & Hwang, W.-W. (2005). Fault diagnosis of rotating machinery based on multi-class support vector machines. Journal of Mechanical Science and Technology, 19(3), 846-859.

Yang, B.-S., HAN, T., & YIN, Z.-j. (2006). fault diagnosis system of induction motors using features extraction, features selection and classification algorithm. JSME International Journal Series C, 49(3), 734-741.

Yang, B.-S., Han, T., Hwang, W-W. . (2005). Fault diagnosis of rotating machinery based on multi-class support vector machines. J. Mech. Sci. and Tech, 19(3) 846-859.

Yang, B.-S., Hwang, W.-W., Kim, D.-J., & Chit Tan, A. (2005). Condition classification of small reciprocating compressor for refrigerators using artificial neural networks and support vector machines. [doi: DOI: 10.1016/j.ymssp.2004.06.002]. Mechanical Systems and Signal Processing, 19(2), 371-390.

Yang, B. S., Hwang, W. W., Kim, D. J., & Tan, A. C. C. (2005). Condition classification of small reciprocating compressor for refrigerators using artificial neural networks and support vector machines. Mechanical System and Signal Processing, 19(2), 371–390.

Yang, B. S., Hwang, W. W., Ko, M. H., & Lee, S. J. (2005). Cavitation detection of butterfly valve using support vector machines. Journal of Sound Vibration, 287(1-2), 25–43.

Yang, J., Pu, L., Wang, Z., Zhou, Y., & Yan, X. (2001). Fault detection in a diesel engine by analysing the instantaneous angular speed. Mechanical Systems and Signal Processing, 15(3), 549-564.

Yang, W.-X. (2006). Establishment of the mathematical model for diagnosing the engine valve faults by genetic programming. Journal of Sound and Vibration, 293(1), 213-226.

Yuan, J., Wang, K., Yu, T., & Fang, M. (2007). Integrating relevance vector machines and genetic algorithms for optimisation of seed-separating process. Engineering Application of Artificial Intelligence, 20, 970-979.

Zhang, Y., Jiang, J., Flatley, M., & Hill, B. (2001). Condition monitoring and fault detection of a compressor using signal processing techniques. Paper presented at the American Control Conference, 2001. Proceedings of the 2001.

Zheng. (2005). Numerical Simulation of a Multi-Cylinder reciprocating compressor for condition monitoring. MPhil Thesis, School of Mechanical, Aerospace and Civil Engineering, the University of Manchester, UK.

Zhu, D., Bai, J., & Yang, S. (2009). A Multi-Fault Diagnosis Method for Sensor Systems Based on Principle Component Analysis. Sensors, 10(1), 241-253.