

GOLDFIELDS LIBRARY

VAAL UNIVERSITY OF TECHNOLOGY



VIBRATION CONDITION MONITORING AND FAULT
CLASSIFICATION OF ROLLING ELEMENT BEARINGS
UTILISING KOHONEN'S SELF-ORGANISING MAPS

by

J.S.R. Nkuna

(Student number : 9817336)

A Dissertation

Submitted to the Faculty of Engineering and Technology in Complete Fulfilment of
the Requirements for the Degree of *Magister Technologiae* in Mechanical
Engineering, Department of Mechanical Engineering

September, 2006

Supervisor : J. Enslin

Co-Supervisor : Prof. D.F. van der Merwe

VAAL UNIVERSITY OF TECHNOLOGY	
Bib No:	11153994
Item No:	11408081
Order No:	Donation
2006-08-29	
Price:	R350.00
Call No.	b20.3 NKU
LIBRARY STOCK	

© J.S.R. Nkuna 2006

VAAL UNIVERSITY OF TECHNOLOGY
FACULTY OF ENGINEERING AND TECHNOLOGY

The undersigned certify that they have read, and recommend to the Faculty of Engineering and Technology for acceptance, a dissertation entitled “VIBRATION CONDITION MONITORING AND FAULT CLASSIFICATION OF ROLLING ELEMENT BEARINGS UTILISING KOHONEN’S SELF-ORGANISING MAPS” submitted by J.S.R. Nkuna in partial fulfilment of the requirements for the degree of *Magister Technologiae* in Mechanical Engineering.



Mr. J.G. Drotsky

Head of Mechanical Engineering Department



Prof. P. Mendonidis

Faculty of Engineering and Technology

Prof. H. J. de Jager

Faculty of Engineering and Technology

15/09/2006
Date

Declaration

I Jay Shipalani Rhulani Nkuna declare that this dissertation is the result of my own independent work and investigation, except where otherwise stated, and no words are copied directly from reference texts. Other sources are acknowledged by giving explicit references. The references are appended. No part of this dissertation has been previously accepted or submitted in substance for any degree, and is not being concurrently submitted in candidature for any degree in any university or other institution. The dissertation highlights experiences and gain of knowledge that I have attained. I have received guidance from my Supervisor and Co-supervisor in writing of the dissertation, and evaluation of technical content.



Mr. J.S.R. Nkuna

Researcher



Mr. J. Enslin

Supervisor

Prof. D.F. van der Merwe

Co-Supervisor

15/09/2006
Date

Acknowledgements

I would like to express my appreciation towards the lecturers of the faculty of Mechanical Engineering that helped me complete this dissertation. I am most grateful to the following people for reviewing the dissertation and offering comments and suggestions : Supervisor Mr J. Enslin, Co-supervisor Prof. D.F van der Merwe, manufacturing technicians of the mechanical workshop in R-block.

It has been gratifying to work with them through-out this project. I would like to thank the Vaal University of Technology for granting me the opportunity of compiling the dissertation. Appreciation is also expressed towards the following:

- The national research foundation, and the council for scientific and industrial research for the funding of this research.
- The Vaal University of Technology for the use of the Iscor innovation centre and the maintenance workshop of the Faculty of Engineering.
- The staff of the maintenance workshop for their support.

I wish to thank my Father, Mother, Sister and Brother without whose patience, encouragement and support this dissertation might not have been completed. Finally, I wish to thank the Almighty God for the opportunity, strength and courage that He gave me to complete this dissertation.

Dedication

This dissertation is dedicated to Wisani Nkuna, Khensani Nkuna, Manini Nkuna, Vongani Nkuna and Ramasele Masonto.

Abstract

Bearing condition monitoring and fault diagnosis have been studied for many years. Popular techniques are applied through advanced signal processing and pattern recognition technologies. The subject of the research was vibration condition monitoring of incipient damage in rolling element bearings. The research was confined to deep-groove ball bearings because of their common applications in industry. The aim of the research was to apply neural networks to vibration condition monitoring of rolling element bearings. Kohonen's Self-Organising Feature Map is the neural network that was used to enable an automatic condition monitoring system.

Bearing vibration is induced during bearing operation and the main cause is bearing friction, which ultimately causes wear and incipient spalling in a rolling element bearing. To obtain rolling element bearing vibrations a condition monitoring test rig for rolling element bearings had to be designed and built. A digital vibration measurement acquisition environment was created in Labview and Matlab. Data from the bearing test rig was recorded with a piezoelectric accelerometer, and an S-type load cell connected to dynamic signal analysis cards. The vibration measurement instrumentation was cost-effective and yielded accurate and repeatable measurements. Defects on rolling element bearings were artificially inflicted so that a pattern of bearing defects could be established. An input data format of vibration statistical parameters was created using the time and frequency domain signals. Kohonen's Self-Organising Feature Maps were trained in the input data, utilising an unsupervised, competitive learning algorithm and vector quantisation to cluster the bearing defects on a two-dimensional topographical map.

A new practical dimension to condition monitoring of rolling element bearings was developed. The use of time domain and frequency domain analysis of bearing vibration has been combined with a visual and classification analysis of distinct bearing defects through the application of the Self-Organising Feature Map. This is a suitable technique for rolling element bearing defect detection, remaining bearing life estimation and to assist in planning maintenance schedules.

Table of Contents

	Page Number
Declaration	ii
Acknowledgements	iii
Dedication	iv
Abstract	v
List of Figures	xiii
List of Tables	xvi
List of Annexures	xvii
Glossary of Abbreviations	xviii
Nomenclature	xix
Chapter 1 : Introduction	1
1.1 Introduction	1
1.2 Background	1
1.3 Introduction to Neural Networks	5
1.4 Importance of the Research	7
1.5 Purpose of the Study	8
1.6 Research Problem Statement	8
1.7 Research Objectives	9
1.8 Delimitations	9
1.9 Assumptions	10
1.10 Scope of Work	11
1.11 Process Flow Chart	11
1.12 Dissertation Overview	12

Chapter 2 : Review of the Related Literature	14
2.1 Introduction	14
2.2 Evolution of Rolling Element Bearings	14
2.3 Vibration Condition Monitoring of Rolling Element Bearings	15
2.4 Rolling Element Bearings Condition Monitoring Techniques	17
2.5 Human Versus Artificial Intelligence Condition Monitoring	19
2.6 Signal Processing of Bearing Vibration Waveforms	20
2.7 Statistical Analysis on Rolling element Bearing vibrations	22
2.8 Induced Rolling Element Bearing Defects	23
2.9 Historical Development of Kohonen Neural Networks	25
2.10 Biological Analogy to Artificial Intelligence Neural Networks	27
2.11 Development of the Self-Organising Map Algorithm	28
2.11.1 Competitive Algorithm	28
2.11.2 Data Pre-processing and Feature Extraction	29
2.11.3 Output Topology and Classification	30
2.11.4 Learning Vector Quantisation Neural Network	31
2.12 Applied Neural Networks on Vibration Monitoring Systems	32
2.12.1 Rolling Element Bearing Condition Monitoring	32
2.12.2 Machining Tool Condition Monitoring	35
2.12.3 Machinery Condition Monitoring	35
2.13 Summary	36
Chapter 3 : Bearing Vibration Fundamentals	37
3.1 Introduction	37
3.2 Types of Rolling Element Bearings	38
3.3 Deep-Groove Ball Bearing Nomenclature	38
3.4 Design and Manufacture of rolling element bearings	40

	Page Number
4.8	Time Waveform Statistical Parameters 62
4.9	Data Acquisition and Signal Processing Software 67
4.10	Filtering of the Bearing Vibration Measurements 69
4.11	Induced Bearing Frequencies 70
4.12	Loading Configuration 71
	4.12.1 Rolling Element Bearing Life 71
4.13	Vibration Measurement Acquisition 73
	4.13.1 Natural Responses of the Rolling Element Bearings 73
	4.13.2 Natural Responses of the Condition Monitoring Rig 74
4.14	Bearing Vibration Defect Pattern Measurements 75
	4.14.1 Outer Race Incipient Defect Bearing Noise 75
	4.14.2 New Bearing Noise 76
	4.14.3 Ball Bearing Defect 78
	4.14.4 Inner Race Spall Defect 79
	4.14.5 Outer Race Crack Defect 80
4.15	Bearing Vibration Defect Severity Measurements 81
	4.15.1 Stage one Bearing Defect 81
	4.15.2 Stage Two Bearing Defect 81
	4.15.3 Stage Three Bearing Defect 83
	4.15.4 Stage Four Bearing Defect 84
4.16	Summary 85
Chapter 5	Application of Self-Organising Feature Maps 86
5.1	Introduction 86
5.2	Methodology 86
5.2	Self Organising Feature Map Architecture 87
5.3	Unsupervised Learning Algorithm 90
	5.3.1 Hebbian Algorithm 90

	Page Number
5.3.2 Oja’s Rule	91
5.3.3 Competitive Learning Algorithm	91
5.3.4 The Learning Vector Quantisation Algorithm	92
5.3.5 The k-means Clustering Algorithm	93
5.4 The Self-organising Feature Map Algorithm	94
5.4.1 Competitive Transfer function	94
5.4.2 Competition	95
5.4.2.1 Vector Quantisation	95
5.4.3 Cooperation	97
5.4.4 Adaptation	98
5.5 Training Procedure	100
5.5.1 Normalisation	101
5.5.2 Initialisation	102
5.5.3 Adaptive Adjustment	104
5.5.3.1 Quantisation error	105
5.5.3.2 Topological Error	105
5.6 Parameter Classification	106
5.7 Summary	107
Chapter 6: Experimentation	108
6.1 Introduction	108
6.2 Creation of an Inclusive Parameters Input Data Set	108
6.3 Compared Error Performances	111
6.4 Kohonen’s Feature Map in Parameters Input Data	114
6.5 Prognosis and Visualisation of Rolling Element Bearing Defects	118
6.6 Summary	127

	Page Number
Chapter 7 : Conclusion and Recommendations	128
7.1 Summary and Conclusions	128
7.2 Recommendations	129
7.3 Fields for Further Study	130
Annexures 1	132
Annexures 2	137
Annexures 3	138
Annexures 4	139
Annexures 5	140
Annexures 6	142
Annexures 7	143
Annexures 8	148
References	150

List of Figures

Page Number

1.1	Bearing Life Span : ‘The Bath-Tub Curve’	3
1.2	Kohonen’s Map Topological Architecture	7
1.3	The proposed condition monitoring process flow chart	11
3.1	Deep-Groove Ball Bearing Nomenclature	39
3.2	Front View of a Deep-Groove Ball Bearing	44
3.3	Side View of a Deep-Groove Ball Bearing	44
3.4	Rolling Element Bearing Condition Monitoring Rig	48
3.5	Test Bearing Shaft and V-Belt Drive System	49
3.6	Bearing under a Bearing Seat and Stepped Chuck Mounting Mechanism	49
3.7	Test Bearing Loading Mechanism with Load Cell and Accelerometer	49
3.8	An Artificially Inflicted Bearing Spall of $\pm 2\text{mm}$	52
4.1	Time Domain (a) and Frequency Domain (b) Waveform Representation	55
4.2	A Three Sinusoid Waveform Representation	57
4.3	Random Vibration, Time and Frequency Response	58
4.4	Virtual Instrument Construction Flow Diagram	68
4.5	Online Signal Processing Virtual Instrument Display Window	69
4.6	New Bearing Unfiltered Signal Time Response	69
4.7	New Bearing Filtered Signal Time Response	70
4.8	New Bearing	70
4.9	New Bearing Filtered Signal	70
4.10	Bearing 6012ZZ3C Impact Time Response	73
4.11	Bearing 6012ZZ3C Impact Frequency Response	73
4.12	Loading Mechanism Impact Time Response	74
4.13	Loading Mechanism Impact Frequency Response	74

	Page Number
4.14 Noisy Bearing Time Response	75
4.15 Noisy Bearing Frequency Response	76
4.16 New Bearing Time Response	77
4.17 New Bearing Frequency Response	77
4.18 Ball Cut Time Response	78
4.19 Ball Cut Frequency Response	78
4.20 ± 1 mm Inner Race Defect Time Response	79
4.21 ± 1 mm Inner Race Defect Frequency Response	79
4.22 Outer Ring Cut Time Response	80
4.23 Outer Ring Cut Frequency Response	80
4.24 Stage One New Bearing	81
4.25 Stage One Noisy Bearing	81
4.26 Stage Two Outer Race Defect Time Response	82
4.27 Stage Two Outer Race Defect Probability Density Distribution	82
4.28 Stage Two Outer Race Defect Frequency Response	82
4.29 Stage Three Outer Race Defect Time Response	83
4.30 Stage Three Outer Race Defect Probability Density Distribution	83
4.31 Stage Three Outer Race Defect Frequency Response	83
4.32 Stage Four Outer Race Defect Time Response	84
4.33 Stage Four Outer Race Defect Probability Density Distribution	84
4.34 Stage Four Outer Race Defect Frequency Response	84
5.1 Symbolic Architecture of Kohonen's Self-organising Feature Map	87
5.2 Map Topology	102
5.3 Trained Map in Euclidean Space	103
5.4 Labelled Classification of Trained Map	104
5.5 Parameter Classification Maps	107
6.1 Bearing 6008ZZ3C Error vs. Dimension	112

	Page Number
6.2 Bearing 6010ZZ3C Error vs. Dimension	113
6.3 Bearing 6012ZZ3C Error vs. Dimension	114
6.4 Trained Map Projection in B12 Input Data Euclidean Space	116
6.5 B12 Trained Map Cluster Distributions and Defect Labelling	116
6.6 Trained Map Projection in B12_15Hz_5kg Input Data Euclidean Space	117
6.7 B12_15Hz_5kg Trained Map Cluster Distributions and Defect Labelling	117
6.8 B12 Nine Parameters Self-Organising Feature Maps	118
6.9 B12_15Hz_5kg Nine Parameters Self-Organising Feature Maps	120
6.10 B12 Nine Parameters Defect Projections, SOM	121
6.11 B12_15Hz_5kg Nine Parameters Defect Projections, SOM	121
6.12 B12 Speed Clustering and Defect Projection	122
6.13 B12 Load Clustering Defect Projection	122
6.14 B12 Unit Matrix Clusters with Best Matching Units	123
6.15 B12_15Hz_5kg Unit-Matrix Clusters with Best Matching Units	123
6.16 B12 Davies Bouldin Clustering	124
6.17 B12_15Hz_5kg Davies Bouldin Clustering	125
6.18 B12 Learning Vector Quantisation Defect Clustering	125
6.19 B12_15Hz_5kg Learning Vector Quantisation Defect Clustering	126
A5.1 Bearing 6008ZZ3C Frequency Response	140
A5.2 Bearing 6010ZZ3C Frequency Response	140
A5.3 Electric Motor Casing Frequency Response	140
A5.4 Base-plate Frequency Response	141
A5.5 Belt Resonance Frequency Response	141
A5.6 Bearing Misalignment Frequency Response	141

List of Tables

	Page Number
2.1 Causes of Rolling Element Bearing Failure	17
3.1 Symbols for Figure 3.2 and Figure 3.3	45
4.1 Condition Monitoring Testing Speeds	71
4.2 Experimentation Induced Bearing Frequencies	71
5.1 Input Data Set of a 6008ZZ3C Deep-Groove Ball Bearing	96
6.1 Statistical Parameters Input Training Data	110
6.2 Dimension Size of Input Data Sets	111
6.3 Training Errors for Bearing 6008 Input Data Sets	111
6.4 Training Errors for Bearing 6010 Input Data Sets	112
6.5 Training Errors for Bearing 6012 Input Data Sets	113
A2.1 Bearing Characteristic Equations	137
A3.1 Bearing Dimensions	138
A4.1 Test Rig Loading Adjustments	139
A4.2 Bearing Design Load Ratings	139

List of Annexures

		Page Number
1	Derivation of Bearing Characteristic Equations	180
2	Induced Bearing Frequency Equations with Measurement and Applications	180
3	Induced Bearing Frequencies	182
4	Loading Configurations	183
5	Natural Responses	185
6	Input Data Feature Parameters	186
7	Automatic Bearing Defect Classification	187
8	SOM Training Report	189

Glossary of Abbreviations

Symbol	Description
ART	Adaptive Resonance Technique
BPFI	Ball Pass Frequency of Inner race
BPFO	Ball Pass Frequency of Outer race
BSF	Ball Spin Frequency
Cf	Crest Factor
EVAM	Evaluated Vibration Analysis Method
FTF	Fundamental Train Frequency
HFD	High Frequency Detection
K	Kurtosis
KNN	Kohonen's Neural Network
LVQ	Learning Vector Quantisation
PDD	Probability Density Distribution
PDF	Probability Density Function
Pk	Peak Amplitude
RA	Rectified Average
RBF	Radial Bias Function
RMS	Root Mean Square
Skw	Skewness
SOM	Self-Organising Map
SOFM	Self-Organising Feature Map
Spd	Speed
SPM	Shock Pulse Method
Std	Standard Deviation

Nomenclature

Symbol	Description	Metric units
N_s	Shaft rotational speed	rev/min
V_b	linear speed of the ball	m/s
V_c	linear speed of the cage	m/s
V_i	linear speed of the inner race	m/s
V_o	linear speed of the outer race	m/s
ω_b	rotational speed of the ball	rad/s
ω_c	rotational speed of the cage	rad/s
ω_i	rotational speed of the inner race	rad/s
ω_o	rotational speed of the outer race	rad/s
r_c	rotational radius of the cage	mm
r_i	rotational radius of the inner race	mm
r_o	rotational radius of the outer race	mm
r_r	rotational radius of the balls	mm
α	Contact angle	Degrees
B_d	Ball Diameter	mm
D_i	Bore diameter	mm
D_o	Outer ring diameter	mm
P_d	Pitch circle diameter	mm
P_c	Pitch circle diameter	mm
P_e	End Play	mm
Z	Number of ball elements	
W_n	normalised cut-off frequency	Hz
W_p	pass-band frequency	Hz
W_s	Stop-band or cut-off frequency	Hz
R_p	pass-band ripple	dB
R_s	stop-band attenuation	dB
P	filter frequency order	
A	Amplitude, milli-gravitational acceleration, ($g = 9.81\text{m/s}^2$)	mg
T	Time	s

Chapter 1

Research Introduction

1.1 Introduction

The subject of the research is vibration condition monitoring of rolling element bearings. The scope of the research and the layout of the dissertation are given in this chapter. Some of the relevant literature is mentioned, but the proper literature review is in chapter 2.

A review of the relevant literature provides the background to the discussions of the problems associated with bearing damage and failure methods of analysis. The development of bearing damage is explained with the assistance of the bath-tub curve. An introductory overview to neural networks and the Self-Organising Map neural network is provided. This is followed by an explanation of the objectives, purpose, delimitations and assumptions for the research. Then the scope and expectations of the research are discussed, giving the layout of the dissertation.

1.2 Background

Shiroishi *et al.* (1997:693) wrote that bearings are of paramount importance to almost all forms of rotating machinery and are among the most common of machine elements. As a consequence of their importance and widespread use, bearing failure is one of the foremost causes of breakdown in rotating machinery. Research on the cause and analysis of bearing failure has been conducted for over four decades. Bearing vibration monitoring and analysis provides many benefits in terms of optimising process efficiency. Costly downtime can be minimised or eliminated by providing early warning of potential problems, thus preventing catastrophic machine failures.

By studying vibration changes over time, the condition of the monitored bearing can be determined and a maintenance schedule planned. A decision on plant shut down can be made with minimal disruption to production.

SKF (2004:31) reported that in South Africa the cost of maintenance on machinery overhauls for bearing replacements on industrial rollers in conveyor idlers is very high, dictating the necessity for a condition monitoring system that can be applied in predictive maintenance. Great savings in time, effort and expense can be achieved if the bearing user can establish the reason for premature failure and undertake corrective action to prevent further failure. Information on the extent of the bearing damage can provide good guidance on scheduling maintenance routines.

Chen (2000:2) reported that the most important and expensive task in terms of labour time and cost in machinery maintenance is fault detection and diagnostics. Without accurate identification of machine faults, production maintenance and scheduling cannot be effectively planned for necessary repair work. In addition, accurate fault detection and diagnosis is essential for reducing troubleshooting and repair time. Moreover, machine reliability is improved by correct fault diagnosis.

Li *et al.* (1997:693) noted that several fault detection techniques employing time and frequency domain analysis have failed to consistently predict bearing damage and failure. This observation is corroborated by Al-Gahmd *et al.* (2004:1) who states that the extent of bearing damage frequently eludes the diagnostician.

Barkov *et al.* (1995a:1) suggested that the operational life of rolling element bearings can be divided in four main stages or periods according to the extent of bearing damage conditions: running-in, operation without defects, development of wear defects and increased bearing damage degradation.

During the first three stages of bearing lifetime, up to bearing damage degradation, it is possible to make an accurate assessment of bearing condition and lifetime. To accomplish this successfully it is necessary to detect and identify all the bearing defects that influence its operational life. At the final stage of bearing degradation, the remaining operational life is predicted mainly by the magnitude and rate of change of measured parameters. Bearing condition assessment and lifetime prediction does not depend on the methods and results of condition assessment used for other machine components. The bearing life span is illustrated in figure 1.1 with the use of the bath-tub curve.

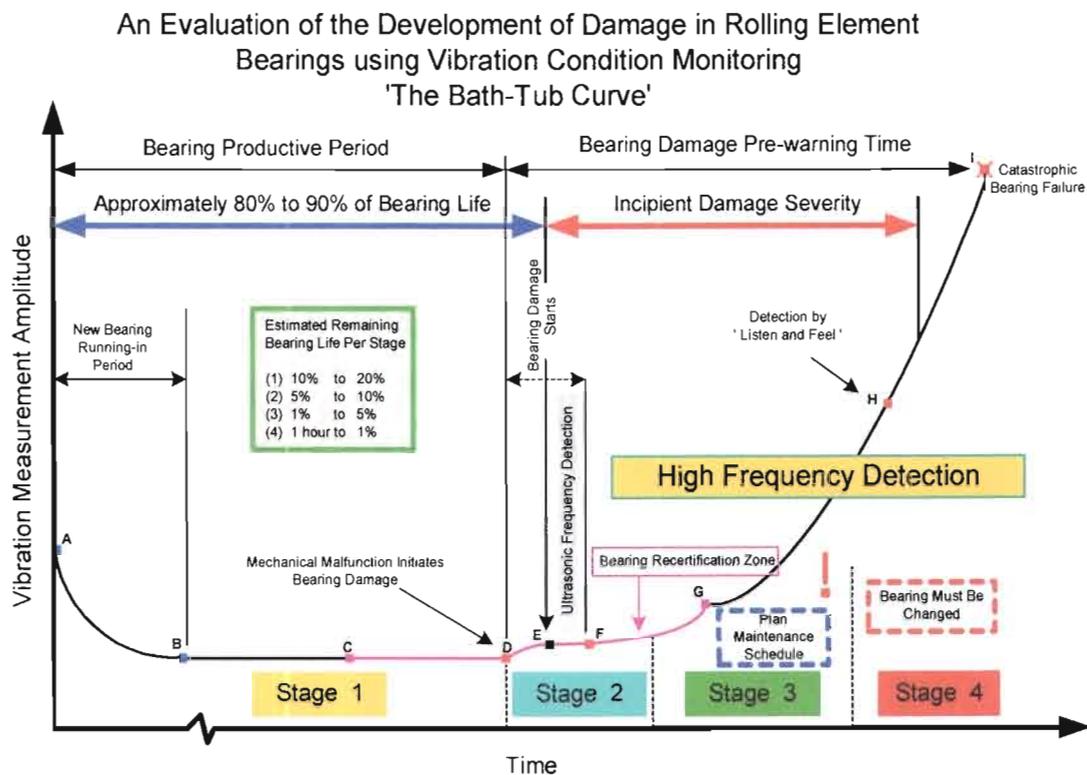


Figure 1.1: Bearing Life Span : ‘The Bath-Tub Curve’

The bearing life begins at point A, when the bearing is first mounted on the shaft and machine operation takes place. The first bearing vibration measurement may show peaks as the lubrication spreads between the surfaces of the bearing components.

The earliest incipient damage initiates at point D. Ultrasonic methods may detect such damage. High frequency detection techniques indicate damage from point F onwards. The severity of the incipient damage increase through stages 2, 3, to stage 4.

BAT Services (2004:5) also described four stages of rolling element bearing damage progression to failure. Stage one is recorded when a new bearing is installed. During stage two race way surface defects begin and bearing noise increases. During stage three, induced bearing frequencies are clearly noticeable on a frequency spectrum. Stage four is characterised by noise from the bearing, which is an unmistakable sign that the bearing must be changed. The vibration frequencies then appear with side bands on the frequency spectrum.

Taylor *et al.*(2004:10) noted that, for bearings with rotating inner race and stationary outer race, the bearing normally fails in the following order: (1) the outer race in the load zone fails since the load zone does not shift on the outer race; then (2) the inner race, since a single location on the inner race shifts in and out of the load zone with each shaft rotation; followed by (3) ball elements or rollers; and lastly (4) cage damage. This is true for a correctly installed bearing, because it stands to reason that if the cage was accidentally damaged during installation then the cage would probably be the first to fail. The patterns of bearing damage requires analysis from a skilled personnel who should be familiar with spectrum analysis and all the techniques most commonly used. Accurate measurements and accurate analysis are difficult to achieve and are the reason for uncertainty when analysing the vibration.

Rao (1995:821) wrote that vibration measurements that are not similar or identical even though the measurements were taken under similar conditions, are random vibrations. Bearing vibrations are an example of random phenomenon. Bearing vibrations are not identical in magnitude of excitation in either force or motion, and the measurement reveals this.

A neural network is trained to perform a particular function by adjusting the numerical values of the connections between elements. The network is adjusted, based on a comparison of the output and the targeted output. The network is repeatedly adjusted until the network output either matches the target or differs by an extremely small amount. When the input is compared to a specific target, the learning is known as supervised learning. Neural networks utilising supervised learning are commonly used, but other networks can be obtained from unsupervised learning or from direct design methods.

Unsupervised learning networks can be used to identify groups of data. Bearing vibration data can be classified into groups of different bearing defects, using unsupervised learning networks. The chosen self organising feature map network uses unsupervised learning and unsupervised training techniques. The supervised learning network can also be used on bearing vibration data where a solution or target output is known. In this case, the supervised learning network is limited to a two conditional output where the data is recognised according to known desired output, and the solution from the network points to a bearing which is either damaged or not damaged, irrespective of the extent of the damage sustained.

Kohonen's Neural Network (KNN) may be referred to as Self-Organising Map (SOM), or Self-Organising Feature Map (SOFM). The SOFM enables good cluster visualisation of a data set. The architectural construction of Kohonen's map is illustrated in figure 1.2. The main components of the architecture are shown, where neurons are arranged in a two dimensional topological layer, and are grouped into clusters.

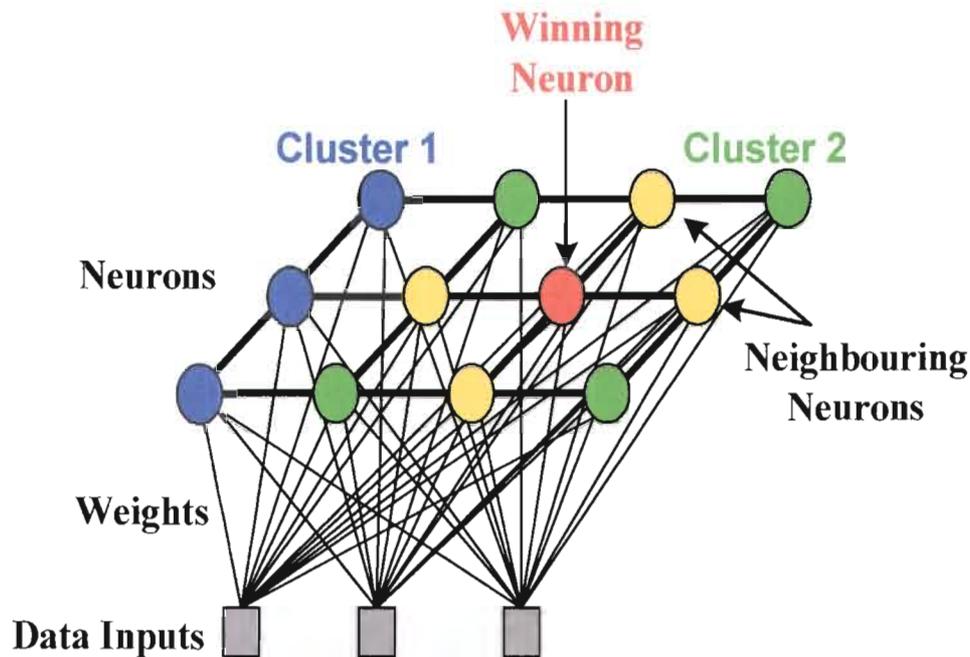


Figure 1.2: Kohonen's Map Topological Architecture

1.4 Importance of the Research

Brown (1977:41) observed that until recently the most common method for condition monitoring relies on either temperature, vibration or noise measurement. Common to all these methods is that by the time a significant change has occurred to be quantifiable as a measurement, the bearing is probably on the verge of collapse.

Brown (1977:42) reported on the importance of rolling element bearings and their performance in modern industry. Rolling element bearings have a limited life that is directly influenced by such factors as their mountings, operational load condition and the routine maintenance they receive.

To avoid plant breakdown due to plant failure, plant engineers have two options open to them:

- To replace bearings well within the bearing manufacturer's recommended lifetime. If this particular technique is used, some bearings will fail before the set number of hours have been reached and other bearings will be replaced unnecessarily.
- To monitor continuously the condition of the bearing and only replace the bearing, regardless of time used, when the monitoring technique indicates failure. In applying this method, the bearing could be on the verge of collapse by the time failure shows.

These observations and stipulated options, highlight the need for a monitoring system that enables early detection of bearing damage and provides automatic reliable analysis of the extent of damage as well as the location of damage on the bearing. The system should enable well timed maintenance routines.

1.5 Purpose of the Study

The purpose of this research project is to develop a condition monitoring system for rolling element bearings that is reliable in providing early detection of bearing defects and quantifies the extent of defect development. The monitoring system will enable the collection of vibration and noise measurements for automatic analysis utilising Kohonen's SOM. The analysis involves signal processing, automatic feature extraction, and fault classification.

1.6 Research Problem Statement

To apply Kohonen's Self-Organising Maps with the automatic evaluation of statistical parameters, to monitor the condition of rolling element bearings.

1.7 Research Objectives

The objective of the research is to create a condition monitoring system that will be able to provide early detection of rolling element bearing damage and provide reliable information on the extent as well as the location of the damage.

The research concentrates on creating a condition monitoring system that is based on artificial intelligence and provides assistance to highly skilled personnel that monitor bearings. The system is to operate automatically while it receives a vibration signal. The monitoring system will enable the collection of vibration and noise measurements to be automatically analysed by Kohonen's Neural Network known as Self-Organising Maps (SOM). The analysis will involve signal processing, automatic feature extraction, fault classification and extent of bearing damage. This condition monitoring system will enable accurate detection of rolling element bearing damage through reliable bearing vibration measurement analysis. The system aims to detect incipient damage of rolling element bearings, in the four stages of bearing damage as described by Barkov *et al.* (1995a:1).

To test the monitoring system, a laboratory scale monitoring rig was built to accommodate deep-groove ball bearings.

1.8 Delimitations

Only rolling element bearings are considered and data acquisition and processing are limited to bearing vibrations.

Tests were performed on deep-groove ball bearings, because of their common application in industry. The size of the test rig constrained the size of bearing that could be tested.

Laboratory tests were performed at different shaft rotational speeds achievable by the bearing monitoring rig. Vibration signals were measured at a specific constant shaft speed and constant applied radial load conditions. An oil lubricant instead of a grease lubricant was used in order to achieve the best bearing vibration measurements due to bearing defects, at the same time avoiding a dry bearing which would result in excessive bearing vibrations. The common type of bearing defect is spalling, thus the considered bearing defects in developing the monitoring system, were induced spalls on specific locations on the bearing. Distinct spalling at specific bearing locations could be established.

1.9 Assumptions

Vibration signals, though random, were captured at a constant rotational speed and assumed to be stationary. In stationary operational conditions, the shaft and inner ring rotational speeds are constant and do not vary, the static loading on the bearing or bearing load when the shaft is not rotating is constant and does not vary. (For varying operational conditions other methods of analysis could be applied such as wavelets or the Gabor's transform).

Loading and lubrication are assumed not to interfere with the vibration measurements for the detection of bearing damage.

All bearings considered for experimentation are assumed to be of the same design and have similar material and manufacturing properties that remain constant during experimentation. Since the shape of a spall defect is irregular, an average diameter of the spall size is assumed.

1.10 Scope of Work

A condition monitoring test rig was designed and fabricated. Labview was used to perform signal pre-processing, and MATLAB was used to implement the neural network programs. The program enabled automatic extraction of features from the measured signal, and processing of the feature vectors that were used as input to the (SOM) algorithm. The Self-Organising Maps used for fault classification were to illustrate the type and extent of the bearing damage and allow a final decision on whether to either keep or replace the bearing. The condition monitoring system was developed and tested in the laboratory for reliability by using bearings with known types of damage deliberately applied before experimentation.

1.11 Process Flow Chart

The following diagram illustrates the steps in performing automatic bearing vibration defect detection and classification.

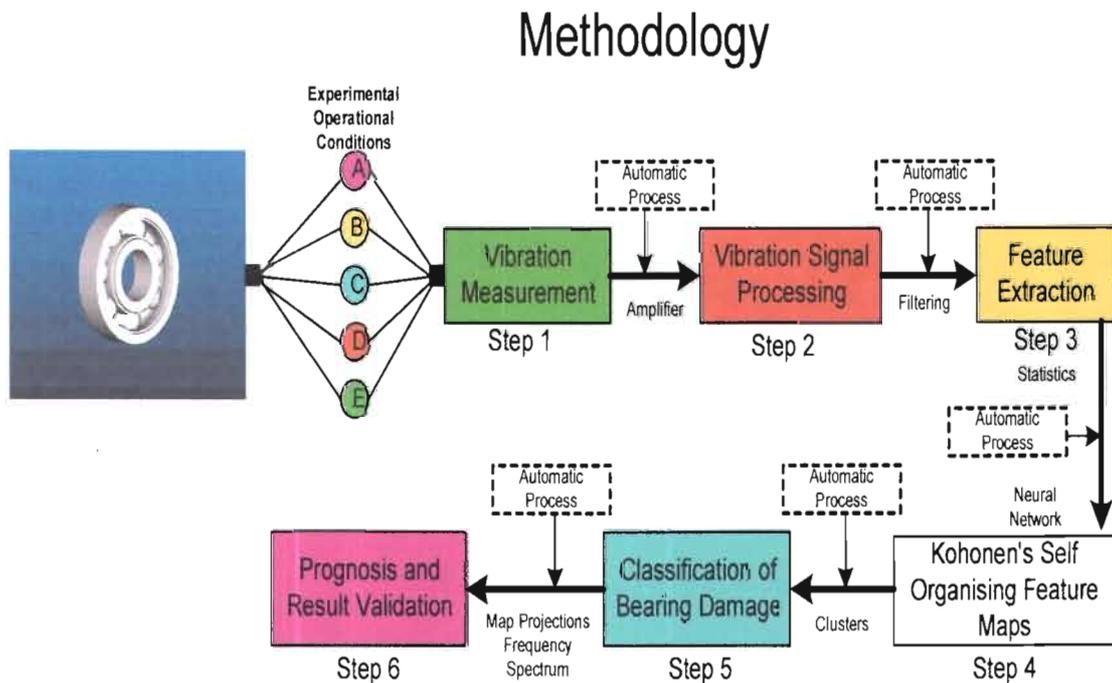


Figure 1.3: The proposed condition monitoring process flow chart

1.12 Dissertation Overview**Chapter 1: Introduction**

Introduces the subject content and research topic. The plan and scope of the research is outlined.

Chapter 2: Review of the Related Literature

Reviews the literature related to vibration condition monitoring using neural networks. A bearing maintenance approach is followed.

Chapter 3: Bearing Vibration Fundamentals

Introduces the definition of a rolling element bearing, operational principles, and vibration characteristics. The design of the monitoring rig is introduced, so that the relationship between the monitoring system and the monitoring rig is established.

Chapter 4: Signal Pre-processing Techniques

Introduces digital signal processing and the creation of the time domain and frequency domain waveform displays. Signal processing tools are developed in the Labview software. Vibration characteristics of the monitoring rig and different bearing defects are displayed on the time domain and frequency domain. This is to illustrate the difference in signal processing patterns of different rolling element bearing defects. Statistical parameters are determined through signal processing.

Chapter 5: Application of Self-Organising Maps

Self-Organising Map algorithms are introduced. A single input data set is created. The process of clustering and classification are explained. A single data set is processed through the learning and training procedures of the algorithm, and the results of the procedures are illustrated.

Chapter 6: Experimentation

Different sizes of input data sets are considered so as to compare the performance of the Self-Organising Map. Ultimately, two sizes of input data sets are recommended, and illustrated in their effectiveness in monitoring rolling element bearing damage. Map classification of the bearing defects is displayed for visualisation and analysis. Classification matrices are interpreted according to the severity and location of the bearing defect considered.

Chapter 7: Conclusions and Recommendations

Achievements of the research are reviewed, and recommendation are made, with suggestions for further studies.

Chapter 2

Review of the Related Literature

2.1 Introduction

Chapter 2, reviews the relevant literature on vibration condition monitoring of rolling element bearings. The literature is focused on experiments conducted on rolling element bearings, with the aim of applying neural networks to vibration condition monitoring of rolling element bearings. Kohonen's Self-Organising Feature Maps will be utilised. The development of Kohonen's Self-Organising Map algorithm is studied and reviewed. The advantages of the self-organising map neural network are determined by reviewing the related literature.

The literature review starts of with an introduction to the development of rolling element bearings and conventional vibration condition monitoring applied to rolling element bearings. The literature review ends with a presentation of case histories where neural networks, specifically Self-Organising Feature Maps are applied to vibration condition monitoring of rolling element bearings. The literature study is aimed at supporting the purpose of the research.

2.2 Evolution of Rolling Element Bearings

Singer (1954:26) an English historian noted that the invention of the wheel gave birth to the journal bearing, where the wheel would rotate on the axle. This bearing was of the rubbing plain type, where the material of the wheel would rub against the axle material.

The study of friction by Leonardo da Vinci (1452 - 1519) resulted in the observation that the analogy of rolling logs situated between two flat surfaces, as in the sledge and log mechanism, provided easier motion of the top surface as compared to when the two surfaces were in contact with one another without the presence of rolling logs. Leonardo da Vinci attributed this phenomenon to friction, and so from his studies, da Vinci conceived the basic construction of the modern rolling element bearing. The rolling logs could be substituted by spherical balls, within a circular surface.

During the twentieth century, the events of the First World War (1914 - 1918) and the Second World War (1939 - 1948) brought a rapid increase in the field of science engineering and manufacturing. During the two world wars the demand for steel rolling bearings on army vehicles and fighter aeroplanes increased. The strength of these bearings was low, owing to their low material strength, and so the machines were subjected to regular repairs which was costly.

Since the 1960's the development of superior rolling bearing steels and constant improvement in manufacturing provided accurate geometry and long life to rolling bearings. This development was triggered mainly by the bearing requirements for high speed aircraft turbines. During the 1970's competition for world-wide markets increased between roller and ball bearings which provided consumers with low cost high endurance bearings. Also during this period maintenance engineering was becoming a structured discipline.

2.3 Vibration Condition Monitoring of Rolling Element Bearings

Birchon (1975:55) states that maintenance covers all activities undertaken to keep equipment in a particular condition or to return it to such a condition.

According to SKF (1992:2), preventive maintenance involves the overhaul of a machine or machine parts on a regular basis regardless of the condition of the parts

before failure is reached. This process however results in excessive downtime due to unnecessary overhauls and the excessive costs of replacing good parts along with worn parts. SKF defined vibration condition monitoring of rolling element bearings as the process of determining the condition of a rolling element bearing by measuring the emanating vibrations as the bearing operates, so that the most efficient and effective repair of the bearing prior to failure may be scheduled.

Van Wyk (1998:42) noted that vibration condition monitoring of rolling element bearings is indirect predictive maintenance, in that overhaul of machinery and bearings is not necessary before a decisive fault is detected. The state of operation of a rolling element bearing in a machine system is predicted through objective vibration condition monitoring with the use of digital signal processing techniques, and subjective condition monitoring which involves human senses.

In the 1970's sophisticated vibration condition monitoring began to be widely applied in industrial maintenance, although some simple inspection techniques using thermocouples and pressure transducers were being used before this. The main objectives of a vibration condition monitoring policy for rolling element bearings were outlined by Birchon (1975:65) as follows: to minimise the risk of unexpected machine breakdown; to diagnose the state of bearing damage, if any, in a bearing, and to specify the work to be performed to restore a pre-determined level of reliability.

Robson (1989:47) wrote an indirect predictive maintenance philosophy with regard to vibration condition monitoring of rolling element bearings. The philosophy states that vibration condition monitoring will provide adequate warning to avert catastrophic bearing failure, and that diagnostic data will be available so that when a warning is given, the bearing will have visible damage.

Table 2.1 lists a number of causes of rolling element bearing failures identified by Mathew and Alfredson (1984:447).

Table 2.1: Causes of Rolling Element Bearing Failure (Adapted from Mathew & Alfredson 1984:447)

Failure Total = 100%	Root Causes of Rolling Element Bearing Failure
43%	Improper lubrication
26%	Improper mounting, or mounting technique
20%	Contamination, excessive external vibration, improper application
10%	Design bearing life wear out
1%	Brand new bearing, with metallurgical, or material, or quality control, defects

Sohoel (1994:2) inventor of the shock pulse method (SPM) in 1970, focuses on shock pulses emanating from the bearing. The main difference between conventional spectrum analysis and the shock pulse method is that the shock pulse method implements a transducer with a high sensitivity. The shock pulse method is implemented in the evaluated vibration analysis method, (EVAM), developed by SPM instruments. The EVAM method combines the SPM with statistical evaluation and pre-determined amplitude and statistical levels for rolling element bearing damage.

2.4 Rolling Element Bearings Condition Monitoring Techniques

Mathew and Alfredson (1984:447) conducted a brief review on techniques of condition monitoring of rolling element bearings. In the review, two approaches were considered to be applicable in industry. The first approach relies on the analysis of lubricating oils and greases. Temperature, thermographic, spectrographic, or a ferrographic analysis can be conducted to determine the condition of the lubricant. The second approach involves the measurement of bearing vibration and bearing noise. There are many techniques that can be applied such as acoustic emission, high frequency detection technique, shock pulse technique, ultrasonic technique and stress wave analysis. The vibration signal can be processed in the frequency domain or the time domain so as to highlight certain aspects of the whole vibration signal.

Hansford (2001:39) observed that vibration is usually one of the first parameters to show changes in a rolling element bearing, as well as indicating the severity of the damage and the potential cause.

The Vibration Institute of South Africa (2003:5) observed that, although all the techniques and methods have been proven to diagnose bearing faults in some form or another, there is no single method that is successful over the entire frequency range of bearing vibration. However due to past experience the institute has come to trust high frequency detection (HFD) spectrum analysis and the haystack effect.

The Technical Associates of Charlotte (1994:7) reported on case surveys conducted on various organisations using the high frequency detection technique. It was found from the survey that the HFD was not as effective as either shock pulse, ultrasonic flaw detection, nor spike energy in trending the earliest incipient damage. This was because of the constant magnitudes from one measurement to the other, reflecting no changes in incipient damage, even though incipient bearing damage was developing. Some users have reported that the HFD had not detected bearing defects until the second stage of degradation. This makes sense, since this is the same stage in which the HFD accelerometer responds to bearing vibrations.

Shiroishi *et al.* (1997:694) observed that HFD takes advantage of the large amplitude of a defect signal in the range of a high frequency bearing resonance and provides a demodulated signal with a high defect signal to noise ratio in the absence of low frequency mechanical noise.

Shiroishi *et al.* (1997:697) noted that the HFD demodulation involves three steps. Firstly, the measured signal is band passed around a selected high frequency band with the centre at a chosen resonant frequency of the bearing assembly. Secondly, the band passed signal is demodulated with a non linear rectifier.

Thirdly, a low pass filter is used to cancel high frequency components and retain the low frequency information associated with bearing defects.

Barkov *et al.* (1995b:1) showed that the HFD envelope or demodulation of the random vibration excited by friction forces and shock pulses in the bearing, is the most effective, accurate and least expensive method of rolling element bearing vibration condition monitoring. When bearing defects develop, the high frequency bearing vibration acquire amplitude modulation. Bearing defects can be detected by the analysis of a demodulated spectrum of this vibration. With this method, vibration condition monitoring can be accomplished with only 10 to 20 vibration measurements during the bearing's entire operational life. The demodulated signal of a damaged bearing obtained by the HFD is contaminated by broadband noise making it difficult to detect the earliest incipient damage of the bearing.

2.5 Human Versus Artificial Intelligence Condition Monitoring

Mathew and Alfredson (1984:450) recommend that the analysis of acquired rolling element bearing vibration measurements requires the skill of a vibration specialist to make final evaluations and suggestions of the results. However, it was observed that this may be a lengthy process which is prone to human error. Even with highly skilled personnel the margin of error is reduced but not overcome. A convinced vibration specialist on a bearing fault, would overhaul and replace the bearing. Proceeding with machine overhaul would result in loss of revenue and production time. Re-lubricating a damaged bearing results in a very short time of proper functioning and then malfunctioning of the machine is repeated.

Sharkels (2004:38) supports rolling element bearing vibration analysis by skilled personnel, and warns against relying on artificial intelligence alone in detecting bearing damage, especially the reliance on predetermined alert and alarm levels that indicate bearing damage.

However, he recommends a combined application of artificial intelligence and human evaluation for an informed decision. This observation was based on a case study on a paper machine bearing that had been diagnosed to have an inner ring crack, but predetermined alert and alarm levels had failed to indicate the damage. The crack was revealed when a visual inspection was conducted.

Van Wyk (1998:8) credits the recent developments of expert systems. These systems have intricate algorithms with powerful feature extraction and pattern recognition capabilities to self-evaluate the data. An expert system would provide information of the probable mode of failure of a bearing, the predicted time to failure and reasons for such predictions. These expert systems have the same limitations in that they are unable to classify data that show a lot of noise or unrecorded behaviour.

2.6 Signal Processing of Bearing Vibration Waveforms

Ramirez (1985:4) recorded that a great deal of scientific theory is written in terms of time histories. In 1582, Galileo observed the consistency of a pendulum, and Christian Huygens, in 1665, built the first pendulum clock using Galileo's observations. An electronic oscilloscope is in many ways analogous to a mechanical pendulum. The output of a sine wave oscillator has a time history that closely resembles the time history of a pendulum's angular displacement. In all time domain bearing vibration measurements an electronic pulse or a frequency counter is used as a time base. The time base is used to measure the period of the signal or to generate a time axis for recording an amplitude history.

Fourier analysis has been developed since the early 1800s, when Jean Baptiste Joseph Fourier developed the initial concepts and theory. In 1807 Fourier presented a paper on the use of sinusoids to represent temperature distributions. The paper made the controversial claim that any continuous periodic signal could be represented by the sum of properly chosen sinusoidal waves.

J.L Lagrange objected strongly to publication on the basis that Fourier's approach would not work with signals having discontinuous slopes, such as square waves. So the paper was not published until the death of Lagrange 15 years later. Lagrange was correct that a summation of sinusoids cannot exactly form a signal with a discontinuity. However, a very close approximation can be achieved if enough sinusoids are used. This is described by the Gibbs effect.

The Fast Fourier transform, the commonly used algorithm for determining the characteristics of a real vibration waveform on a frequency domain plane, was developed from the Fourier series and Fourier theory. In the 1960's J.W. Cooley and J.W. Tukey published an article on the algorithm for the machine calculation of complex Fourier series. This algorithm became known as the Fast Fourier Transform, (FFT), and has since become the new context for Fourier analysis. This algorithm allows for quick economical application of Fourier analysis and techniques to a variety of practical situations including filter design.

BAT Services (2004:1) observes that even when the most advanced manufacturing technology is used, noise still occurs naturally in rolling element bearings. As such, vibration and sound from a new bearing are accepted as normal bearing characteristics. Raceway noise is the most basic sound in rolling element bearings. It is generated in all bearings and is of a smooth and continuous amplitude.

BAT Services (2004:2) characterises raceway noise as follows:

- The frequency of bearing sound does not change even when rotational speed changes. The frequency of bearing sound is the natural frequency of non-damaged raceway rings.
- An increase in the running speed, causes an increase in sound magnitude.
- Reducing the radial clearance, causes an increase in sound magnitude.

- The higher the lubrication viscosity, the lower the sound magnitude.
- The higher the rigidity of the housing, the lower the magnitude of the sound.

2.7 Statistical Analysis on Rolling element Bearing vibrations

Mathew and Alfredson (1984:451) recommends the application of statistical analysis on bearing vibration. They conducted experiments on rolling element bearings. The bearings were subjected to unfavourable operational conditions. All bearings were allowed to run in for a reasonable period. This way the bearings developed defects according to the initiated defect.

Martin *et al.* (1994:76) confirmed that the statistical approach shows a great deal of promise in achieving tracking of bearing damage from the incipient stage. The use of statistical moments on the rectified raw data, especially the skewness, has been shown to be effective for tracking incipient damage. In using rectified data, both the odd and even moments can be used effectively, giving more flexibility for field operation. The method is also capable of characterising surface finishes in addition to surface damage, using the damage mapping approach.

Mathew and Alfredson (1984:452) analysed the time domain waveforms using the root mean square (RMS) statistical parameter. More information was obtained by measuring the ratio of the peak amplitude to RMS value which is the crest factor. The crest factor indicates deterioration of the bearing, the amplitude increases more rapidly than the RMS levels because of the increase in impulses. Skewness values near zero indicated that the data is symmetric about the mean. A kurtosis value of 3, indicates a good bearing. Mathew and Alfredson (1984:452) observed that an increase in incipient damage causes an increase in impulses on the time domain waveform giving kurtosis values of 6. Another observation made was that the sensitivity of the kurtosis value to good bearings seemed to lessen, due to dynamic impulses, when the testing shaft speed was increased. Crest factor and kurtosis patterns were similar.

Dyer and Stewart (1978:229) proposed kurtosis as a measurement parameter of the variation from the Gaussian distribution of a new bearing that would normally produce kurtosis values less than or equal to 3.

Mathew and Alfredson (1984:453) observed that damaged rolling element bearings produce impulsive vibration signals which would result in probability densities that were non-Gaussian. With the probability density normalised to its standard deviation, the shape of the distribution would have spikes indicating damage. A near Gaussian signal resembles a good bearing because of the evenness in amplitude, or lack of impulses in the time domain waveform. Various statistical moments can then be used to describe the shape of distribution. For example, the third moment, which is related to skewness, is used to identify lack of symmetry of the distribution.

Chen (2000:42) noted that the probability density distribution can be represented in the form of a histogram. It can be observed that a good bearing with random vibrations has a Gaussian distribution, while changes in the distribution curve, particularly at the lower values of the probability density distribution, indicate early stages of bearing failure. Changes at low probability have been found to be important in detection of bearing damage. The vertical axis is the density of the data, and horizontal axis is the acceleration of the vibration signal normalised to the standard deviation. A severely damaged bearing may also produce a near Gaussian distribution due to the evenness of many high amplitude impulses in the time domain waveform.

2.8 Induced Rolling Element Bearing Defects

Mathew and Alfredson (1984:449), in experimenting with rolling element bearings, induced bearing damage on seven bearings. The first bearing was contaminated with small amounts of solid particles. The outer raceway was finally worn. The second bearing had a small groove on the outer raceway. Severe damage subsequently occurred in the load zone of the outer raceway.

On the third bearing, small flat surfaces were ground on the rolling elements. The outer raceway in the load zone was subsequently damaged and so did the rolling elements. Bearings 4, 5 and 6 were operated under overload conditions. The cage became loose, the outer raceway was damaged in the load zone and the rolling elements were spalled. The seventh bearing had its lubrication drained out. All components of the bearing had been subjected to very high temperatures and bluish burn marks were evident on the rolling elements and the inner raceway. The outer raceway was damaged in the load zone. The high frequency detection technique was used on analysing the time and frequency domains of these induced damages.

Rubini *et al.* (2000:287) subjected a radial load of 500N on a self-aligning bearing which guarantees a bearing life of over 16 000 million cycles to failure. One bearing was damaged on the inner race, one on the outer race and the last one on a rolling ball. Starting from the undamaged condition, three different pit dimensions were artificially produced by an electric pen to simulate a gradual increase of the damage. A transverse line, approximately 1 mm wide, involving the race of one ball row was created on the raceway. In the case of the outer race, the bearing was mounted taking care to locate the damage at the point subjected to the highest load, where the probability of fault appearance was at a maximum.

Shiroishi *et al.* (1997:694) also conducted tests on rolling element bearings. The acoustic emission technique was used. The recorded defect reached a defect size of 0.0645 mm^2 which is commonly defined to be incipient damage by industry standards, and can be detectable by the acoustic emission technique. In the conducted tests all damage took the form of scratches made to the centre of the raceways with a diamond scribe. All cup and cone scratches were controlled to a length of approximately 2.54 mm. The width of the damage created was controlled by the number of passes made by a scribe made over the raceway. The defect widths examined ranged from $15.40 \mu\text{m}$ to $408.48 \mu\text{m}$.

The defect sizes were such that they were well below industry standards for the definition of incipient failure as detectable by the acoustic emission technique. The acoustic emission technique is more sensitive than the high frequency detection technique.

Al-Gahmd *et al.* (2004:5) also applied the acoustic emission technique. The test bearing employed was a Cooper split-type roller bearing. The split-type bearing was selected as it allowed defects to be seeded onto the raceways, furthermore, assembly and disassembly of the bearing was accomplished with minimum disruption to the test sequence. Five test conditions of varying severities were simulated with point defects engraved on the outer raceway of the test bearing.

2.9 Historical Development of Kohonen Neural Networks

Bermington (2001:194) reported that in the early 1900s, electronic devices began to replace manually operated mechanical machines. In later years, the inventions of the transistor and the silicon chip in the 1950s, allowed the computer to develop. Binary code is the language that all computers use in performing calculations. Through binary coding the programs that carry the functions and operations of neural networks is understood by the computer.

McCulloch and Pitts (1943:117) published the first papers on neural networking. A significant event was recorded in 1956 at Dartmouth College that modelled the basics of artificial intelligence (AI). The Rockefeller Foundation sponsored a conference on the topic of potential use of computers and simulation in every aspect of learning and any other feature of intelligence. It was at this conference that the term artificial intelligence was formally defined. Artificial intelligence was defined as: computer processes that attempt to emulate the human thought processes and are associated with activities that require the use of intelligence.

This definition includes the fields of automatic learning, understanding natural language, vision-image recognition, voice recognition, game playing, mathematical problem solving, robotics and expert systems. In recent years researchers have accepted neural networks as a legitimate field of artificial intelligence. The reason for the slow acceptance of this artificial intelligence is the similarity of neural networks to biological entities.

Rosenblatt (1958:386) initiated the development of the first perceptron neural network along with the training algorithm. A perceptron is the most basic structure of a neural network. A single neuron or perceptron consist of a transfer function with a number of inputs, with each input having its own weight. The perceptron training algorithm allows for the weights to be adjusted so that the perceptron produces a target output. The target output, is the reason the perceptron is grouped into supervised neural networks. A single neuron can model a linear relationship. Applying neurons in parallel results in a neural network that can model non-linear relationships.

In the following publications by Rosenblatt (1958:386), numerous upgrades were introduced for faster and better learning and training. Upgrades such as the Steinbach's learning rule and Widrow-Hoff learning rule brought rapid development in supervised neural networks. As a result Widrow and Lehr (1990:1415) recorded that Werbos developed the back propagation learning rule in 1971.

The earliest unsupervised neural network learning rule was the Hebbian learning rule discovered by D. Hebb in 1949, a neurophysiologist who found that 'if one neuron repeatedly excites another neuron, the amount of excitation is lowered'. Hebbian learning extract and copies the pattern of the input data vector space, so that the output is similar to the input.

Real developments in unsupervised neural networks began in the late 1950s and early 1960s with Stratonovitch (1957:416) studied vector quantisation and suggested that data vectors can be grouped with relating vectors by a cluster vector. Glenn (1996:1) recorded studies on vector quantisation conducted in 1964. Von der Malsburg (1973:85) developed the first simulation of self-organising. In 1976 Von derMalsburg and Willshaw (1976:121) suggested the idea of a self-organising map (SOM). In the 1980s refinement of the structure and computational algorithms of the self-organising map were developed by Kohonen (1982a:1).

2.10 Biological Analogy to Artificial Intelligence Neural Networks

Haykin (1994:4) provided a good comparison of biological neurons and artificial intelligence neural networks. The development of neural networks is based on neurobiological modelling. The processing elements that make up the neural network are named neurons, after the term for biological cells that make up the nervous system.

Haykin (1994:6) elaborated that biological neurons and artificial intelligence neurons are stimulated by an input, the neurons then process and transmit the data either to other neurons or produce an output. In biological terms the output would be for example contraction of the muscles to produce body motion. The input in artificial intelligence neural networks is transmitted to the neuron through connection weights, which in biological terms would be the synapse. The input is admitted into the neuron and processed by the transfer function, this would be the cell nucleus in biology. The output of the transfer function is transmitted to another neuron or is itself the expected output. The input progresses through repeated calculations until the output is reached. The transmission path of the stimulation in a biological neuron is through the axon, and then spreads to other neurons through the dendrites, this would resemble the repeated steps of calculations to reach an output. Many neurons can work together in parallel to process the same input, in a similar fashion to biological neurons that work in parallel.

Haykin (1994:20) and Uhrig (1996:1) noted that the accuracy of a neural network output, is increased by the number of neurons working in parallel. Factors affecting this include the strength of the weight connections and the strength of an additional weight connection known as the bias connection. Similarly, the strength of the response from a biological neuron also depends on the number of neurons transmitting the stimuli amongst and on other factors such as the synapse strength and the amount of iron content in the synapse which resembles the bias factor.

2.11 Development of the Self-Organising Map Algorithm

2.11.1 Competitive Algorithm

The idea of competitive learning may be traced back to the early works of von der Malsburg (1973:90) on the self-organisation of orientation sensitive nerve cells in the striate cortex of the brain; Fukushima (1975:121) on a self-organising multiplayer neural network named the cognitron; Willshaw and von der Malsburg (1976:431) on the formation of patterned neural connections by self-organisation; and Grossberg (1972:49, 1976a:121, 1976b:187) on adaptive pattern classification. Durbin et al. (1989:133) highlight the importance of competitive learning in the formation of a topographic map, and the work by Ambros-Ingerson et al. (1990:1344) provides further neurobiological justification for competitive learning.

Grossberg (1976a:187) defined hard competition, and soft competition. The competition in unsupervised neural networks allows for exploring the input data without the need for controlling the output. With hard competition, only the closest neuron to an input data is active and thus has its weights updated. With soft competition the closest neuron is active, and its neighbours are also active, although somewhat less than the winning neuron. This allows for a topological mapping of neighbouring neurons to the input data space.

Grossberg (1976a:8) noted that the major difference between competitive learning and Hebbian learning is that Hebbian learning tends to extract information from the input space, whereas competitive networks are used to cluster similar inputs. Hebbian networks are neuron associators to the input data, while competitive networks compete for input data resources. Grossberg introduced most of the ideas of competition in mathematical terms. Kohonen took a more engineering oriented approach to these ideas and introduced an enhanced set of principles that are easily implemented in digital systems. Kohonen's approach is used in this research.

2.11.2 Data Pre-processing and Feature Extraction

Martin-del-Brio and Serrano-Cinca (1993:193) referred to data pre-processing as the operations performed on data before information can be extracted from the data. Data pre-processing includes normalisation, linear scaling, and linear regression. As far as normalisation is concerned, input and output vectors should be normalised to unity if correlation is used to determine convergence, unless Euclidean distance or squared error is used, where upon normalisation need not be to unity. For a data set that is normally distributed the data can be normalised so that the mean becomes zero and the standard deviation becomes unity.

Kohonen (1982b:5) notes that data processing includes data pre-processing and feature extraction. Feature extraction is referred to as the process of determining characteristics of data that can easily be modelled. Mao and Jain (1995:297) wrote that if a mathematical model can duplicate a data set, indicating a certain condition, and the same model can duplicate an unknown data set, then the assumption that the same condition exists for both data sets is valid. Feature extraction may also be viewed as a data projection by means of mapping an input data space to an output space. The result is a reduction in dimensionality, enabling easier access to large data sets, as well as separating different data state conditions.

De Backer *et al.* (1998:711) studied feature extraction and four feature classification of high dimensional input data sets. Large feature vectors that are able to describe complex objects and to distinguish between them are generated. The four classification methods studied were: multi dimensional scaling approach, Sammon's mapping, Kohonen's Self-Organising Map, and an auto associative feed-forward neural network. All four neural networks yield better classification results than the optimal linear approach of the back propagation neural network, and therefore can be utilised in feature classification schemes. The findings were that the multi dimensional scaling and Sammon's mapping algorithms, are most suited for high dimensional data sets with a limited number of data points, while self-organising map and the auto-associative feed-forward neural network are more appropriate for low dimensional problems with a large number of data points.

2.11.3 Output Topology and Classification

Topology is the lattice structure of the output layer. The neurons in the output layer are arranged in a one or two dimensional lattice structure producing a topology that ensures that each neuron has a set of neighbouring neurons. There are two topological models inspired by von der Malsburg (1973:85). The first model, proposed by Willshaw and von der Malsburg (1976:431), comprises two separate two-dimensional lattices of neurons connected together, with one projecting on to the other and a group of neurons win or dominate the map at any one time. The Willshaw-von der Malsburg model is specialised to mappings where the input dimension is the same as the output dimension. The second model was introduced by Kohonen (1982a:59), and comprises a single two-dimensional lattice of neurons. The Kohonen model is more general than the Willshaw-von der Malsburg model because of the ability to perform data compression as there is one winning neuron, resulting in a greatly reduced output dimension.

The Kohonen model may be derived in two ways, one is the consideration of traditional neurobiology and the second approach is vector quantisation. The Kohonen model is used in this research for its data compression capabilities.

Kohonen (1982b:45) used statistical data features to train a classifier for classifying a data set within non linear boundaries. Classification involves grouping input data vectors into clusters and presenting the cluster according to condition cluster label.

MacQueen (1967:281) adopted the k-means clustering algorithm as a vector quantisation method that makes use of the competitive transfer function. The algorithm minimises the sum of squared errors among a number of k clusters.

Bouldin (1979:224) evaluated vector quantisation and cluster algorithms. The k-means algorithm divides out an input data set into, k , number of classes. The process of creating the first data clusters is similar to that of Self-Organising Maps. However, as the number of classes to be recognised is initially not known, the process of cluster creation must be repeated for different values of, k . The value of k classes is then estimated by means of a validity criterion such as the Davies–Bouldin index. The Davies–Bouldin index is used as a measure of the validity of the number of clusters obtained with the k-means method.

2.11.4 Learning Vector Quantisation Neural Network

Kangas *et al.* (1990:93) observed that the best results for pattern classification are achieved by the use of an unsupervised Kohonen self-organising map together with a supervised learning algorithm or neural network.

Kangas *et al.* (1990:99) recommends the learning vector quantisation network, which is a supervised network in that it has a non-linear competitive layer, and a linear output layer transfer function.

This linear transfer function enables the output to be placed in groups or classes. A learning vector quantisation network performs similar functions to that of the self-organising feature map network. The learning vector quantisation network however creates classes or groups on the topological map instead of only classifying the input on the topological map. The learning vector quantisation network can be used to analyse the input data that is analysed by the self-organising feature map network, and thus enables support evaluation.

2.12 Applied Neural Networks on Vibration Monitoring Systems

2.12.1 Rolling Element Bearing Condition Monitoring

Shao and Nezu (1995:1543) used the factor of the degree of creditability and the basic principle of the expert system on an online monitoring and diagnostic method of rolling element bearings with artificial intelligence, namely fuzzy logic. The technique enhanced traditional vibration analysis and provided a means of automating the monitoring and diagnosis of vibrating devices. It was observed that identifying incipient damage was a problem with traditional vibration analysis due to the sensitivity of parameters. The parameters assist in determining alarm limits automatically, however such parameter values are affected by the load variations, speed variations, contamination, and measurement error. The uncertainty of the parameter value, for good and damaged bearings, is the reason Shao and Nezu used the factor of the degree of creditability. New bearings were used for experimentation, and no distinct fault classification was performed. Therefore, visualisation and diagnosis of bearing damage can not be achieved by this method.

Crupi *et al.* (2004:1137) noted a disadvantage of neural networks in that data acquired from real fault conditions is required to train the network. Also noted is that neural networks have replaced the traditional expert systems, used in the past for fault diagnosis, because neural networks are dynamic systems adaptable to continuous variable data.

Toshiya *et al.* (1995:1) used the feed-forward back propagation neural network to monitor the conditions of journal bearings. The feed-forward back propagation neural network is limited to only two numerically classified conditions, because of the tan sigmoid transfer function, and the bearing is classified as either in a ‘good’ or a ‘bad’ condition.

Li *et al.* (2000:1060) concentrated specifically on rolling element bearing faults, and used a machine fault simulator apparatus to simulate bearing faults. The bearing fault vibration was generated by replacing the front bearing of the apparatus with a bearing of known fault condition provided by the bearing manufacturer. Severe faults were simulated on the ball element, inner raceway, and outer raceway. Vibration measurements were measured in the radial and axial directions. Statistical parameters such as: peak amplitude, mean amplitude, kurtosis, amplitudes at the shaft frequency, and amplitudes at the ball pass frequency of the outer and inner ring, were the extracted features that created the input dataset. A three layer feed-forward neural network was trained using the Levenberg-Marquardt algorithm. A decisive ‘good’ or ‘bad’ two conditions numerical classification of bearing condition was achieved, because a tan sigmoid transfer function was used. Bearing defect visualisation was not achieved and defect severity could not be analysed.

Samanta *et al.* (2001:327) extracted statistical feature parameters from the wavelet transform in condition monitoring of rolling element bearings. The use of the wavelet transform is not within the scope of this research, but could be considered for further study. Samanta used a feed-forward neural network. The results of this experimentation, showed that wavelet feature extraction did not improve the results of bearing fault detection significantly. The training was quite fast requiring significantly small number of epochs. This substantial reduction in training epochs was due to pre-processing of vibration data and using a substantially smaller number of extracted features as inputs. Although the number of extracted features was smaller, they were still effective when extracted from raw signals.

Alguindigue *et al.* (1993:209) implemented a methodology for interpreting rolling element bearing vibration measurements based on the combination of the back propagation feed-forward neural network and the re-circulation neural network. This methodology is also an automated monitoring process. A new bearing was used to obtain bearing vibration measurements. The re-circulation neural network classifies bearing faults with six output neurons. A fault excites one of the six neurons, and is numerically classified. The limiting factor is that the faults might not be visualised on map classification, and a definite fault is given instead of the severity of the fault.

Chen (2000:64) researched condition monitoring of rolling element bearings. A three-layered artificial neural network was introduced to accomplish the non-linear mapping from the feature space to the classification space. One advantage of the artificial neural network approach is that a complicated non-linear relationship can be constructed between data spaces. Defect detection was on tapered roller bearings used in rail traction. To train the neural network for non-linear mapping, a multi defect data set prepared from operating the bearings under various conditions such as different loads and speeds, was used. Severity of the defects was also reflected by single and multiple spalls. Natural fatigue spalls instead of artificially induced spalls were considered.

The back propagation feed-forward neural network was combined with a two-dimensional mapping rule that allows the application of piecewise linear boundaries. The cluster arrangement was somewhat arbitrarily placed evenly on a unit circle in the quadrant of a two-dimensional Cartesian plane. A unit circle was chosen so that the outputs would fit into the range of a sigmoid function, that is, between 0 and 1. The desired cluster centre coordinates were chosen on the Cartesian plane. A distribution type classification method was used. Visualisation of bearing defects was achieved. However, analysis of the clusters was limited because the cluster distribution could be controlled. Therefore, only bearing defects could be visualised, and not vibration parameters.

2.12.2 Machining Tool Condition Monitoring

Scheffer *et al.* (2001:1185) implemented a tool condition monitoring strategy based on vibration measurement. The techniques considered were: time and frequency domain techniques, statistical feature extraction, and wavelet technique. Statistical parameters extracted were used to create a neural network input dataset. The Self-Organising Map could achieve progressive classification of tool wear considering all statistical parameters. The condition of the cutting tool was not limited to a two conditions numerical classification.

Godin *et al.* (2004:299) statistically analysed acoustic emission signatures of various types of damage in composite materials during tensile tests on cross-ply laminate composites. A Kohonen's self-organising map was used as an unsupervised pattern recognition method suitable for the discrimination of acoustic emission signals, and clusters were identified by application of the k-means method.

2.12.3 Machinery Condition Monitoring

Kowalski *et al.* (2003:1) developed a monitoring method for induction motors in the case of rotor, stator, and rolling element bearing faults. A feed-forward multi-layer perceptron neural network, was compared to the Self-Organising Map neural network. Bearing failure was shown to be the major contributor of motor failures. Kohonen neural network were found to be more efficient for initial classification of motor faults. The multi-layer perceptron is more applicable in severe motor faults and a two condition classification.

Van der Merwe *et al.* (2003:139) conducted experiments on machine imbalance and rolling element bearing outer ring fault detection. Kohonen self-organising maps were used and compared to the radial base function neural network and the feed-forward multi-layer perceptron. The radial base function network was better than the multi-layer perceptron, on a dataset that had well-defined clusters.

2.13 Summary

An introduction in which condition monitoring of rolling element bearings were described, with the focus on the application of Self-Organising Maps to automatic fault detection and classification. The purpose and application of condition monitoring in the maintenance of rolling element bearings was discussed. An overall review of previous studies and developments in the application of neural networks to condition monitoring of rolling element bearings was presented. The techniques used for the purpose of this research were introduced through a discussion of the related literature, as in the case of bearing damage infliction and signal processing. The literature on induced bearing damage highlighted different severities of incipient damage detectable by the high frequency technique.

The literature research on Self-Organising Maps for condition monitoring purposes, revealed the applicability of the technique and is the reason for the decision to develop a similar monitoring system for rolling element bearings. Feature extraction techniques like statistical moments, spectral and amplitude parameters of the raw data can be used as features, serving as input to the Self-Organising Map.

The literature review reflects the dynamic nature of neural networks, and that the self-organising map is an excellent tool in the visualisation of high dimensional data. As such, it is most suitable for the data understanding phase of the analysis process. However, Self-Organising Maps can also be used for data preparation, modelling and classification.

Chapter 3

Bearing Vibration Fundamentals

3.1 Introduction

This chapter provides a description of the deep-groove ball bearing on which vibration condition monitoring experiments were done in the research. More specifically, the characteristics and properties of the rolling element bearing that influence bearing vibration are discussed. The design of the monitoring rig is presented with reference to how the bearing is accommodated in the rig, as well as the operation mechanism of the rig. Characteristics of the operations of a bearing include bearing friction, bearing lubrication, bearing loading, bearing speed, and bearing life. Barkov *et. al.* (1995b:1) highlighted that bearing friction is the main cause of wear and ultimately spalling in a bearing. This is the reason bearing friction receives so much attention in discussions of the operations of a bearing.

Different types of rolling element bearings are noted and the reasons for researching the deep-groove ball bearing are mentioned. The nomenclature and vibration characteristics of the rolling element bearing components are explained so as to assist in locating bearing defects. The characteristics of complex phenomena of bearings such as bearing waviness and the fluctuation of bearing stiffness are also mentioned so that limits to the scope of the research are established. The discussion on the design of the monitoring rig is accompanied by illustrations of the structure of the rig. The concept of machine vibration in relation to bearing vibration, is discussed with specific attention to rolling element bearing vibration.

3.2 Types of Rolling Element Bearings

Harris (2001:4) wrote that the term *rolling element bearings* includes all forms of bearings that utilize the rolling action of ball or roller elements to attain minimum friction between one body relative to another. Today, technical and environmental requirements for bearings are very complex. In certain conditions, ball bearings are more applicable than roller bearings. The selection of a bearing requires the expertise of skilled personnel, and the selection is supported by mathematical calculations.

There are different classes of ball bearings, and these are classified according to general construction, rather than being grouped in terms of shields, seals, and snap rings. The ball bearings can be classified into nine main groups. These groups are: deep-groove, maximum capacity, angular contact, self-aligning, double-row, Y-bearings, cam rollers, thrust, and thrust angular contact.

SKF (2004:31) lists the advantages of deep-groove ball bearings that have led to their widespread use as: (1) lower friction torque than roller bearings, (2) higher grease life than roller bearings, (3) reduced sensitivity to misalignment, and (4) lower production costs per volume of production achieved. The research is focussed on bearings that are most commonly used and have wide-spread applications as is the case with deep-groove ball bearings.

3.3 Deep-Groove Ball Bearing Nomenclature

The nomenclature of a modern type of ball bearing is shown in figure 3.1. This bearing is named a deep-groove ball bearing because of its characteristic deep groove raceway on the outer surface of the inner ring and on the inner surface of the outer ring. A deep-groove ball bearing consists of an inner ring, an outer ring, ball elements, a cage or separator, and seal caps if the bearing is to be sealed. The seals may not be required, where upon the seal grooves on the two rings are not necessary and this

may be changed according to specific design requirements. Another additional feature is the lubrication groove on the outer surface of the outer ring, with an inlet hole for the lubrication to pass into the rolling element area. This may be required for lubricating bearings that are difficult to reach, and may not be a necessary feature of a bearing because some bearings are greased only once by the manufacturer and then sealed. Additional features can be changed according to design requirements.

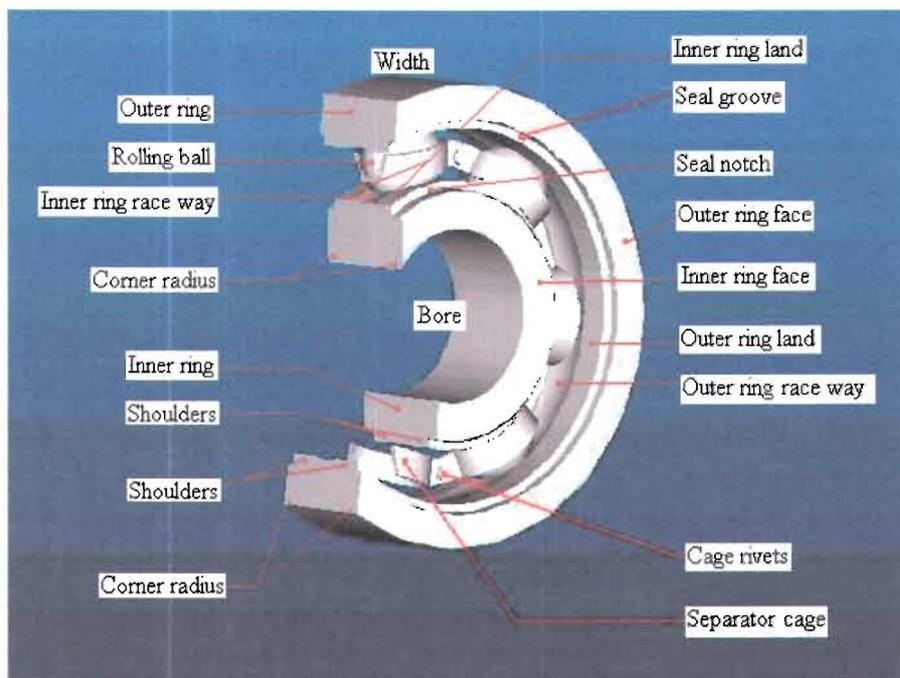


Figure 3.1: Deep-Groove Ball Bearing Nomenclature

The ball cage or ball separator has two halves which are used to space the ball elements at equal distance from each other and thus prevent one ball from rubbing against another. The cage follows the shape of the ball element so that there is smooth rolling between the ball element and the cage. At half the distance between two ball elements, the cage is pin riveted thus bringing the two cage halves together. The cage is placed halfway between the space separating the inner ring and the outer ring and the pitch circle diameter lies along the cage. The cage also helps to absorb dynamic forces that may occur when the ball elements roll in and out of the bearing load zone.

The deep grooves on the inner race and outer race are accurately ground to eliminate sliding friction between the balls and the grooves, while simultaneously providing proper support for the balls. The rolling balls are similarly ground like the raceways for the same purpose of eliminating sliding friction. The deep-groove ball bearing in figure 3.1 has a single row of rolling elements. In other designs there may be two or more rows to increase the support load. The bore has rounded edges to easily accommodate the shaft diameter and the outer ring is similarly rounded to reduce stress concentration when mounted in the bearing housing.

3.4 Design and Manufacture of rolling element bearings

The design and manufacturing processes of bearings and how these processes affect the vibration characteristics of the bearing are discussed below.

3.4.1 Rolling Element Bearing Design

SKF (1996:8) consider the following factors when designing rolling element bearings: load zone, contact areas, contact stress, elastic deformation, and rolling and sliding friction. While rolling friction may approach zero, a small amount of sliding friction will exist in all rolling element bearings. Although bearing components are made of hardened steel, they are elastic bodies under load and this property of the bearing material is partly responsible for the vibration response of a bearing, because elasticity is essentially the stiffness property of a bearing.

3.4.2 Rolling Element Bearing Materials

Avallone *et al.* (1989:739) recorded that for many years the most important bearing material was Babbitt metal, patented by Isaac Babbitt in 1839. Babbitt metal, an alloy of tin with small amounts of antimony, copper and lead, has a number of varieties, depending on the proportions of the constituents. Other metallic materials are a combination of Babbitt metal and steel. The type of material used in a bearing affects the stiffness properties of the bearing and the natural frequency of the bearing.

3.4.3 Rolling Element Bearing Manufacturing

SKF (1996:20) implements manufacturing of rolling element bearings in four basic process stages:- first form, heat treat, final form and assembly. Rings are formed from solid bars by numeric controlled machines. Balls and rollers are turned from bar stock. Taylor *et al.*(2004:68) warns that the grinding process, to remove peaks of surface roughness that initiate spalling, may result in magnetising the bearing, which attracts metal particles that can initiate incipient spalling.

3.4.4 Rolling Element Bearing Preload

Bearings are designed with a certain amount of internal clearance between the component elements. Once the bearing has an applied load, the internal clearances reduce. To minimise the internal clearance without depending on the applied load, a rolling element bearing can be preloaded. A preload is applied when the bearing is first installed and then maintained under a constant load. Preloading a rolling element bearing influences the stiffness and damping properties of the bearing which influences the amplitude of bearing vibration. Therefore amplitude alone cannot be relied upon to indicate defect severity.

3.4.5 Rolling Element Bearing Stiffness

The bearing stiffness influences the frequency response of the bearing. This frequency is higher than the frequencies of the other components of the bearing, such as contacting bearing rings and rolling elements. On a time domain waveform there is a component sinusoid of the waveform that has the same frequency of the effective bearing frequency response.

3.4.6 Rolling Element Bearing Waviness

Bearing waviness refers to the surface contours on the raceways of the outer ring and the inner ring. Contours with long wavelengths are produced during machining. Contours with short wavelengths result from preload and loading conditions as the rolling elements apply pressures at different points in the raceways.

The relationship between the surface geometry and bearing vibration level is complex and is not within the scope of this research. What is of concern here is the fact that bearing noise of a new bearing is mainly due to bearing waviness among other factors such as lubrication, surface roughness or finish, load and speed.

3.5 Operating Principles of Rolling Element Bearing Friction

Ruiyun (1984:1) studied friction, and proposed a theory explaining the causes of friction. The theory is based on the fact that the surfaces of the pieces of solid in contact are not completely smooth and when they move relative to each other uneven surface peaks will prevent the bodies from moving thus producing resistance to motion. The resistance to relative motion in rolling element bearings is due to many factors, the basic one being rolling friction. This was long assumed to be the only resistance to motion in this type of bearing, however it has been established that rolling friction is a small amount of the whole bearing friction. Nevertheless, the contribution of rolling friction to wear and tear and operating temperature is important.

The major motion of a rolling element bearing is concerned with rotation since it supports a rotating shaft. The presence of friction between the rolling elements and the inner ring allow the rolling elements to rotate in the opposite direction to the rotation of the inner ring and the shaft, and thus reduce sliding friction. The unpredictable rotational path or direction of the rolling element is the reason why there are different types of frictional force in rolling element bearings. In bearings, this frictional force is termed a friction torque, because a torque or a couple sets the shaft into a rotational motion.

There is friction torque due to:

- differential sliding while the rolling element rotates.
- elastic hysteresis depending on the elastic properties of the rolling element and

ring materials and the amount of deformation that takes place under preload and loading conditions.

- geometric errors, for example a geometric error in the shape of the outer race may result in inconsistent surface contacts.
- shearing of the lubricant, which depends on the viscosity of the lubricant.
- temperature changes which result in lubricant viscosity changes.

3.6 Deep-Groove Ball Bearing Vibration Characteristics

Bearings first begin to wear when incipient damage starts, causing the natural frequencies of bearing components to become excited. One set of their natural frequencies is concentrated within the range up to 2000 Hz. Another set is found within ultrasonic frequency range between 20 KHz and 100 KHz. Bearing vibrations are induced when the natural frequency of the bearing and its components are excited, as a rolling element rolls over an incipient damage or due to external mechanical input such as misalignment.

Frequencies associated with bearing components include: the characteristic defect frequencies, and the frequencies of the outer race, inner race, ball elements and cage. The frequency at which the ball element rolls over a defect, is a characteristic bearing frequency.

Bearing characteristic frequencies can be calculated based on bearing rotational speed and bearing geometry. Figure 3.2 shows the geometry of a rolling element bearing with its relation to the velocities of the inner race, outer race, ball and cage as defined by Taylor *et al.* (2004:20). The meaning of the symbols used in figures 3.2 and figure 3.3 are shown in table 3.1.

Figure 3.3 shows the side view of a rolling element bearing illustrating the change of contact angle between the rolling element and the raceway. The change of

contact angle affects the length of the inner radius and length of the outer radius with reference to the position of the rolling element. The symbols in the figures are applied in the derivation of bearing defect frequency equations. The derivations of the equations for determining characteristic bearing frequencies are shown in annexure 1 (p132) , and the formulas are tabulated in annexure 2 (p137).

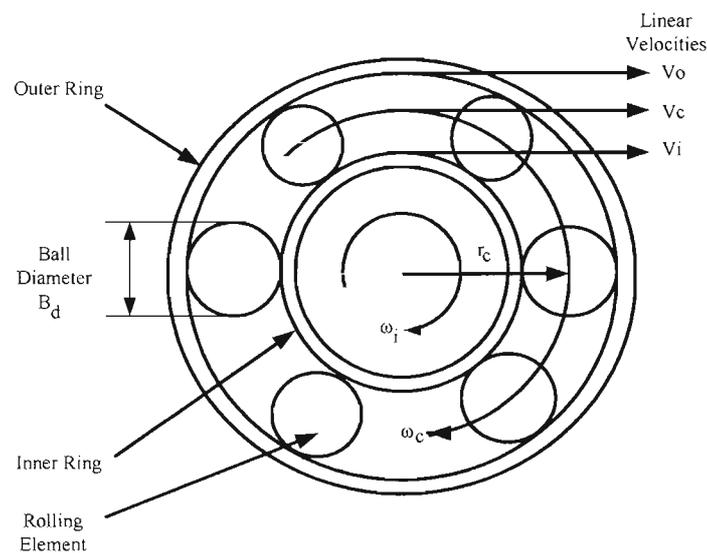


Figure 3.2: Front View of a Deep-Groove Ball Bearing

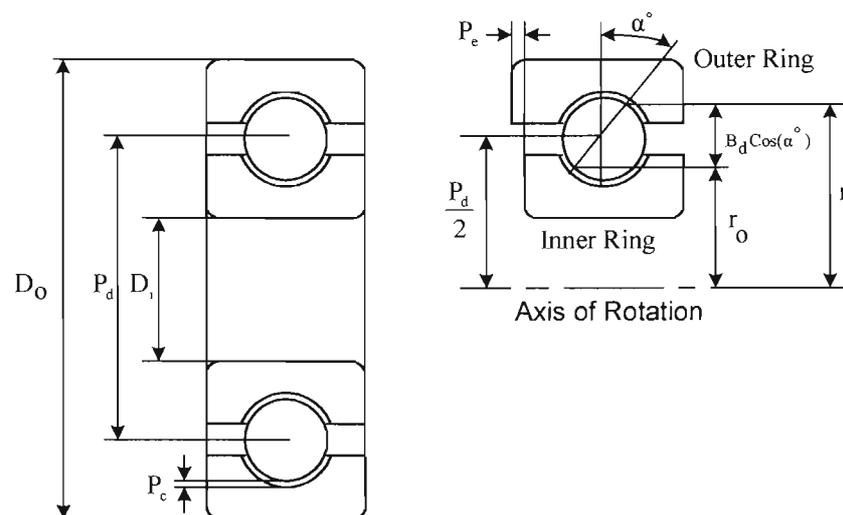


Figure 3.3: Side View of a Deep-Groove Ball Bearing

Table 3.1: Symbols for Figure 3.2 and Figure 3.3

V_c = linear speed of the cage	ω_c = rotational speed of the cage
V_i = linear speed of the inner race	ω_i = rotational speed of the inner race
V_o = linear speed of the outer race	ω_o = rotational speed of the outer race
V_b = linear speed of the ball	ω_b = rotational speed of the ball
r_c = rotational radius of the cage	P_d = Pitch circle diameter
r_i = rotational radius of the inner race	B_d = Ball diameter
r_o = rotational radius of the outer race	α = contact angle
r_r = rotational radius of the balls	Z = number of balls P_e = end play

3.7 Rolling Element Bearing and Machine Vibrations

For a well balanced and aligned machine with an old damaged rolling element bearing, most of the vibration energy is generated by the shock force and frictional force emanating from the bearing. Under such circumstances, the machine must be overhauled and the bearing replaced, to prevent consequent detrimental effects on the machine such as the generation of excessive noise and malfunction of resulting unaligned components. Since general rolling element bearing vibrations account for a very small part of the whole machine's vibration energy, special care must be taken when monitoring bearings to ensure that measurements of the vibration are those generated from the bearing and reflect the vibration conditions of only the bearing.

A rolling element bearing generates shock forces throughout the bearing's life. Shock forces create a transient pressure wave that is transmitted and absorbed by the bearing material and adjacent bodies and the surrounding air. SPM instruments (2000:12) confirm that the intensity of the shock pressure wave is directly related to the thickness of the oil film between the rolling element and raceway, and to the mechanical condition of the bearing surface.

An initial shock wave generated on a bearing will cause a pressure wave to be transmitted through the bearing and machine body. The beginning of a second shock wave during the first wave transmission, initiates induced bearing vibration. For a bearing in rotational motion, the repetition of the first shock wave starting and its combination with the second wave, creates a continuous induced bearing vibration which can be measured. The induced bearing vibration excites the natural frequency response of the bearing and its components. Induced bearing vibration and bearing natural frequencies are dangerous if they result in high amplitudes. Induced bearing vibration is mainly responsible for the progression of incipient damage into severe damage in a bearing, assuming that external sources of vibration such as imbalance are at a minimum.

3.8 Design of the Condition Monitoring Rig

The first problem was to design and build the test rig. The rig was used for defect testing on rolling element bearings, and was located in the Iscor Innovation Centre at the Vaal University of Technology. The rig was used to conduct experiments in acquiring rolling element bearing vibration measurements. Safety was a priority in the design and operation of the condition monitoring rig, so the South African occupational health and safety regulations acts of 2003 for rotary machines were followed.

3.8.1 Structural Support Framework

In order for the rig to be ergonomically designed for quick exchange of tested bearings, the base-plate was elevated to near human waist height of 740mm. The support structure had a rectangular shape with a top view length of 790mm and width of 685mm. The structural support frame work is illustrated as part of the rig in figure 3.4. The support structure stood on four elastomeric mounts which had adjustable nuts, for setting a horizontal level of the base-plate. The elastomeric mounts dampen the vibrations transmitted from the test rig to the floor.

3.8.2 Variable Speed Drive

A 2.2KW, single phase, capacitor start, capacitor run electric motor was used for the test rig. The electric motor received power from a frequency inverter, which in turn was connected to the 230V, 50Hz line voltage. The frequency inverter enabled smooth speed variation of the electric motor unlike a manual gear system. A variable speed drive system was used for the electric motor and the frequency inverter. The frequency inverter used was a CFW-08 series that ran on version 3.9X0899.4690 E/5 digital software.

3.8.3 Test Bearing Mounting Mechanism

The electric motor drives a test bearing shaft through a pulley and V-belt system. The test bearing shaft has a diameter of 30mm and length 490mm with respect to a safety factor of, $n = 2$, which is conservative to minor impact loading. The shaft pulley and electric motor pulley both have an outside diameter of 105mm. Equal pulley sizes enable equal speed transmission from the electric motor to the test bearing shaft. The speed is set by the frequency inverter. The test bearing shaft is supported by two Y-bearings.

The shaft is elevated on a rigid shaft-base, giving an allowance for adequate mounting and dismounting of the tested bearing. The shaft pulley and the tested bearing are located at the ends of the shaft, with the pulley on the one end and the tested bearing on the other, thus creating overhangs on both ends of the shaft. The test bearing shaft and V-Belt drive system are illustrated in figure 3.5. The tested bearing is however mounted on a stepped chuck and it rests at the back of the step pushing against the chuck. Figure 3.6 shows the tested bearing under a bearing seat and mounted on a stepped chuck mounting mechanism. This method of mounting, greatly reduces misalignment of the bearing on the chuck. A tapered mounting washer strengthens the mounting and alignment when the taper pushes on the bearing and is locked in place by lock nuts at the end of the shaft.

3.8.4 Loading Mechanism

The loading mechanism is mounted on the base-plate, and aligned at right angles to the tested bearing. The mechanism is assembled on two parallel round bars. The whole loading assembly is able to slide over the round bars. The round bars are 25.4mm in diameter and 760mm in length. There are two locking spring bushes one on each bar. The springs enable smooth loading of the bearing. A hydraulic jack is attached to the upper arm with the cylinder extending towards the base-plate. A load cell is attached on the clevis of the cylinder and on the lower arm. The hydraulic cylinder and the load cell are located along the centre between the two round bars. The lower arm also holds the bearing seat. The bearing seat is manufactured in a way that resembles a bearing housing. The bearing seat covers the top half of the bearing so that the load zone is within the bearing seat. Figure 3.7 illustrates the test bearing loading mechanism with the load cell and attached accelerometer.

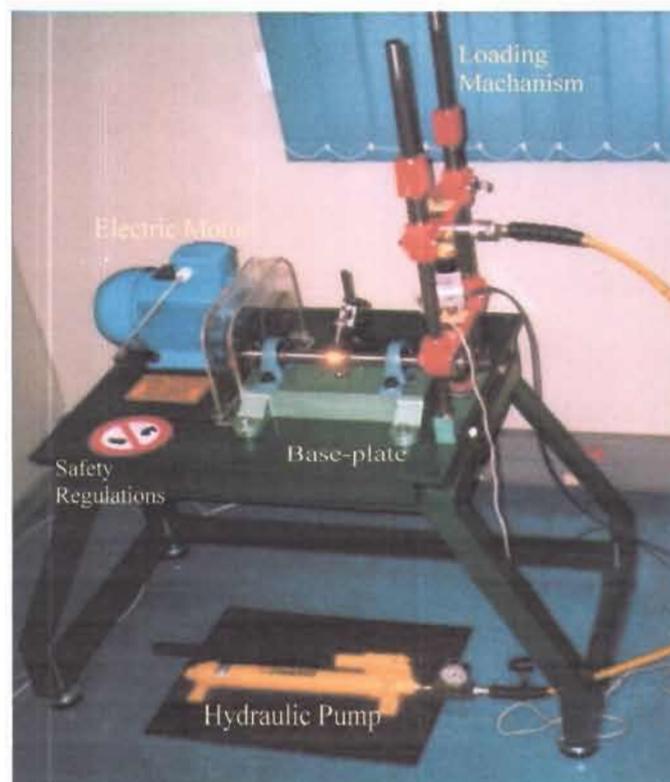


Figure 3.4: Rolling Element Bearing Condition Monitoring Rig

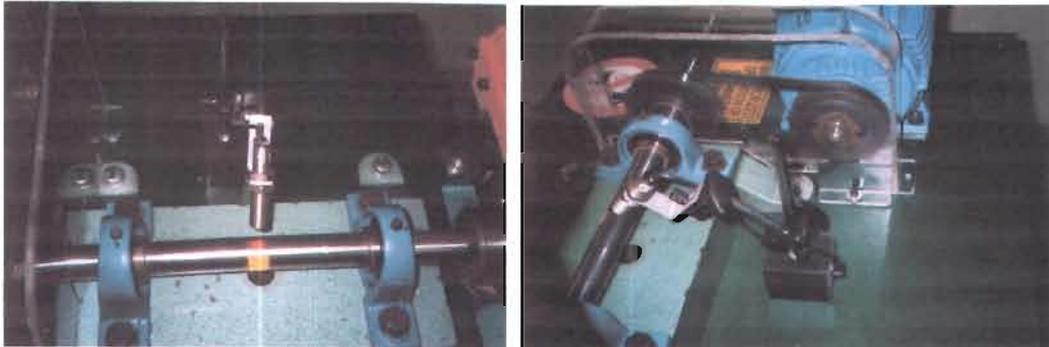


Figure 3.5: Test Bearing Shaft and V-Belt Drive System

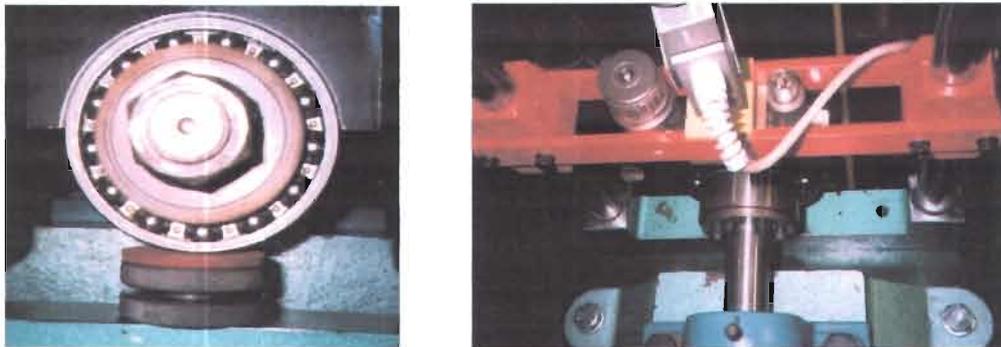


Figure 3.6: Bearing under a Bearing Seat and Stepped Chuck Mounting Mechanism

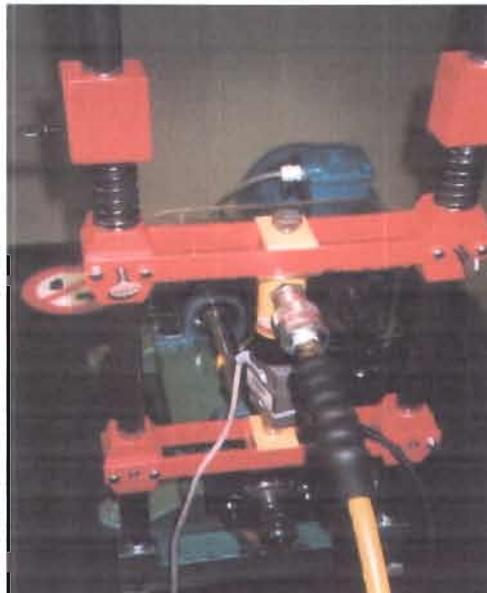


Figure 3.7: Test Bearing Loading Mechanism with Load Cell and Accelerometer

3.8.5 Data Acquisition Instrumentation

The vibration transducer, or vibration measuring instrument, chosen for use on the monitoring rig was an industrial type, model 624A11 accelerometer. The accelerometer was selected according to its sensitivity and mounting applicability, while ambient temperature was not critical since experimentation took place at room temperature. Since the high frequency detection technique was used in the monitoring of rolling element bearings, the sensitivity of the vibration transducer was important. The sensitivity of the piezoelectric quartz crystal accelerometer, was a hundred millivolts per unit gravitational acceleration, that is 100mV/g, where g is 9.81m/s^2 . Two accelerometers with the same sensitivity were used, namely, a magnetic base and a threaded stud accelerometer. The threaded stud accelerometer is accommodated on the bearing seat, and is located in the radial direction to the bearings rotational axis.

An 'S' type model LT400 load cell with a maximum load capacity of 200 Kg was used. This load cell was most suitable for this rig because of its small size, $76 \times 50 \times 25$ mm, and light weight of 248.5g as it is manufactured from anodised aluminium. The connection configuration allows the strain gauges to measure bending strain within the load cell. Applying a load on the tested bearing subjects the load cell to a compressive load. A remote optical sensor was applied as a tachometer.

The piezoelectric accelerometer, load cell and the remote optical sensor tachometer are connected to a data acquisition card. The data acquisition card used is the NI-DAQ 7.x and a driver software from National Instruments Corporation. There are two plug in devices, the accelerometer that is linked to device PCI-4474 and the tachometer that is linked to device PCI-6220. A PXI module was installed at the position where the signal conditioning box with configurable connectors links the load cell to the SC-2345 signal conditioning channel.

3.9 Artificial Infliction of Defects on Rolling Element Bearings

The requirement for condition monitoring of rolling element bearings was to have standard rolling element bearing vibration data that could be compared with data acquired on any monitored bearing at hand. The standard vibration data was to indicate the presence of a common spalling defect on the three components of a rolling element bearing, namely, the outer ring, inner ring and ball element. The standard vibration data of bearing defects was also required for the creation of an input data set that would be applied in the training of the Self-Organising Feature Map.

Condition diagnosis of previously damaged bearings was used as a guide, in terms of spall size, when inflicting artificial spalls through the application of a corrosive technique. The damage inflicted bearings was then used to obtain standard vibration data of a specific spall defect.

3.9.1 Application of the Corrosive Technique Method

Aquaregia was the corrosive solution used in the infliction of artificial defects. Aquaregia is a yellow acid mixture solution of hydrochloric acid and nitric acid, mixed in a ratio of 3H:1N. The bearings considered for experimentation were the 6012.ZZ3C, 6010.ZZ3C and 6008.ZZ3C deep groove ball bearings, with their respective bore diameters of 60, 50 and 40mm. The difficulty of dismantling a bearing for the creation of spalls was the main reason the corrosive technique was considered. The bearing remained assembled, but the seals were removed and the lubricant cleaned off. Copper tape with a hole template was placed on the raceway between two ball elements. The size of the template hole on the copper tape determines an estimated size of the corroded spall. The bearing was pre-heated to 30°C and kept constant for 30 minutes during which aquaregia was applied. Since surface components in a bearing are close together, tiny droplets of aquaregia would settle on these surfaces leaving stained marks as the aquaregia bubbles. The created spall is more significant than the additional stained marks.

For a single bearing size, spalls of $\pm 1\text{mm}$ and $\pm 2\text{mm}$ were created on the inner and outer raceways respectively. An inflicted spall of diameter $\pm 2\text{mm}$ is shown in figure 3.8 with a magnification of $\times 10$. A problem arose when trying to create a spall on the ball element, because of the limited space for placing the copper tape and the rounded surface of the ball element. So, to create a fourth stage damage and a ball damage, a 0.8mm slit was cut on each raceway for all bearing sizes. The slits were cut with an induction wire cutter.



Figure 3.8: An Artificially Inflicted Bearing Spall of $\pm 2\text{mm}$

Damage infliction was conducted on eight bearings of the same size to provide rolling element bearing vibration data. The first was a brand new bearing, the second and third bearings were inflicted with a $\pm 1\text{mm}$ spall on the inner and outer ring respectively, the fourth and fifth bearings had a $\pm 2\text{mm}$ spalls on the inner and outer ring respectively, the sixth and seventh bearings received a 0.8mm slit cut on the inner and outer ring respectively, and the eighth bearing a slit cut spall on the ball element.

3.10 Summary

Rolling element bearing vibration fundamentals were discussed in detail with explanations of the relation between bearing properties and bearing vibration characteristics. An understanding of the operations of rolling element bearings brought to light the fact that the source of induced bearing vibration is friction between the rolling element and the raceways. The design of the test rig and infliction of bearing defects was discussed. Different bearing defects were created so that a data set of bearing vibration measurements could be created.

Chapter 4

Signal Pre-processing Techniques

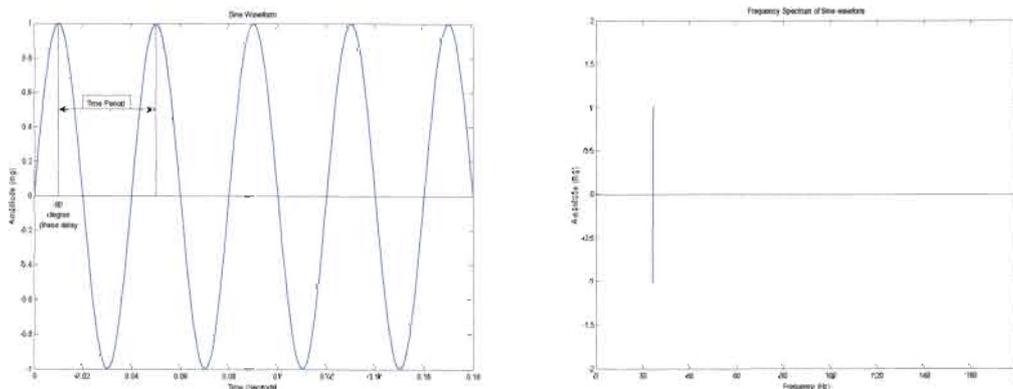
4.1 Introduction

The acquisition of bearing vibration measurements and analogue to digital signal processing of the measurements are discussed in this chapter. The application of the Fourier series, Fourier integral, Discrete Fourier Transform and ultimately the Fast Fourier Transform, to the bearing vibration measurement is discussed. The application of Fourier theory leads to better analysis of the measurement. Application of mathematical theory governing analogue to digital signal processing enables the application of statistical processing on bearing vibration measurements. Furthermore, this chapter focuses on the analysis of different formats of bearing vibration measurement through the signal processing software implemented.

Data acquisition and processing are performed. Here, data refers to the digital information obtained from a vibration measurement. Signal pre-processing assists in achieving the best results from vibration measurements. The relevance of the Fourier series and the progression to Fourier Transform, Discrete Fourier Transform and Fast Fourier Transform, is explained for better understanding of time domain and frequency domain analysis. The measurement parameters that characterise bearing vibration measurement are statistically defined. The signal processing software is developed to perform continuous bearing vibration analysis, and acquisition of the vibration measurements. A series of time domain and frequency domain measurements of bearing defects is illustrated.

4.2 Time waveform Development

Ramirez (1985:2) showed that the sine wave output on an oscilloscope can be described according to frequency. The time amplitude record of the sine wave, is termed the *time domain* waveform, (figure 4.1a), and the frequency amplitude record of the sine wave, is termed the *frequency domain* spectrum, (figure 4.1b). A *spectrum* is an arrangement or display of the frequency range contained in a waveform. The frequency representation of the sine wave is the reciprocal of the waveform period. The projected representation of the sine wave on to the frequency domain is in the form of an impulse that instantaneously rises and falls, with a peak amplitude equal to that of the sinusoid. A description of the time domain and frequency domain is illustrated in figure 4.1, with the time domain on the left and frequency domain on the right. The time domain shows a sine wave. A frequency impulse can be seen on the frequency domain, at a frequency of 25Hz, which is the frequency of the sine wave. The amplitude in the y-axis is in milli-gravitational units, $1mg = 0.00981m/s^2$.



(a)

(b)

Figure 4.1: Time Domain (a) and Frequency Domain (b) Waveform Representation

The time domain sinusoid is characterised by positive and negative peak amplitudes with a reference median amplitude at zero. Similarly, the frequency domain spectrum is characterised by impulses with positive and negative peak amplitudes. The symmetry of the impulse about the zero amplitude reference, makes it appealing to

project only the positive peak rather than the full peak to peak swing, in representing the sinusoid on the frequency domain. The position of the impulse on the frequency spectrum indicates the frequency of the sine wave.

At a time of zero seconds the sine wave does not have to begin with an amplitude of zero, but can begin with any amplitude. The position at which the sinusoid begins is characterised by its phase shift. In the case of the closest positive peak to time zero, occurring after time zero, the sinusoid is said to be delayed. A delay is denoted by a negative phase. In the case of the closest positive peak to time zero, occurring before time zero, then the sinusoid would have been advanced. An advance is denoted by a positive phase. The total range of phase shift that any wave can have is from -180° to $+180^\circ$ or 360° . The sine wave in figure 4.1 has a 90° phase delay.

Ramirez (1985:10) noted that the system of representing phase within a 2π range is referred to as *modulus 2π phase*. A sinusoid that is advanced by $360^\circ + 90^\circ = 450^\circ$ is not different from the same sinusoid advanced by 90° , so it can be represented as having a $+90^\circ$ phase shift. In the case of having a reference phase shift, then shifts beyond the 2π range can be represented as such. This system of representing phase beyond the 2π range is referred to as continuous phase representation.

In the above discussion, a single sinusoid was considered. Now consider three sinusoids having different frequencies and phase shifts, were presented on the same time domain plane. A waveform of the combination of all the sinusoids would result. Without a look at the frequency domain, changes in the shape of a sinusoid wave are the only indication that some frequency components have been modified. The frequency domain would show the specific frequencies of each sinusoid as impulses with positive peak amplitudes of the respective sinusoids.

Ramirez (1985:12) represented different frequencies on the same domain or plane, similar to the representation of different colours of the rainbow, and referred to the frequency domain plane as a frequency spectrum. The representation of physical waveforms in the time or frequency domain without the application of Fourier theory is referred to as analogue representation of a waveform. The phase shifts of the three sinusoids can only be determined mathematically using Fourier analysis, and not through analogue representation of waves on the time nor the frequency domain. Fourier analysis provides characteristic information of the waveform on the frequency domain plane, that is; the frequencies of the three sinusoids, their positive peak amplitudes, and their phase shift are determined. Figure 4.2 illustrates a combination waveform of three sinusoids on the time domain, and three frequency components on the frequency domain. The waveform was composed by sinusoids with frequencies of 20Hz, 50Hz and 60Hz respectively.

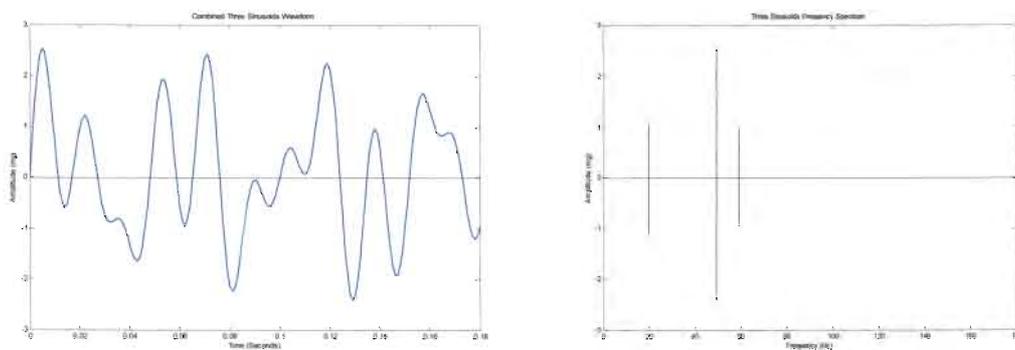


Figure 4.2: A Three Sinusoid Waveform Representation

Unfortunately, however, the classical approach to Fourier analysis is frustrating for all except the simplest of waveforms. The simplest of waveform would be a sinusoid. A waveform that is non-periodic and random cannot be mathematically formulated or expressed as an equation, and classic Fourier techniques cannot be applied. Special Fourier transforms are required for random waveforms.

4.3 Stationary Random Bearing Vibrations

Rao (1995:17) wrote that, if the magnitude of excitation acting on a vibratory system is known at any given time, the resulting system vibration is known as deterministic vibration. In a case where the magnitude of excitation is not known and cannot be predicted, then the resulting vibration is non-deterministic, or random, or stochastic. This is the random phenomenon of mechanical vibrations.

Rao (1995:821) recorded that vibration measurements that are not similar or identical even though the measurements were taken under similar conditions, give practical indication that the vibration is random. Bearing vibrations are an example of random phenomenon in that they are not identical in the magnitude of excitation in either force or motion for the same bearing condition, Dyer and Stewart (1978:227).

Bendat and Piersol (1986:542) wrote that a random process is a collection of time history records that can be described by appropriate statistical parameters. A stationary random process has statistical parameters that do not vary with respect to time. The latter is considered appropriate since constant speed conditions were applied in the experimental work for this study. Figure 4.3 illustrates a random bearing vibration of a new bearing in the time domain and in the frequency domain.

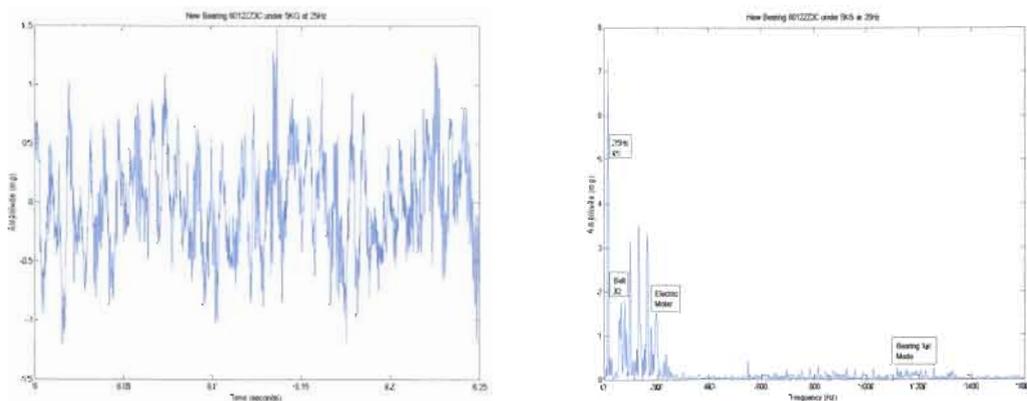


Figure 4.3: Random Vibration, Time and Frequency Response

4.4 The Fourier Series

Fourier analysis is performed through the application of Fourier series. The initial mathematical series of sine and cosine terms that could describe a periodic system were presented by Fourier in 1822, and is as follows;

$$y = \frac{a_0}{2} + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \dots \quad (4.1)$$

This expression is termed the Fourier series, and describes the frequency domain content of a periodic waveform. Where, y , is the periodic function; x , is the fundamental frequency term in radians per second; in equation 4.1, a_0 , is the value of the peak amplitude of the function over one period; a_1 , b_1 , a_2 and b_2 , are the coefficients of the sine and cosine terms and represent peak amplitudes at multiple frequencies of x . Fourier expanded the series to include the Fourier integral before his death in 1830. Each sinusoidal term in the Fourier series is some integer multiple of the fundamental frequency and is referred to as a *harmonic*. The fundamental frequency is sometimes referred to as the *first harmonic*, but integer multiples greater than one are always referred to as harmonics. These harmonics are presented on the frequency domain plane as impulses at the frequencies of the smaller amplitude sinusoids.

4.5 The Fourier Integral

The Fourier integral is the tool used to investigate the frequency spectra of non-periodic waveforms. Kreyszig (1999:559) noted that the Fourier integral is developed from the Fourier series. Consider a waveform represented by a function $x(t)$ in a certain time interval (t), whose time period T is allowed to approach infinity. The time period ranges from negative infinity to positive infinity, that is, $-\infty$ to ∞ .

This means that the waveform does not repeat itself. To get the Fourier integral of this waveform, the general form of the Fourier series shown in equation 4.1 should be re-expressed in exponential form. Assignment of new variables and manipulation of Fourier coefficients results in a more compact exponential form of the Fourier series.

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad (4.2)$$

c_n is evaluated for the number of sinusoids that compose the waveform. In this case $n = -\infty, \dots, -2, -1, 0, 1, 2, \dots, \infty$. When the Fourier coefficients have become a function of a continuous frequency variable ω , and f in Hertz so that $\omega = 2\pi f$. The Fourier coefficients are given by the *Fourier transform*;

$$f(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt \quad (4.3)$$

Although the Fourier integral is an improvement as compared to the Fourier series, both these methods of analysis require a mathematically defined waveform, which is complex for bearing vibrations. The solution to this problem is the Discrete Fourier Transform, (DFT).

4.6 The Discrete Fourier Transform

Proakis and Ingle (2004:120) recognised that digital machines do not work with infinite values, but work with samples of time variables that are finite. A waveform is sampled if its amplitude values are determined at discrete points along the time axis. This means that the waveform can be plotted for these determined amplitude values, within the considered time interval. This plotted waveform is considered to be windowed because it has been limited within a finite time interval. Each sample exists at a single discrete time point, and its value is the amplitude of the waveform.

An analogue waveform $x(t)$, recorded over a time interval, t , is digitised by taking, N , number of samples. Each sample, n , is part of a sequence comprising the total number of samples taken, that is $n = 0, 1, 2, \dots, N - 1$. Sample $n = 0$ is the direct current sample, equivalent to an average amplitude term in a Fourier series. Each sample, n , is measured for the amplitude of the analogue waveform $x(t)$ after every sampling period, ϕ , over the total time interval, t . This means that the sampling rate or sampling frequency, $f_s = 1/\phi$. The Nyquist frequency, f_n , is the largest frequency sinusoid of the waveform that can be defined at a given sampling rate f_s . This sinusoid has a frequency $f_n = f_s/2 = 1/2\phi$.

Stander (2003:10) referred to the Nyquist theorem which states that if a waveform $x(t)$ is sampled at a rate f , that is equal to or more than twice the value of the waveform's largest frequency f_n , then the analogue waveform $x(t)$ can be recovered from the digital windowed waveform $x(n\phi)$ precisely, and the digital time series contains all the information required to reconstruct the analogue waveform.

Although the discrete Fourier transform can be expressed in terms of the Fourier integral, a more general mathematical expression of the *discrete Fourier transform* is as follows, with $\lambda = 2\pi/N\phi = 2\pi/t$ radian per second;

$$f(k\lambda) = \phi \sum_{n=0}^{N-1} x(n\phi) e^{-jk\lambda n\phi} \quad (4.4)$$

Equation 4.4, enables the transformation of a time series of samples, to a frequency series of samples. Now, a random vibration waveform can be transformed to provide discrete frequency and amplitude values of sinusoids comprising the waveform. The performance of large numbers of repetitive calculations and the consequent time taken is the disadvantage of the discrete Fourier transform.

4.7 The Fast Fourier Transform

The *fast Fourier transform* (FFT) is an algorithm for computing the discrete Fourier transform in a time efficient way. The fast Fourier transform reduces the number of major operations of the Discrete Fourier Transform. To simplify the algorithm, the length of the transform or total number of samples taken N should be a power of two, that is $N = 2, 4, 8, 16, \dots 2^b$, and b is an integer. A large number of samples provide better time resolution and results in the calculation of more frequency domain points.

The fast Fourier transform performs, $N \log_2 N$, major operations as compared to N^2 operations of the Discrete Fourier transform. The number of major computational operations are reduced resulting in a fast algorithm for determining Fourier coefficients of the waveform, hence the algorithm's name Fast Fourier Transform.

The original Cooley-Tukey algorithm uses a process referred to as a decimation in time, the Sande-Tukey algorithm uses decimation in frequency. These two algorithms apply where the number of samples $N = 2^b$, and b is an integer. While the two methods differ, the result produced is the same. The fast Fourier transform algorithm used for signal processing here, is the Cooley-Tukey algorithm. Since fast Fourier transform computes a discrete Fourier transform, equation 4.4 applies.

4.8 Time Waveform Statistical Parameters

Oppenheim and Willsky (1983:85) noted that a digital waveform can be statistically analysed according to the variation of the amplitudes with time. The average and mean value of the amplitudes of one waveform can be determined. A number of vibration measurements taken on one bearing operating under the same conditions for each measurement, will have similar statistical parameters, but will not necessarily be equal.

The statistical parameters will not be equal, because the vibration measurements are taken at different times. A group of vibration waveforms taken on one type of bearing defect, will have statistical parameters that are characteristic of the type of bearing defect concerned. A group of vibration waveforms and their statistical parameters form a data set, this is the plan for creating an input data set.

The computation of statistical parameters on one waveform requires that the amplitude values be normalised. Normalisation is a process of scaling the numbers in a data set, in this case amplitude values, to improve the accuracy of the numerical computations that follow. A way to normalise data is to find the difference between an amplitude value and the mean amplitude value, and then scale each data value to unit standard deviation as follows;

$$\text{Amplitude Data} = \frac{[\text{Amplitude Data} - \text{Mean (Amplitude Data)}]}{\text{Standard Deviation (Amplitude Data)}} \quad (4.5)$$

Once the waveform is normalised from amplitude values that are instantaneously too large or too small, the application of statistical formulas can be considered.

The following statistical parameters are computed from a bearing vibration time waveform:

(1) Rectified average, RA, is a computation of a positive valued average, which avoids having a negative sign on the average. Symmetry of the time waveform about the time axis allows for the consideration of only the positive amplitudes. The number of samples taken for each waveform was set at $N = 2048$. With the amplitude values defined as, y , the rectified average is;

$$RA = \frac{\sum_{n=0}^{N-1} |y_n|}{N} \quad (4.6)$$

(2) The mean, μ , is the average amplitude and the first central moment of distribution.

$$\mu = \frac{\sum_{n=0}^{N-1} y_n}{N} \quad (4.7)$$

(3) The range, rng , is the difference between the highest and lowest recorded amplitudes.

$$\text{rng} = \max(y) - \min(y) \quad (4.8)$$

(4) The median, med , is the centre amplitude, if the amplitudes are placed in ascending order.

(5) Variance, va , is the second central moment of distribution, and a measure of how much the data differ from the mean. The deviation ($y_n - \mu$) is squared to avoid a negative sign value.

$$\text{va} = \frac{\sum_{n=0}^{N-1} (y_n - \mu)^2}{N} \quad (4.9)$$

(6) The standard deviation, σ , is a measure of the data spread, range or deviation from the mean. Standard deviation is the square root of the variance.

$$\sigma = \sqrt{\frac{\sum_{n=0}^{N-1} (y_n - \mu)^2}{N}} \quad (4.10)$$

(7) Root mean square, RMS, estimates the mean of the normal distribution. The root mean square is equivalent to the standard deviation, so $\text{RMS} = \sigma$.

(8) The maximum peak, P_k , is the largest amplitude value in a waveform, and due to symmetry only the positive peak is considered.

$$P_k = \max|y| \quad (4.11)$$

Bunch *et al.* (1995:412) noted that RMS and Peak values can be applied to reflect the energy level of the vibration, however they cannot be used for single snapshot detection of bearing damage as the expected values generally exhibit a wide range depending on the operating conditions such as shaft speed and bearing load.

(9) Crest factor, C_f , is a ratio between the peak amplitude and the root mean square or standard deviation. The crest factor is used to evaluate the sharp peaks in the vibration waveform, like those generated from rolling element bearing defects. The crest factor can be trended for evaluation purposes. Akedson *et al.* (1985:1) highlighted that the crest factor is relatively insensitive to changes in bearing speed and load. This enables direct assessment of bearing conditions with minimal knowledge of its history.

$$C_f = \frac{P_k}{\text{RMS}} \quad (4.12)$$

(10) Skewness, Skw , is the ratio between the third central moment of distribution and the cube of the standard deviation. The third central moment of distribution, which is related to skewness, is used to identify lack of symmetry in the data distribution due to bearing defect impulses. Akedson *et al.* (1985:265) noted that for a theoretically perfect normal distribution that follows the Gaussian function, all the odd moments will be zero and the even moments take on finite values. Rectifying the third central moment by the standard deviation, results in the odd moments taking on finite values and no longer the value of zero. Hence, the third central moment can be used to monitor the bearing condition in the form of a variable of skewness.

$$\text{Skw} = \frac{\sum_{n=0}^{N-1} (y_n - \mu)^3}{N \sigma^3} \quad (4.13)$$

(11) Kurtosis, K , is the ratio between the fourth central moment of distribution and the fourth power of the standard deviation. The fourth central moment is rectified by the standard deviation to enable odd moments taking on finite values. Akedson *et al.* (1985:270) showed that an undamaged bearing has a kurtosis value of 3.0, indicating a Gaussian probability density function. Applying the high frequency detection technique, a value as high as 10.0 signals incipient bearing damage.

$$K = \frac{\sum_{n=0}^{N-1} (y_n - \mu)^4}{N \sigma^4} \quad (4.14)$$

(12) Dyer *et al.* (1978:232) and Alfredson *et al.* (1985b:4) expressed the amplitude characteristics of a vibration waveform $x(t)$ in terms of a probability density function (PDF), $P(x)$. This is estimated by determining the time duration for which a signal remains in a set of amplitude windows, that are equal to the total samples, N , of the waveform.

$$P[x \leq x(t) \leq (x + \Delta x)] = \sum_{n=1}^N \frac{\Delta t_n}{T} \quad (4.15)$$

Where, Δt_n , is the time duration of a single sample, n , that falls in a single sample amplitude window Δx of the vibration waveform $x(t)$. Here, T , is the window time.

4.9 Data Acquisition and Signal Processing Software

The data acquisition instrumentation used in the test rig is configured to the LabVIEW version 7.1 signal processing software. LabVIEW was used to perform online signal processing because it is applicable to continuous condition monitoring. The graphic user interface in labVIEW provides a suitable environment for online signal processing. LabVIEW operates through the creation of virtual instruments.

Gani *et al.* (2004:64) referred to a virtual instrument as an instrument based on a graphical concept of data flow programming. The layout of the signal processing path in a virtual instrument begins with acquiring an electrical signal from a data acquisition card device. Then the signal is linked to a signal processing function such as the time domain and frequency domain functions, and through the application of signal controls and calibration factors, the signal can then be displayed in a graphical window spectrum or stored in a digital storage device. Signal windowing, anti-aliasing, averaging and amplification are automatically performed before the signal is displayed in the time and frequency domain. The frequency domain function implements the Cooley-Tukey algorithm for the Fast Fourier Transform.

A virtual instrument for the rolling element bearing condition monitoring rig was created, of which the construction flow diagram can be seen in figure 4.4. From the diagram it can be seen that the vibration signal and the load cell signal were acquired from different data acquisition cards and processed through different paths. The complete virtual instrument display window shows; a time domain and frequency domain vibration spectrum, a statistical probability density distribution histogram and the variation of statistical parameters concerned, a tachometer and load cell time domain window together with the digital speed and load scale display. The load scale was calibrated manually using a standard 5kg weight.

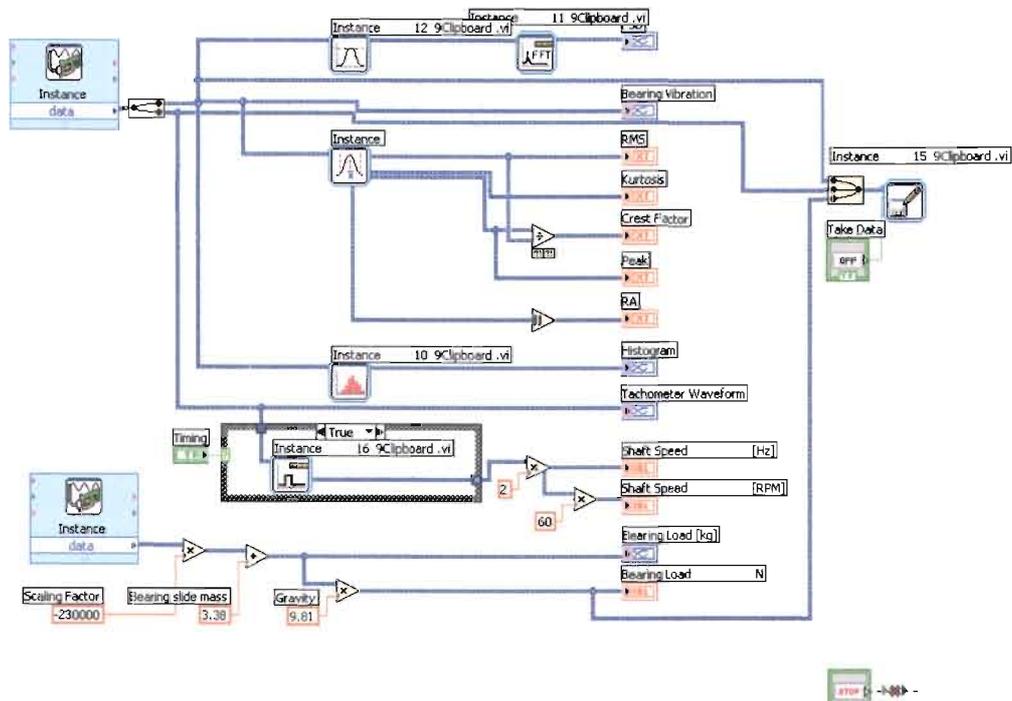


Figure 4.4: Virtual Instrument Construction Flow Diagram

The virtual instrument display window can be seen in figure 4.5. Matlab version 7.1 was used for offline signal processing which is suitable for periodic condition monitoring. Matlab was however most effective in experimental computations and neural network applications. The self-organising feature map toolbox was accessible with classes of programs that perform the Kohonen Self-Organising Feature Map algorithm.

The virtual instrument and signal processing tools were set for a waveform having 2048 samples, the number of samples in the positive frequency spectrum $k = 2048/2.56 = 800$. Frequency resolution = $20000/800 = 25$ Hz. The resolution was estimated to have at least 3 times the resolution between the bearing frequencies. Therefore the bearing frequencies could be closer than 3×25 Hz = 75Hz. To improve the resolution, a narrower frequency bandwidth needed to be selected.

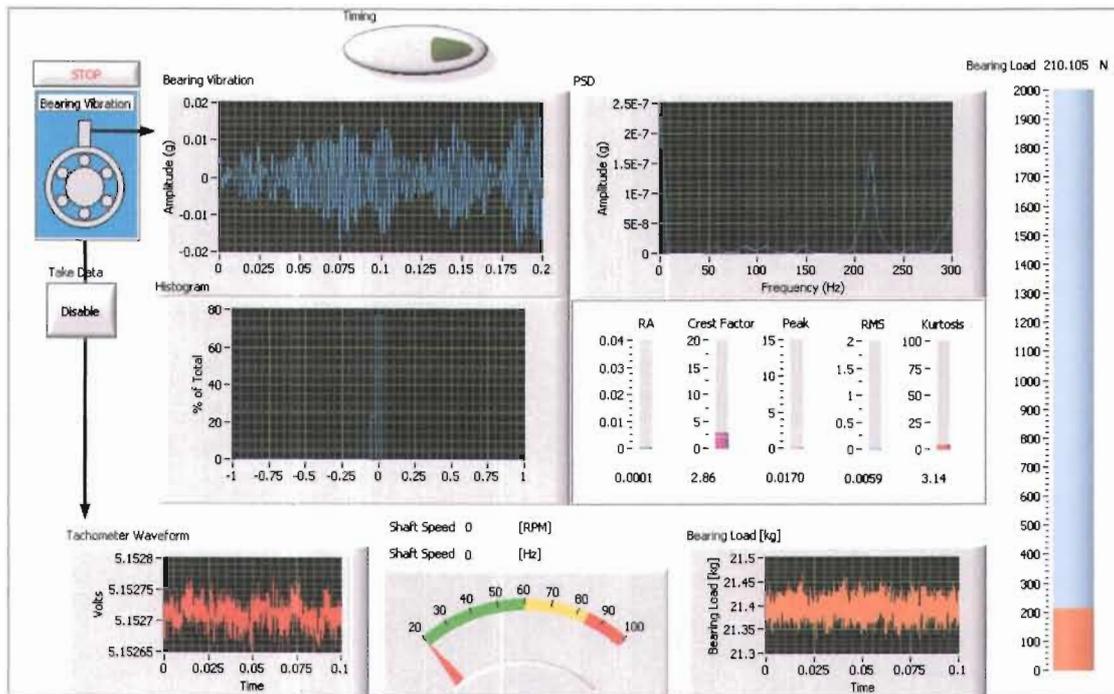


Figure 4.5: Online Signal Processing Virtual Instrument Display Window

4.10 Filtering of the Bearing Vibration Measurements

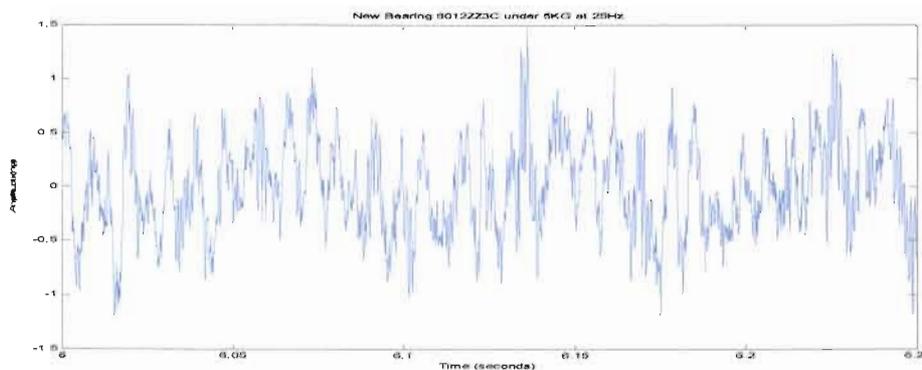


Figure 4.6: New Bearing Unfiltered Signal Time Response

Filtering reduces signal noise or conditions a waveform. The Butterworth infinite impulse response filter was used to remove the low frequency bands or sinusoids that are not required so that only bearing frequencies are present for analysis. The filtered waveform in figure 4.7 shows more closely spaced high frequency pulses than the

unfiltered waveform in figure 4.6. Notice that the low frequencies up to 50Hz have been cut off in the frequency domain of the filtered waveform, (figure 4.9).

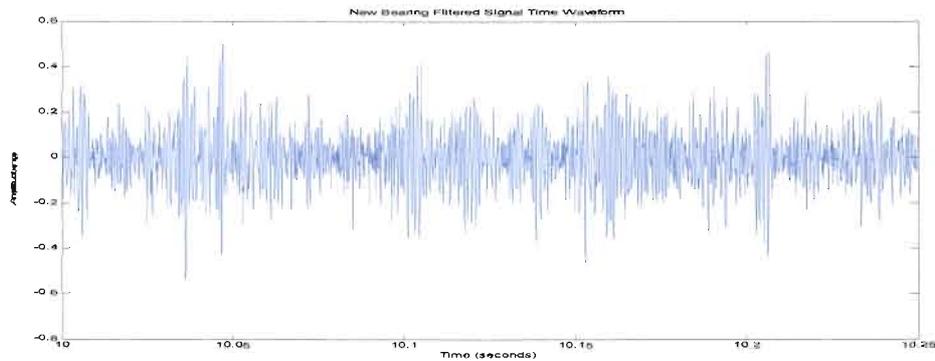


Figure 4.7: New Bearing Filtered Signal Time Response

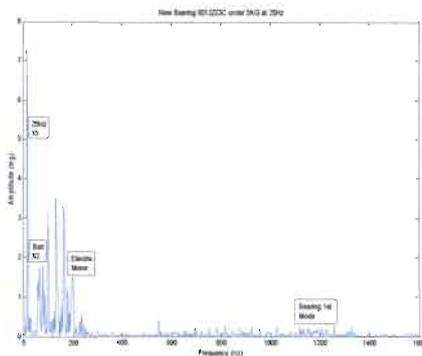


Figure 4.8: New Bearing

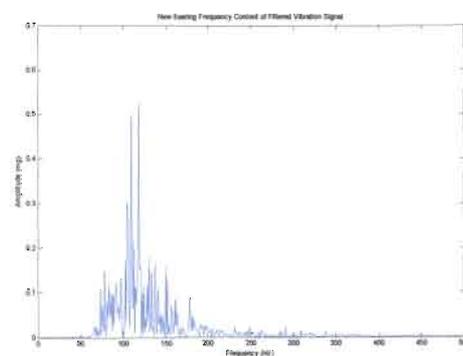


Figure 4.9: New Bearing Filtered Signal

4.11 Induced Bearing Frequencies

The monitoring test conditions were kept stationary. This means that the applied load was kept constant, and so was the test shaft speed. There was, however, the inevitable small dynamic load due to the uneven rotation motion of the ball elements in and out of the load zone even when all misalignment is removed. The monitoring test speeds were those considered to be safe for the rig thereby avoiding the natural resonance frequencies of the loading mechanism and belt drive. The condition monitoring test speeds are shown in table 4.1. Table 4.2 shows the induced bearing frequencies at 25Hz shaft speed according to the program in annexure 3 (p138).

Table 4.1: Condition Monitoring Testing Speeds

Speed 1	Speed 2	Speed 3	Speed 4	Speed 5
900 rpm	1200 rpm	1500 rpm	1800 rpm	2400 rpm
15 Hz	20 Hz	25 Hz	30 Hz	40 Hz

Table 4.2: Experimentation Induced Bearing Frequencies

Bearing	Shaft	BPFO Hz	BPFI Hz	2BSF Hz	FTT Hz	Bspeed
6008ZZ3	25Hz	127.9444	172.0556	166.3493	10.6620	83.1746
6010ZZ3	25Hz	151.4962	198.5038	182.7821	10.8212	91.3910
6012ZZ3	25Hz	151.6968	198.3032	184.4132	10.8355	92.2066

4.12 Loading Configuration

The bearing loads considered for experimentations had to be determined through bearing design calculations and rig configuration. The selection of the load was based on the minimum operating or static load that could be sustained by any one of the bearings considered, which would also be safe for the monitoring rig.

The monitoring rig configurations refer to the structural component mass that is additional to the hydraulic induced load applied on the bearing. The configurations take into consideration the mass of the different size bearing seats for different bearing sizes and whether the accelerometer used is that of the magnetic type or stud mount type. The configurations are tabulated in table A4.1 of annexure 4 (p139). The rated design static loads are in table A4.2 of annexure 4 (p139), of which the lowest was 61.16KG. The operational load was calculated for an operational period up to failure of three years, two years, and one year.

4.12.1 Rolling Element Bearing Life

To determine the testing loads, bearing life calculations had to be considered. The L_{10} bearing life is a prediction of the operational period of the bearing if the bearing is used under proper application conditions, lubrication, alignment and temperature.

The following formulas were used to show the relationship between load and bearing life.

$$L_{10} = \left(\frac{C}{P} \right)^3 \quad (4.16)$$

Where L_{10} = Rating life in millions of revolutions, and the base of ten refers to the rating life in terms of powers of ten, C = Static load that will allow one million revolutions before bearing failure, and P = Applied static load on the bearing. The bearing life rating may be calculated in terms of hours as in equation 4.17.

$$Lhr_{10} = \frac{16.667}{N} \left(\frac{C}{P} \right)^3 \quad (4.17)$$

Where N = Shaft rotational speed in revolutions per minute, Lhr_{10} = Rating life in hours that are in powers of ten. The equations above for the design of bearing life hold true for operational conditions where the shaft rotational speed, N , is considered to be constant and not accelerating.

The programmed calculations shown in annexure 4 (p139), indicated that for the three years operational period the bearings would have to be subjected to the lowest loads. With the highest load being 31.7KG and the lowest being 18.03KG. Thus, the highest rig monitoring load was chosen to be 20KG. 20KG was safe for the rig and could be achieved by the hydraulic system. Monitoring tests were performed at the configuration load, that is when the hydraulic system is not activated. For the 6012ZZ3C bearing the configuration load was 4kg. The other loads considered for monitoring tests were 5, 10, 15 and 20 kilograms.

4.13 Vibration Measurement Acquisition

4.13.1 Natural Responses of the Rolling Element Bearings

Vinh (1997:269) used impact testing with a manual hammer in experimental structural dynamics. Usually the test is applied at the beginning of a testing campaign to create an overall view of natural frequencies of a mechanical structure. The ability to overlay fault frequencies helps to accurately identify a particular bearing fault. Bearing faults cause the bearing and other machine parts to respond at their natural frequencies. So knowledge of the bearing and machine's natural frequencies assist in detecting bearing faults. Dominant frequencies are the bearing's natural frequencies. Figure 4.10 and figure 4.11 show the impact test, time and frequency response of the 6012ZZ3C deep-groove ball bearing.

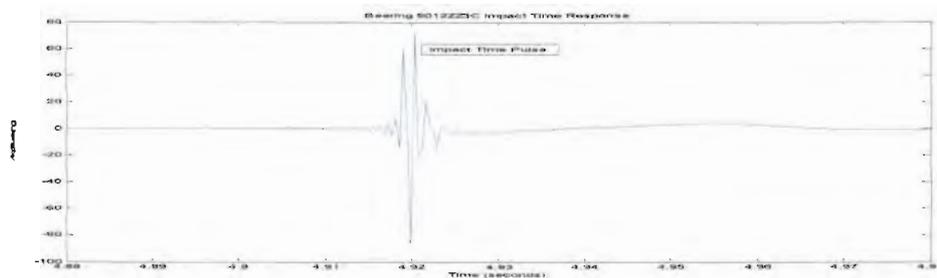


Figure 4.10: Bearing 6012ZZ3C Impact Time Response

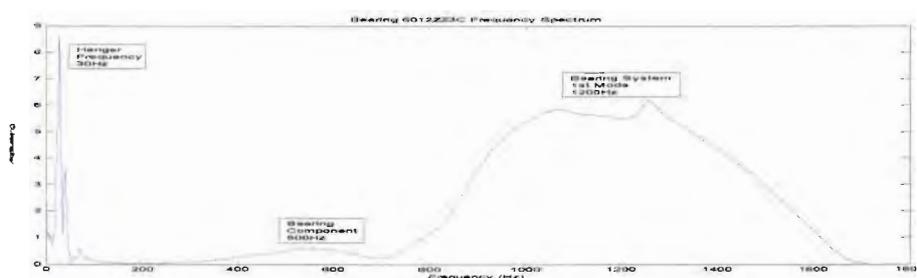


Figure 4.11: Bearing 6012ZZ3C Impact Frequency Response

The time domain displays an instantaneous impact response, and the frequency response displays a bearing component frequency at 500Hz this is most likely that of the cage in contact with ball elements since the overall bearing frequency is at 1200Hz. The natural frequencies of bearing 6008ZZ3C and bearing 6010ZZ3C,

were also determined through the impact test method, and their frequency responses are displayed in annexure 5(p140).

4.13.2 Natural Responses of the Condition Monitoring Rig

Impact test method was used to determine the natural response of the critical structural component of the rig, the loading mechanism, because the accelerometer is located in this structure and the load is transmitted through the structure. Figure 4.12 and figure 4.13 show the time and frequency response of the loading mechanism.

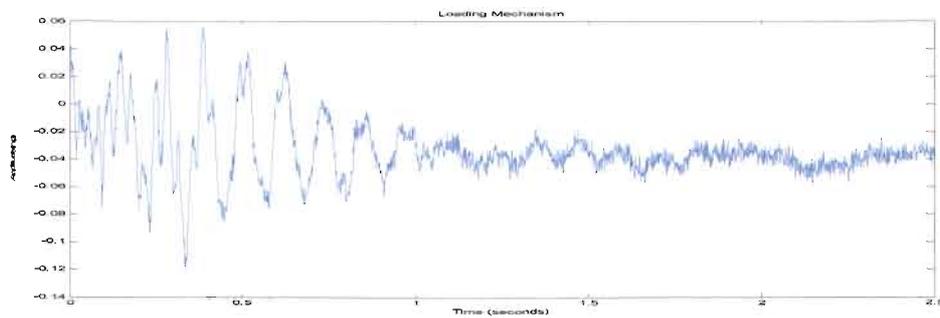


Figure 4.12: Loading Mechanism Impact Time Response

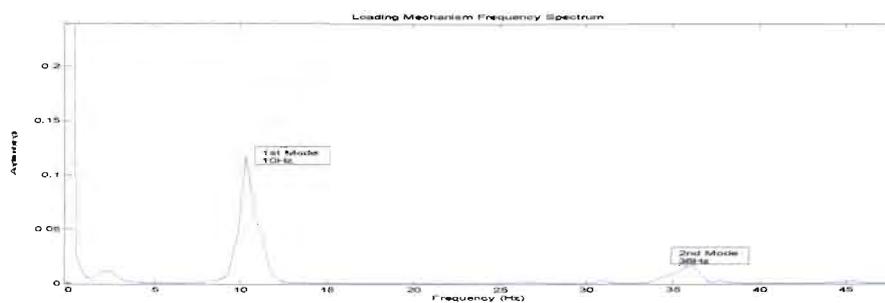


Figure 4.13: Loading Mechanism Impact Frequency Response

The time domain shows a broad impact response, while the frequency domain reveals that the first mode of the load mechanism's natural frequency at 10Hz and the second at 36Hz. Rotational testing shaft speeds of 10Hz and 35Hz were avoided. The frequency response of the electric motor casing [shown in annexure 5 (p140)], indicates a response at 200Hz, and that of the base-plate at 350Hz.

4.14 Bearing Vibration Defect Pattern Measurements

The bearing vibrations presented below were acquired at 25Hz with the 6012ZZ3C bearing subjected to an operating load of 5kg. For this reason the induced bearing frequencies are presented for a shaft speed of 25 Hz, whereas the other induced bearing frequencies are determined from annexure 3 (p138). The purpose here is to show that there are differences in the vibration pattern depending on where the spall is located on the bearing. The raceways were slightly oiled but not excessively lubricated, just enough to prevent metal to metal contact and excessive noise, but also to ensure that only bearing defects are measured and not the effects of lubrication on measurements. Since all machine malfunctions were corrected, it was practically assumed that the vibration measurements were of bearing defects only.

4.14.1 Outer Race Incipient Defect Bearing Noise

The time domain in figure 4.14 is of a new bearing that was not lubricated with oil, to illustrate how the earliest incipient damage would occur. Although the noise amplitudes were not high the noise peaks appear close together because of the metal to metal contact. On the frequency domain, the metal to metal contacts excited the 10Hz frequency of the loading mechanism. At twice the shaft frequency, 50Hz, the contacts triggered the belt resonance frequency.

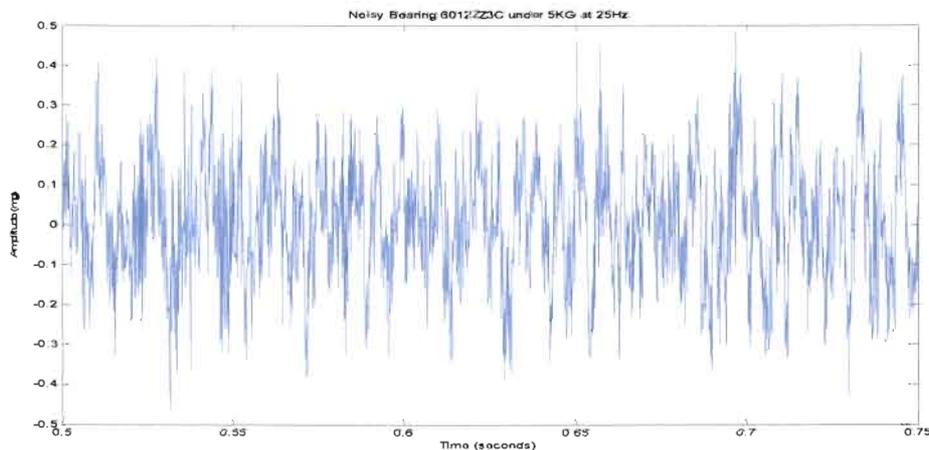


Figure 4.14: Noisy Bearing Time Response

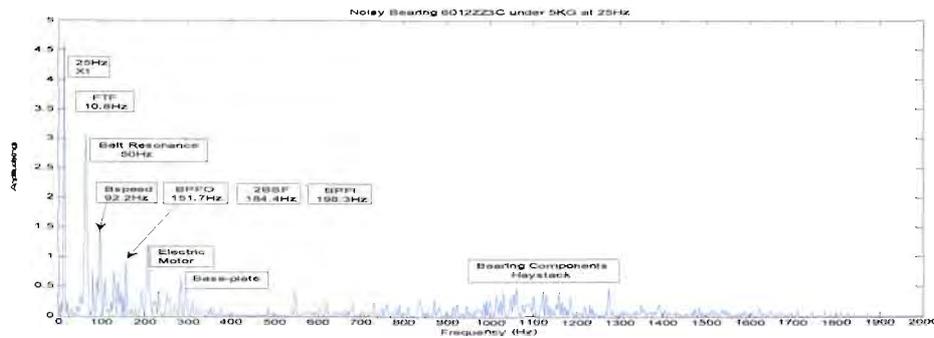


Figure 4.15: Noisy Bearing Frequency Response

The sidebands of the belt resonance are very low, indicating that the problem is not with the belt but that excitation is emanating from the bearing. The amplitudes at the induced bearing frequencies were low indicating that there are no definite bearing defects, and that the electric motor casing and the base-plate are responding at 200Hz and 300Hz frequency range respectively. The fundamental train frequency (FTF) might have influenced the response of the loading mechanism because they are of equal frequencies. The important feature is the rise in amplitude at the natural frequency of the bearing, in the range 1000Hz to 1200Hz. This rise in amplitude at 1000Hz - 1200Hz, may be referred to as the haystack, and indicates the most earliest incipient bearing damage occurring. Another indication is that machine malfunctions, such as misalignment, were minimum if not completely removed.

4.14.2 New Bearing Noise

Figure 4.16 and figure 4.17 are the time and frequency responses of a new bearing however this time adequate lubrication had been provided. In the time domain the noise amplitude was low, the noise peaks are within a low frequency carrier waveform as displayed. Due to the low frequency carrier the amplitude is slightly higher than that shown in figure 4.14.

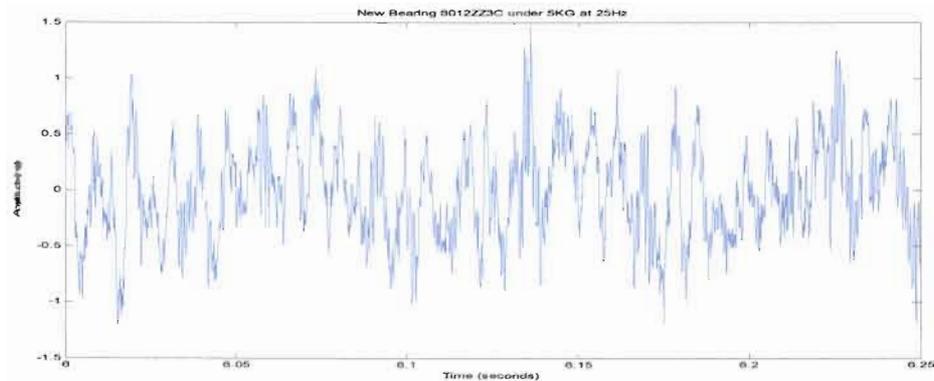


Figure 4.16: New Bearing Time Response

The frequency domain shows that the low carrier frequency is 10Hz, which was that of the loading mechanism. The peak in the frequency domain is an indication of the energy of the 10Hz waveform. This 10Hz response was the result of the inevitable dynamic load from the bearing. The shaft frequency, 25Hz, and belt resonance frequency, 50Hz, have very low amplitudes due to the dynamic loading only. Without dynamic loading these amplitudes would be similar to signal noise. The induced bearing frequencies are between the 50Hz and 200Hz frequency band cause by the dynamic load and slight metal to metal contact. Electric motor casing frequencies are also responding. This is a baseline measurement to which incipient bearing damage can be compared. There were no machine faults and bearing faults present.

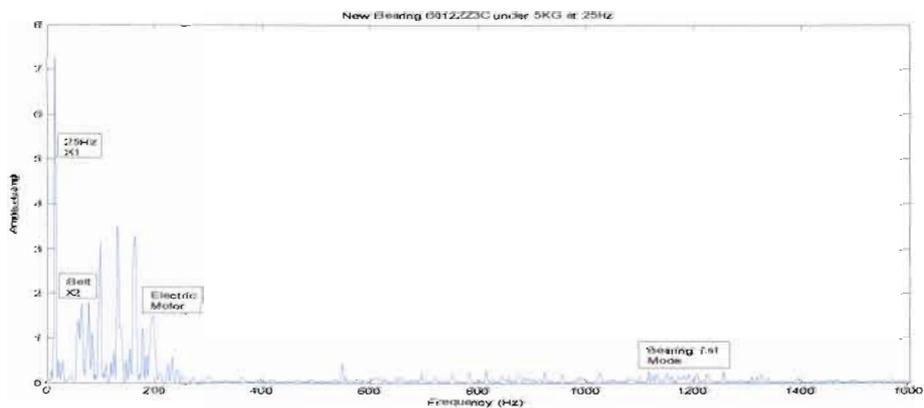


Figure 4.17: New Bearing Frequency Response

4.14.3 Ball Bearing Defect

A vibration measurement of a spall defect on the ball element is shown on the time domain in figure 4.18, and the frequency domain in figure 4.19. Two definite quick impulses can be seen at the ball speed or ball spin frequency. The uneven rotation of the ball is the reason the pulses are not evenly repetitive.

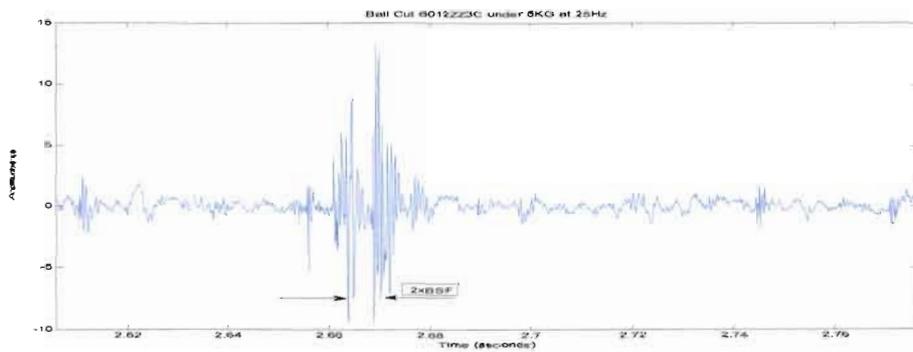


Figure 4.18: Ball Cut Time Response

On the frequency domain, there is a peak at the ball speed, or ball spin frequency (BSF) of 92.2Hz, and at the frequency at which the ball excites the inner and outer race, twice ball spin frequency ($2 \times \text{BSF}$), of 184.4Hz. The whole frequency band 50Hz - 200Hz responds with visible side bands. The impulse energy is high, at 30mg's (Amplitude is in milli-gravitational acceleration, $1\text{mg} = 0.00981\text{m/s}^2$). The amplitude in the bearing resonance range 1000Hz - 1500Hz is high. The bearing component frequencies of the contacting cage and ball elements is high at 500Hz, 700Hz.

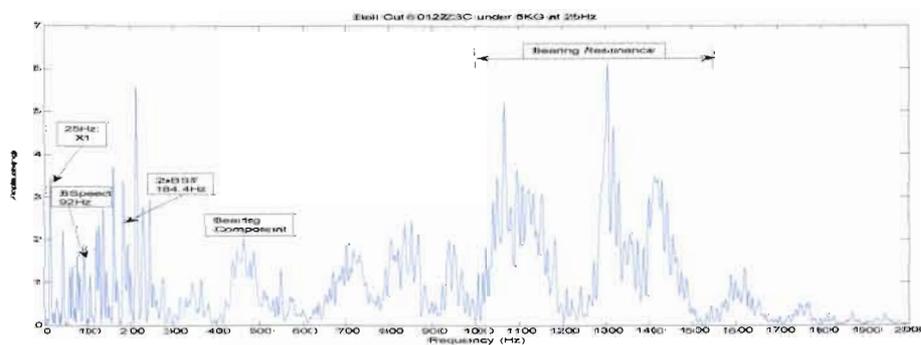


Figure 4.19: Ball Cut Frequency Response

4.14.4 Inner Race Spall Defect

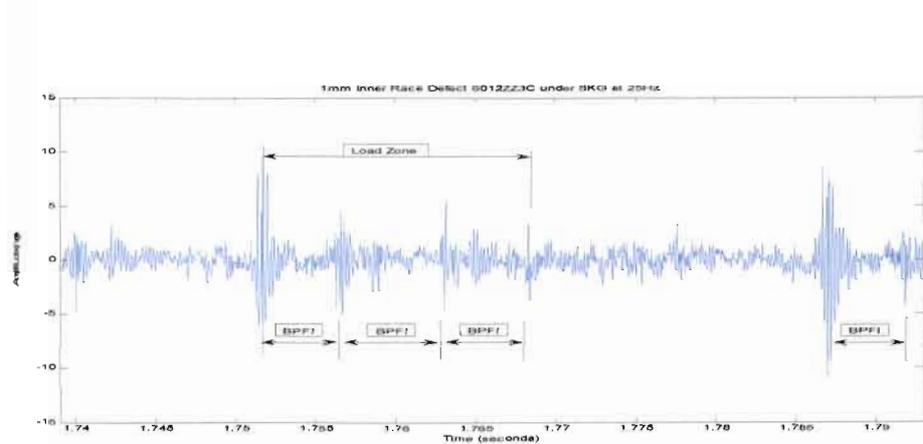
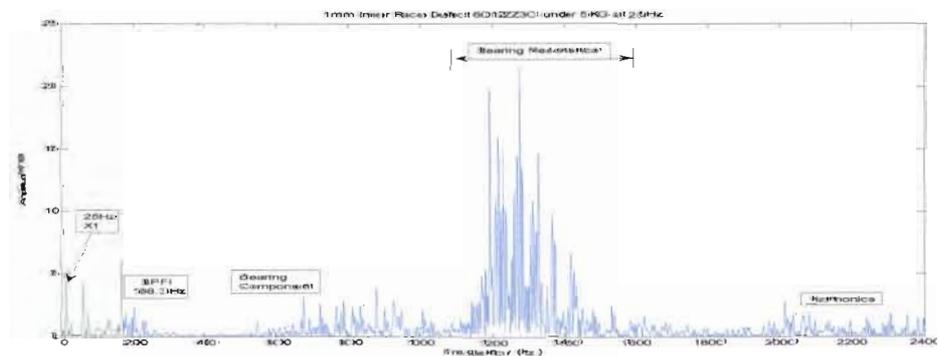
Figure 4.20: ± 1 mm Inner Race Defect Time Response

Figure 4.20, reveals a pulse train of a ± 1 mm inner race spall. There were four distinctive pulses created when the spall passed through the load zone, at the ball pass frequency of the inner ring (BPFI). The pulses out of the load zone generated low amplitudes.

In figure 4.21 the ball pass frequency of the inner race (BPFI) at 198.3Hz has a visible peak indicating an inner race defect. The bearing component frequencies are low indicating that the defect is at a specific location. The amplitude at the bearing resonance frequency (1200Hz) which means that there is a bearing fault. There are visible harmonics of the induced frequencies at 2000Hz.

Figure 4.21: ± 1 mm Inner Race Defect Frequency Response

4.14.5 Outer Race Crack Defect

With the outer race defect, a pulse train through out the whole time domain is visible, since the defect was in the load zone. Figure 4.22 shows the pulses with similar, but not equal amplitudes because of the dynamic load in the bearing and the uneven rotation of the ball elements. The pulses occurred at the ball pass frequency on the outer ring spall (BPFO).

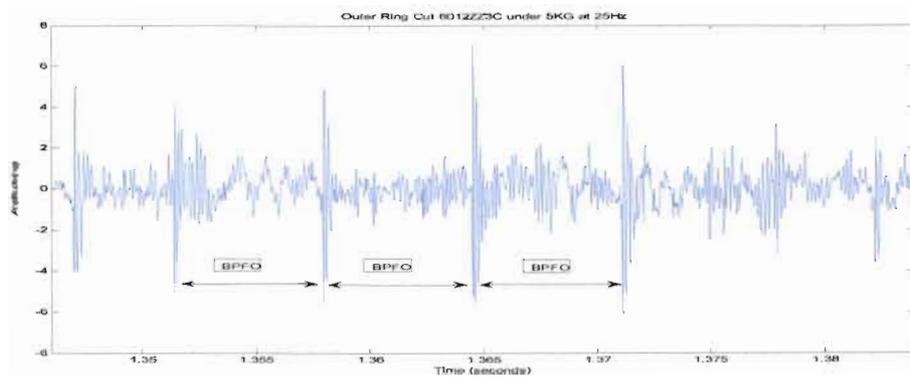


Figure 4.22: Outer Ring Cut Time Response

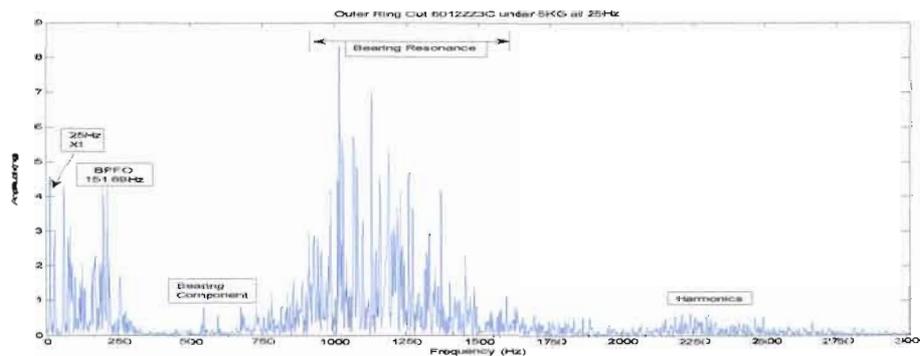


Figure 4.23: Outer Ring Cut Frequency Response

On the frequency domain there is a definite peak at the (BPFO) of 151,67Hz, indicating a defect on the outer ring. The frequency band of the bearing response is wider due to induced harmonics (700Hz - 1600Hz) and higher frequency harmonics are visible. The high amplitudes at the bearing resonance frequencies and at the BPFO indicate an outer race fault on the bearing.

4.15 Bearing Vibration Defect Severity Measurements

The purpose here is to show that there are differences in the vibration pattern depending on the severity of the bearing defect, as the bearing damage progressively passes through the four stages of failure. A single location for all the defects was chosen on the outer race. The testing operational conditions were not changed. Since all machine malfunctions were corrected, it is assumed that the vibration measurement was of the bearing defect only.

4.15.1 Stage one Bearing Defect

Stage one is represented by a new bearing. The amplitude values from the time domain can be represented by a histogram. A lubricated new bearing has a steeper histogram than a bearing that is not lubricated, because there is less noise peaks in the lubricated bearing. Both histograms follow the Gaussian curve, although the variances differ.

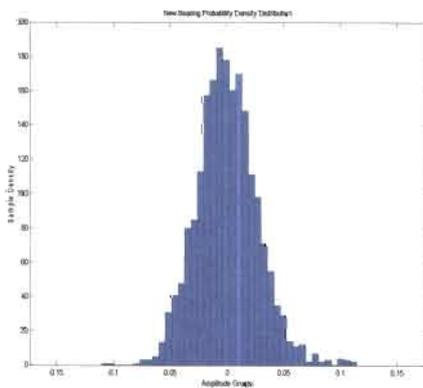


Figure 4.24: Stage One New Bearing

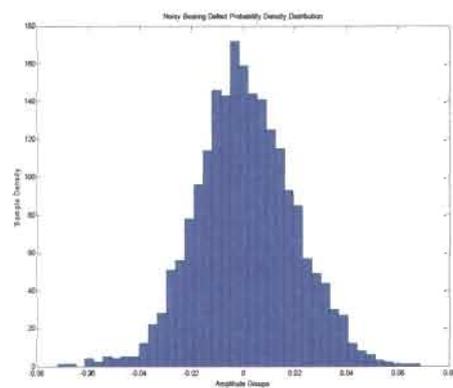


Figure 4.25: Stage One Noisy Bearing

4.15.2 Stage Two Bearing Defect

Stage two of incipient bearing damage is represented by a ± 1 mm spall on the outer ring. The time domain of this defect is shown in figure 4.26, where a pulse train can be seen through out the whole domain. The spall was small enough for the pulse to be sharp because only the outer ring responded.

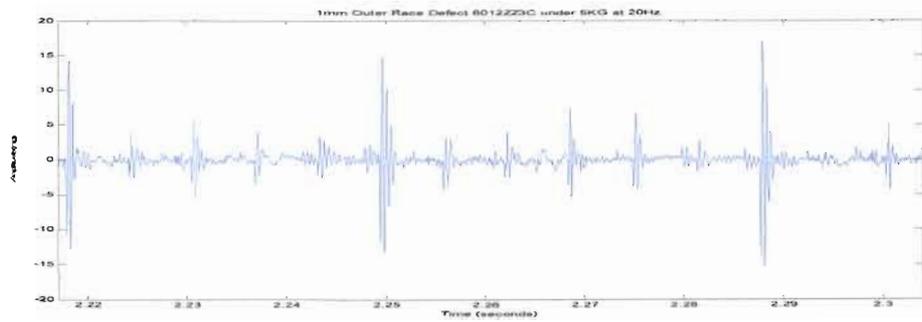


Figure 4.26: Stage Two Outer Race Defect Time Response

The histogram (figure 4.27) indicates that the majority of the amplitudes represent noise. The steepness or low variance of the histogram indicates that only the outer ring pass frequency (BPFO) responded.

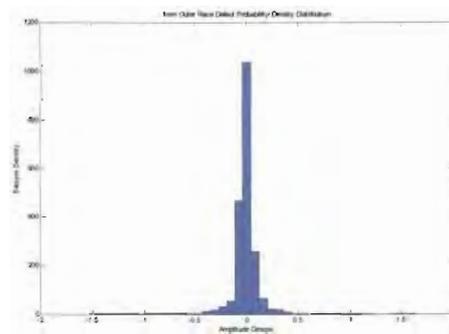


Figure 4.27: Stage Two Outer Race Defect Probability Density Distribution

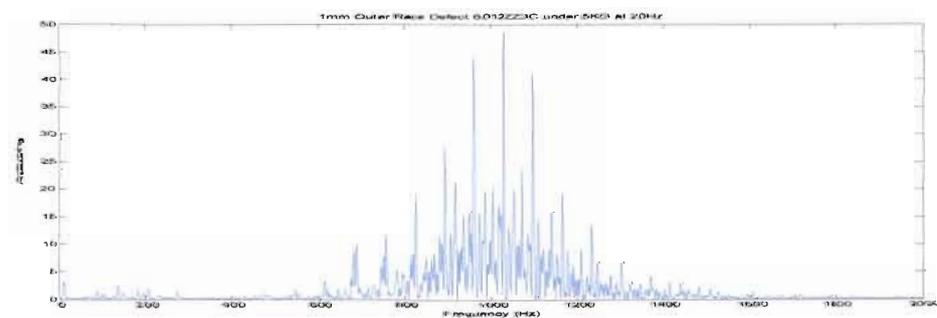


Figure 4.28: Stage Two Outer Race Defect Frequency Response

The frequencies (figure 4.28) are a combination of side bands from 700Hz to 1500Hz and have high amplitudes, and only a bearing defect is present.

4.15.3 Stage Three Bearing Defect

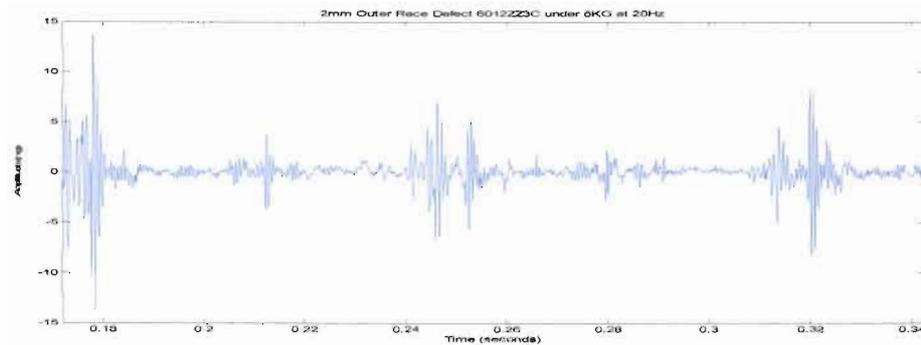


Figure 4.29: Stage Three Outer Race Defect Time Response

The pulses of a stage three defect were broader and uneven (figure 4.29). This indicates that the spall is larger and more time is spent as the ball element passes over the spall. The histogram indicates definite pulses of a bearing defect (figure 4.30). In the frequency domain (figure 4.31) the contact ball elements respond at 800Hz.

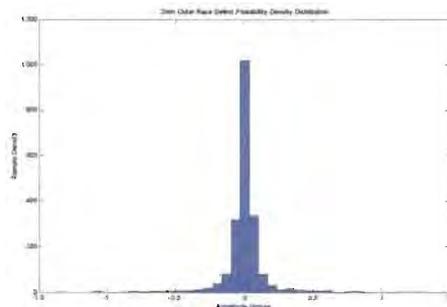


Figure 4.30: Stage Three Outer Race Defect Probability Density Distribution

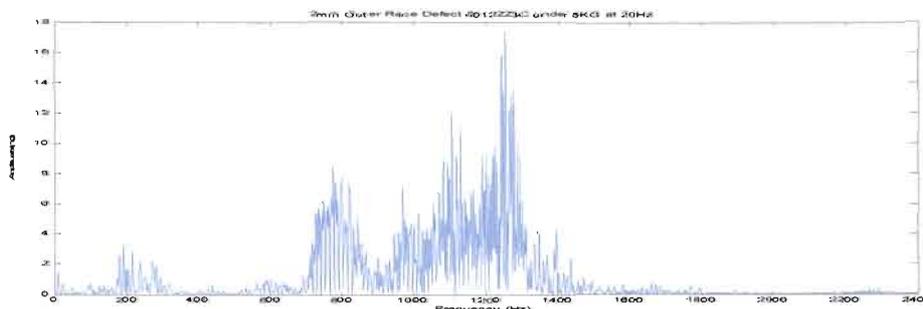


Figure 4.31: Stage Three Outer Race Defect Frequency Response

4.15.4 Stage Four Bearing Defect

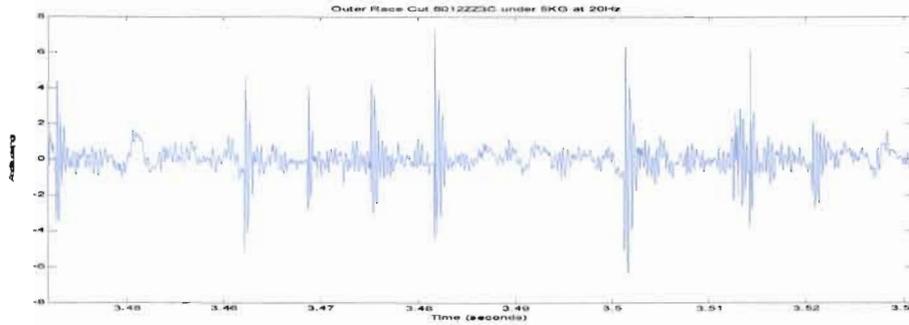


Figure 4.32: Stage Four Outer Race Defect Time Response

Here the pulse train is made up of a combination of sharp and broad pulses (figure 4.32). The histogram is steep but has a larger variance due to the combination of pulses (figure 4.33). This stage is characterised by BPFO harmonics in the frequency domain (figure 4.34).

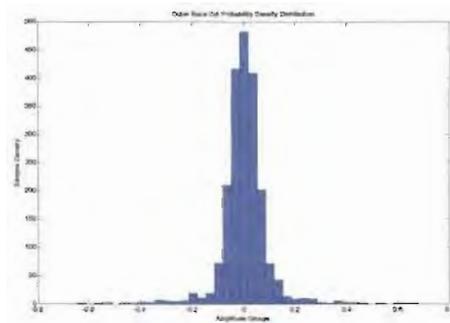


Figure 4.33: Stage Four Outer Race Defect Probability Density Distribution

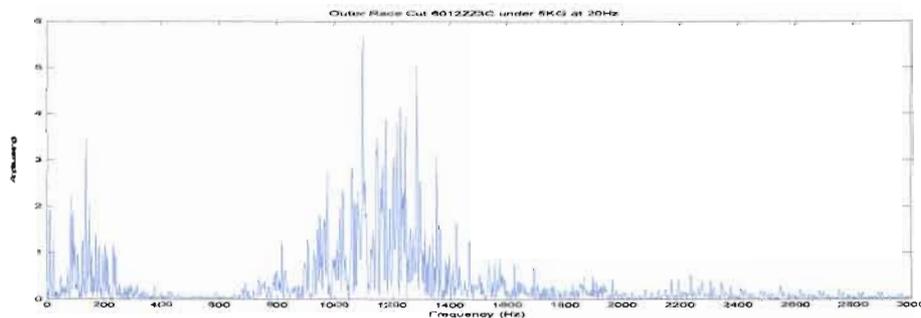


Figure 4.34: Stage Four Outer Race Defect Frequency Response

4.16 Summary

The explanation and relevance of Fourier theory enabled a better understanding of the time domain and frequency domain layouts. The statistical parameters used in the analysis were highlighted. A virtual instrument for signal processing was developed in LabVIEW. The relevance of filtering was discussed. Vibration data acquisition was performed. The natural responses of the monitoring rig and tested bearings were acquired through the impact hammer test. To highlight the difference in bearing defects and defect severity, the bearing defect vibration measurements were illustrated in the time domain and frequency domain. A data set of statistical parameters of different bearing defects can be created as an input data set for the neural networks.

Chapter 5

Application of Self-Organising Feature Maps

5.1 Introduction

Chapter 5 concentrates on the operational principles of Kohonen's Self-Organising Map neural network. Kohonen's Neural Network (KNN) may be referred to as a Self-Organising Map (SOM), or Self-Organising Feature Map (SOFM). The Self-Organising Feature Map enables good cluster visualisation of a data set and is the main reason why this network was chosen, i.e. to create cluster visualisation of bearing defects. The computations that lead to the visualisation of data clusters are governed by the learning algorithm of the Self-Organising Feature Map network. The Self-Organising Feature Map algorithm and architecture of the network are discussed in detail, creating a better understanding of how bearing defect data can be clustered and visualised for analysis.

Architecture is the layout of the components that make up a Self-Organising Feature Map. A comparison of the learning algorithms of unsupervised networks, such as Hebbian and Kohonen algorithms, is performed to enable better understanding of the competition concept that takes place between neurons. An explanation of the Self-Organising algorithm and its training procedure is given, followed by an explanation of data clustering and classification. Analysing the classified clusters enables the identification of bearing defects.

5.2 Self-Organising Feature Map Architecture

Demuth (2001:5) and Dorf (1993:420) described the architecture of a Self-Organising Feature Map neural network as having one output layer and no hidden layers. A layer consisting of elements, or neurons, or processing units, that are arranged in a geometric lattice such that each neuron has a set of neighbouring neurons. The output layer is in the form of a rectangular or hexagonal geometric lattice structure, where neighbouring neurons are connected by lateral connection weights w^L . Each neuron is connected to all input data values through input connection weights w^I . Each neuron has a competitive transfer function linked to the operations of all connection weights to provide an output for the output layer. The input vector \mathbf{p} is not a layer. There are S number of neurons in the output layer. Figure 5.1 shows the described architecture. Included in the figure is the winning neuron, which is created when the self-organising feature map learning algorithm is applied.

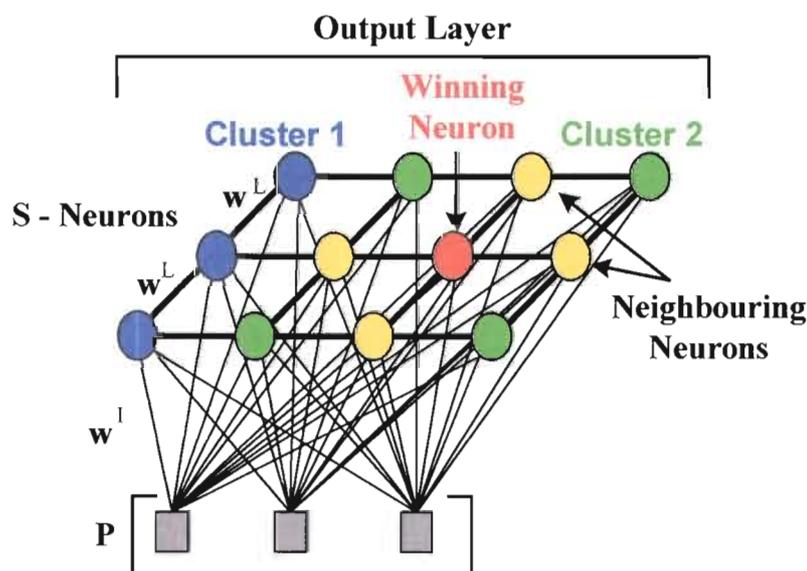


Figure 5.1: Symbolic Architecture of Kohonen's Self-organising Feature Map

The input data is numerical, which enables the network to perform learning and training algorithms. Learning and training in a neural network is the manipulation of the input data with other numerical vectors or matrices to achieve an output. The output of a Self-Organising Feature Map is characterised by the input, and does not have to be a desired output. In the case of Self-Organising Feature Maps, the output is a cluster mapping of the input data. An input column vector \mathbf{p} consist of a number of data values R . R , is the size or dimension of the input vector. The number of neurons S in an output layer and the size or dimension of the input vector R , are different. The input data may also be represented in a matrix form, where each column with R data values in the matrix represents a single input vector or a single considered feature. The number of input vectors Q depends on the number of features considered in the data set. An input data matrix is therefore an R -by- Q matrix. The input data matrix \mathbf{p} of input vectors can be shown as follows;

$$\mathbf{p} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,Q} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,Q} \\ \vdots & \vdots & \vdots & \vdots \\ p_{R,1} & p_{R,2} & \cdots & p_{R,Q} \end{bmatrix} \quad (5.1)$$

Each neuron, j , is part of all the neurons in the output layer, so that there are, $j = 1, 2, 3, \dots, S$ neurons. Each neuron, j , is connected to all the input data values in each input vector, through numerical input connection weights or input synaptic weights, w^j . The number of numeric connection weights for neuron j is therefore equal to the dimension R of each input vector. For one input vector, an input connection weight vector \mathbf{w}^j can be represented in a matrix where each column is a vector of input connection weights for neuron j .

The matrix of input connection weight vectors can be shown as follows;

$$\mathbf{w}^I = \begin{bmatrix} w_{1,j}^I & w_{1,2}^I & \dots & w_{1,S}^I \\ w_{2,j}^I & w_{2,2}^I & \dots & w_{2,S}^I \\ \vdots & \vdots & \vdots & \vdots \\ w_{R,j}^I & w_{R,2}^I & \dots & w_{R,S}^I \end{bmatrix} \quad (5.2)$$

Each input vector in the matrix of equation 5.1 has its own input connection weight matrix similar to the one in equation 5.2. The number of sets of input connection weight matrices is equal to the number of input vectors Q . So an input connection weight vector matrix \mathbf{w}^I may be shown in terms of the different input vectors as follows;

$$\mathbf{w}^I = [\mathbf{w}_1^I \quad \mathbf{w}_2^I \quad \dots \quad \mathbf{w}_Q^I] \quad (5.3)$$

Equation 5.3 is a complete representation of all the input connection weights in a Self-Organising Feature Map neural network according to the number of input features considered.

Willshaw and von der Malsburg (1976:431) referred to the geometric lattice structure of the output layer as its topology. The neurons in the output layer are arranged in a one or two-dimensional lattice structure, a topology that ensures that each neuron has a set of neighbouring neurons. The two-dimensional topology enables better analysis of large data sets than the one dimensional topology. A three, four or more dimensional topology can be created, however since analysis is performed on a two-dimensional surface, a two-dimensional output layer topology is used. With correct computer equipment visualisation and analysis can be done in two and three dimensions, which allows for the creation of a three dimensional topological map.

Two-dimensional topologies can be either rectangular or hexagonal, the hexagonal topology was chosen. The Kohonen topological model is used in this research for its data compression capabilities whereby the number of neurons S in an output layer is much less than the size or dimension R of the input vector.

The size of an input vector R determines the number of coordinates required to plot the vector. A Cartesian plane is a plane in Euclidian space of two coordinates. The size or dimension of an input vector R is equal to the coordinates in Euclidean space.

The input connection weights w^l are adjustable parameters of the neural network. Through training, a network algorithm adjusts the connection weight parameters. Neural networks whose desired output is known use supervised learning algorithms, and those whose output is unknown and depends on the input data, use unsupervised learning algorithms. The Self-Organising Feature Maps use an unsupervised learning algorithm.

5.3 Unsupervised Learning Algorithms

5.3.1 Hebbian Algorithm

Hebb (1949:5) developed the original unsupervised learning algorithm. Hebbian learning extracts correlated information from the input space and creates a topological mapping of the input where data neighbours in the input space are mapped to neuron neighbours in the output space.

For an, $n = 1, 2, \dots, R$, dimensional input vector x , connected to neurons through input connection weights w^l , the output y of neuron j is computed in the following manner;

$$y_j = a_j + \sum_{n=1}^R w_{jn}^l x_n \quad (5.4)$$

where a_j is a fixed bias value and w_{jn}^1 is the connection weight of neuron j to the n^{th} input.

5.3.2 Oja's Rule

Considering relevant constants, the change in connection weight for Hebbian learning as the weights vary with each input, results in a new weight $w_j(n+1)$ computed as follows:

$$w_j^1(n+1) = w_j^1(n) + \eta y_j(n) x_n(n) \quad (5.5)$$

where the variables are the same as for Hebbian learning and η is the learning rate parameter. However this learning rule leads to unlimited growth of the connection w_j . This problem is overcome by normalising the learning rule. Oja (1982:267) introduced the normalisation of the Hebbian learning rule, and is as follows:

$$w_j^1(n+1) = \frac{w_j^1(n) + \eta y_j(n) x_n(n)}{\sqrt{\sum_{n=1}^R (w_j^1(n) + \eta y_j(n) x_n(n))^2}} \quad (5.6)$$

5.3.3 Competitive Learning Algorithm

Haykin (1994:403) noted that in competitive learning, neuron j , can be the winning neuron, if the neuron's internal activity level v_j is the largest among all the neurons in the network. The output y_j , which is the sum of the input connection weights to the winning neuron j , is set equal to one. The output of all the neurons that lose the competition are set equal to zero. A neuron learns by shifting connection weights from inactive to active input nodes at the winning neuron. A neuron that does not respond to a particular input pattern, means that no learning takes place in that neuron. The winning neurons connection weights give up some amount of weight value which is equally distributed among the active input connection weights, that is

the winning neuron's connection weights, so that neuron j , the winning neuron, moves towards the input data pattern. Taking η as the learning rate parameter, the standard competitive learning algorithm defines the change in input connection weight as follows:

$$\Delta W_{jn}^I = \begin{cases} \eta(x_n - w_{jn}^I) & \text{neuron } j \text{ wins} \\ 0 & \text{neuron } j \text{ loses} \end{cases} \quad (5.7)$$

Grossberg (1976a:187) divided competition into two basic types: hard and soft. Hard competition means that only one neuron wins the input resources, as in the competitive algorithm. Soft competition means that there is a clear winning neuron, but its neighbours also share a small percentage of the inputs resources, as in Kohonen's algorithm.

5.3.4 The Learning Vector Quantisation Algorithm

Kangas *et al.* (1990:93) observed that the best results for pattern classification were achieved through the combined use of an unsupervised Kohonen Self-Organising Map with a supervised learning algorithm or neural network such as the Learning Vector Quantisation neural network (LVQ).

However, Haykin (1994:427) noted that the learning vector quantisation neural network (LVQ) is a supervised network, but has a non-linear competitive layer with a competitive transfer function, and a linear output layer with a linear transfer function. This linear transfer function enables the output to be placed in groups or classes.

With the input connection weight vector \mathbf{w}_j^I of neuron j identified to belong to some predefined class label, and the input connection weights w_{jn}^I also labelled to the predefined classes, then, according to Kangas *et al.* (1990:94), a supervised learning algorithm can be used to fine tune the connection weights. With an input vector \mathbf{x} ,

and for each iteration t , the learning vector quantisation algorithm can be written in a compressed form as:

$$\mathbf{w}_j^l(t+1) = \mathbf{w}_j^l(t) + \eta(t)s(t)\delta_{cj}(\mathbf{x} - \mathbf{w}_j^l(t)) \quad (5.8)$$

where $s(t) = +1$ if \mathbf{x} and \mathbf{w}_j^l belong to the same class,

But $s(t) = -1$ if \mathbf{x} and \mathbf{w}_j^l belong to different classes,

Here $\eta(t)$ is the learning rate factor, $0 < \eta(t) < 1$, and δ_{cj} is the class delta function ($\delta_{cj} = 1$ for $c = j$; and $\delta_{cj} = 0$ for $c \neq j$) where c is the defined class of the winning neuron. The neighbourhood set around the winning neuron is the winning neuron itself because of the class delta function.

5.3.5 The k-means Clustering Algorithm

In order to differentiate definite clusters the k-means clustering algorithm is used, as opposed to just classifying any particular cluster using the Self-Organising Map. MacQueen (1967:281) adopted the k-means clustering algorithm as a vector quantisation method. The algorithm minimises the sum of squared errors among a number of k clusters. The cluster centre coordinates are iteratively located. To estimate the validity of the k number of clusters, a validity criterion such as the Davies–Bouldin index, is used. Bouldin (1979:226) defined the Davies Bouldin index (DB) as a measure of the validity of the number of clusters obtained with the k-means method. This index is a function of the ratio of the sum of vector-cluster distance and centre-cluster distance, (equation 5.9), With e_{bh} as the average Euclidean distance of the vectors in class b to the centre vector of class h and d_{bh} as the distance between the centre vectors of clusters b and h :

$$DB = \frac{1}{k} \sum_{b=1}^k \max \left\{ \frac{e_{bh} + e_{hb}}{d_{bh}} \right\} \quad (5.9)$$

5.4 The Self-Organising Feature Map Algorithm

Kohonen (1982a:59) developed and modified the self-organising feature map algorithm. The self-organising feature map algorithm consists of three processes: competition, cooperation and adaptation.

5.4.1 Competitive Transfer Function

Let $x_1, x_2, x_3, \dots, x_R$ be the input data values of an, $n = 1, 2, \dots, R$ dimensional input vector \mathbf{x} . For an output layer with $j = 1, 2, 3, \dots, S$ neurons, the input connection weight vector \mathbf{w}^I , has $w_{j1}^I, w_{j2}^I, w_{j3}^I, \dots, w_{jR}^I$ as the corresponding input connection weights to neuron j . Let the lateral connection weight vector \mathbf{w}^L have $w_{j,-K}^L, \dots, w_{j,-1}^L, w_{j,0}^L, w_{j,1}^L, \dots, w_{j,K}^L$ as lateral connection weights to neuron j , where K is the radius of neighbouring neurons for lateral or layer interaction. Individual neuron outputs $y_1, y_2, y_3, \dots, y_S$ make up the output vector \mathbf{y} of the network, where S is the number of neurons in the output layer. Let the net input to neuron j , be I_j then;

$$I_j = \sum_{n=1}^R w_{jn}^I x_n \quad (5.10)$$

The output of neuron j at time or iteration t , may be expressed as follows;

$$y_j(t+1) = f\left(I_j + \beta \sum_{k=-K}^K w_{j,k}^L y_{j+k}(t)\right) \quad (5.11)$$

where $f(\cdot)$ is a non-linear function that limits the value of $y_j(t+1)$ and ensures that $y_j(t+1) \geq 0$. Thus, $y_j(t+1)$ is the output of neuron j at iteration $t+1$, and $y_{j+k}(t)$ is the output of neuron $(j+k)$ at the previous iteration t . β controls the rate of convergence of the iteration process. The non-linear function $f(\cdot)$ causes the output $y_j(t+1)$ to stabilise in a certain fashion depending on the convergence rate β . The outputs stabilise in a cluster where the initial output $y_j(0+1)$ is maximum.

5.4.2 Competition

Competition takes place between output layer neurons, to find the neuron that best matches the input vector \mathbf{x} . In accordance with the competitive transfer function, a neuron with a response similar to the stimulus I_j wins the competition and determines where a cluster is to be formed. The stimulus is the inner product $[\mathbf{w}_j^1]^T \mathbf{x}$ of the transposed input connection weights vector $[\mathbf{w}_j^1]^T$ and the input vector \mathbf{x} . The largest inner product $[\mathbf{w}_j^1]^T \mathbf{x}$, when comparing the inner products of all the neurons, shows where the winning neuron and the centre of a cluster is to be located.

The neuron with the shortest Euclidean distance is the winning or best-matching neuron and becomes the centre of a cluster. For computing positive Euclidean distance, the norm of the vector difference of input and input connection weights is shown as follows;

$$\begin{aligned} \|\mathbf{x} - \mathbf{w}_j^1\| &= \sqrt{(x_1 - w_{j1}^1)^2 + \cdots + (x_n - w_{jn}^1)^2} \\ &\vdots \\ \|\mathbf{x} - \mathbf{w}_s^1\| &= \sqrt{(x_1 - w_{s1}^1)^2 + \cdots + (x_n - w_{sn}^1)^2} \end{aligned} \quad (5.12)$$

5.4.2.1 Vector Quantisation

Euliano *et al.* (2000:315) showed that vector quantisation reduces the amount of data during the execution of the learning algorithm. A vector quantiser with minimum encoding distortion is a Voronoi quantiser. A reproduction vector defined and encoded by the Voronoi quantiser is a Voronoi vector and the region in input space represented by the Voronoi vector is a Voronoi cell. A Voronoi cell is created using the nearest-neighbour rule based on grouping data with a code that reduces the distortion measure. The Kohonen self-organising feature map is an approximate vector quantisation algorithm, since the output layer has neurons that act similar to Voronoi vectors by representing clusters of the input data set.

Consider an input data set of statistical parameters obtained from bearing vibration measurements. The program used to determine the statistical parameters is shown in annexure 6 (p142). The measurements were obtained from damaged 6008ZZ3C deep-groove ball bearings subjected to a static load of 4kg and a running speed of 15Hz. The statistical parameters investigated were: RA - rectified average, RMS - root mean square, Pk - peak value, Cf - crest factor, Skw - skewness, K - kurtosis and Std - standard deviation. The testing parameters were: Spd - speed and Load - Kg. Table 5.1 shows the input data set with five bearing vibration measurements taken for each defect condition of the bearing. Each statistical vector has a dimensional size, R, equal to 40. The label column indicates the bearing defect concerned.

RA	RMS	Pk	Cf	Skw	K	Spd	Kg	
0.0033076	0.0042144	0.017741	4.2097	-0.048539	3.364	0.9	4	newIa
0.003203	0.0041633	0.02668	6.4084	0.11518	4.794	0.9	4	newIa
0.0034115	0.0042945	0.017263	4.0198	-0.012241	3.1932	0.9	4	newIa
0.0033473	0.0043095	0.016835	3.9064	0.037124	3.3999	0.9	4	newIa
0.0033557	0.00434	0.021973	5.063	-0.031312	4.0867	0.9	4	newIa
0.021337	0.032589	0.3534	10.844	0.41138	16.33	0.9	4	BallcutIa
0.023885	0.03927	0.31283	7.9663	0.036974	13.257	0.9	4	BallcutIa
0.023107	0.031858	0.15913	4.9949	0.105	5.325	0.9	4	BallcutIa
0.021326	0.031123	0.20823	6.6905	-0.0024201	8.7968	0.9	4	BallcutIa
0.024716	0.039949	0.2672	6.6886	0.10674	13.932	0.9	4	BallcutIa
0.080745	0.13172	0.8345	6.3354	-0.0087818	10.187	0.9	4	Outer1mmIa
0.074258	0.10957	0.52632	4.8036	-0.065141	5.8923	0.9	4	Outer1mmIa
0.081118	0.13534	0.70434	5.2042	-0.072013	8.3707	0.9	4	Outer1mmIa
0.093437	0.14872	0.79124	5.3203	-0.021828	8.1492	0.9	4	Outer1mmIa
0.081343	0.12447	0.70632	5.6747	0.026185	7.8233	0.9	4	Outer1mmIa
0.14772	0.19959	0.81394	4.078	0.0044037	4.5498	0.9	4	InnerCrackIa
0.11734	0.15644	0.84745	5.4173	0.042108	4.8325	0.9	4	InnerCrackIa
0.12528	0.16982	0.73235	4.3124	0.041417	4.4982	0.9	4	InnerCrackIa
0.12628	0.17435	0.86378	4.9543	-0.054576	5.4489	0.9	4	InnerCrackIa
0.13729	0.1857	1.1084	5.9689	0.020651	6.422	0.9	4	InnerCrackIa
0.088943	0.15014	0.93155	6.2044	-0.022703	10.998	0.9	4	Outer2mmIa
0.10684	0.18252	1.0575	5.7937	0.1096	10.082	0.9	4	Outer2mmIa
0.10202	0.17887	1.1447	6.3998	0.16724	12.716	0.9	4	Outer2mmIa
0.08584	0.1584	1.0188	6.4318	0.1301	15.094	0.9	4	Outer2mmIa
0.1047	0.18294	1.1308	6.181	0.12176	12.881	0.9	4	Outer2mmIa
0.033446	0.058388	0.64237	11.002	-0.3496	31.173	0.9	4	Inner1mmIa
0.030132	0.044518	0.41866	9.4045	-0.35882	19.311	0.9	4	Inner1mmIa
0.026551	0.034961	0.1965	5.6206	-0.17199	4.4918	0.9	4	Inner1mmIa
0.031549	0.047621	0.42024	8.8247	-0.17814	15.789	0.9	4	Inner1mmIa
0.028083	0.041529	0.37143	8.944	-0.016777	14.39	0.9	4	Inner1mmIa
0.043322	0.067752	0.42186	6.2265	-0.13226	11.491	0.9	4	OuterCrackIa
0.037536	0.06009	0.4065	6.7648	-0.16357	12.948	0.9	4	OuterCrackIa
0.042541	0.06614	0.48427	7.3219	-0.20626	11.612	0.9	4	OuterCrackIa
0.039268	0.061238	0.42741	6.9795	-0.038607	12.956	0.9	4	OuterCrackIa
0.039657	0.063378	0.40375	6.3705	-0.012928	12.689	0.9	4	OuterCrackIa
0.025859	0.040126	0.2697	6.7212	0.32564	10.886	0.9	4	Inner2mmIa
0.033583	0.054775	0.28969	5.2888	0.038474	8.2877	0.9	4	Inner2mmIa
0.025439	0.037356	0.22586	6.0461	-0.20454	7.3008	0.9	4	Inner2mmIa
0.030203	0.048433	0.43286	8.9372	-0.041369	19.822	0.9	4	Inner2mmIa
0.034608	0.059899	0.53305	8.8992	0.46059	16.497	0.9	4	Inner2mmIa

Table 5.1: Input Data Set of a 6008ZZ3C Deep-Groove Ball Bearing

For a three-dimensional input vector $\mathbf{x} = (1, -2, 2)$ and two output neurons A and B. The input connection weight vector of neuron A was $\mathbf{w}_A^1 = (2, -1, 3)$ and that of neuron B was $\mathbf{w}_B^1 = (-2, 0, 1)$. For a learning rate coefficient $\eta = 0.5$ and a neighbourhood function $h_{ij}(d_{ij}) = 1$. The difference in Euclidean distance:

$$\begin{aligned}\|\mathbf{x} - \mathbf{w}_A^1\| &= \sqrt{(1-2)^2 + (-2+1)^2 + (2-3)^2} = \sqrt{3} \\ \|\mathbf{x} - \mathbf{w}_B^1\| &= \sqrt{(1+2)^2 + (-2-0)^2 + (2-1)^2} = \sqrt{14}\end{aligned}$$

gives the winning neuron as, neuron A because $3^{1/2} < 14^{1/2}$. Now from table 5.1, the input vector RA has a dimension of 40. The map to be trained is a [6, 5] hexagonal map, this means there are 30 output neurons in the topology. Each neuron has an initialised input connection weight to input vector RA. The winning neuron is determined as in the example above with constants $\eta = 0.5$ and $h_{ij}(d_{ij}) = 1$.

5.4.3 Cooperation

Letting the winning neuron be neuron i , then the amount of help to neuron j is calculated using the neighbourhood function $h_{ij}(d_{ij})$, where d_{ij} is the distance between neuron i and neuron j in the output layer. The winning neuron also helps itself more than others because $d_{ij} = 0$. The topology defines which neurons in the two-dimensional layer are neighbours. The first neighbouring neurons including the winner neuron are located within the first neighbourhood radius K_1 , and are immediately connected to the winner neuron by lateral connection weights $w_{j,K}^L$. Doubling the neighbourhood radius increases the number of neighbouring neurons. The second neighbouring neurons are immediately connected to the first neighbouring neurons and all these neurons are within the second neighbourhood radius. A hexagonal topology was used, so the neighbourhood function $h_{ij}(d_{ij})$ takes the form of the hexagonal topology.

This way, all the neighbourhoods are represented by the neighbourhood function $h_{ij}(d_{ij})$. Since $h_{ij}(d_{ij})$ is a function of the distances between neurons d_{ij} , and the hexagonal topology was used, the neighbourhood function applied was that of a Gaussian function. Taking K_t as the neighbourhood radius at iteration t , the Gaussian neighbourhood function is represented as follows;

$$h_{ij}(d_{ij}) = \left(K_t \sqrt{2\pi}\right)^{-1} e^{-\frac{d_{ij}^2}{2K_t^2}} \quad (5.13)$$

The neighbourhood function begins by considering all the neurons in the output layer, and then correlates the directions of the weight adjustments according to the winning neurons of different clusters. The neighbourhood function then shrinks with each iteration, because of the decreasing number of neurons whose direction of weight adjustments are correlated. So there is lateral feedback between neurons with each iteration.

5.4.4 Adaptation

After a winning neuron has been located, the weights of the network are adjusted so that a better representation of the input data is established. The change in layer connection weight vector \mathbf{w}^L is dependent on the neighbourhood function $h_{ij}(d_{ij})$ through cooperation. Also, the change in input connection weight vector \mathbf{w}^I is dependent on the initial output y and the rate at which the weights change η . Then the change in input connection weights for neuron j , after receiving input vector \mathbf{x} , can be shown as follows:

$$\frac{d \mathbf{w}_j^I}{dt} = \eta y_j \mathbf{x} - h_{ij}(d_{ij}) \mathbf{w}_j^I \quad (5.14)$$

For an input vector \mathbf{x} that changes at a rate that is extremely slow compared to the input connection weight vector \mathbf{w}^I for all neurons, a justifiable assumption is that

due to the clustering effect, the output y_j of neuron j is either at low or high saturation depending on whether neuron j is in the first, second or third neighbourhood of the winning neuron. A low saturation means neuron j is in a neighbourhood far away from the winning neuron and the output y_j is allocated a null value. A high saturation means neuron j is in a neighbourhood close to the winning neuron and the output y_j is allocated a value of one.

In terms of iterations t , the input connection weight vector $\mathbf{w}_j^I(t)$ for neuron j at iteration t , is combined with the change in the input connection weight vector $\Delta\mathbf{w}_j^I(t+1)$ after iteration $(t+1)$ to give an updated input connection weight vector $\mathbf{w}_j^I(t+1)$. The learning rate function $\eta(t)$ can be linear, inverted or quadratic, but the form is not critical. A linear learning rate function was used with η_0 as the initial learning rate, and the total number of iterations performed as T . After iteration t , the linear learning rate function $\eta(t)$ is defined as follows:

$$\eta(t) = \eta_0 \left(1 - \frac{t}{T} \right) \quad (5.15)$$

With the learning rate coefficient defined by $\eta(t)$ and the neighbourhood function $h_{ij}(d_{ij})$ as a function of d_{ij} , the updated input connection weight vector after iteration t , can be found as follows:

$$\mathbf{w}_j^I(t+1) = \mathbf{w}_j^I(t) + \eta(t) h_{ij}(d_{ij})(\mathbf{x} - \mathbf{w}_j^I(t)) \quad (5.16)$$

To adapt neurons A and neurons B for the first iteration, in the example considered, equation 5.16 was used with constants $\eta = 0.5$ and $h_{ij}(d_{ij}) = 1$.

$$\mathbf{w}_A^1(1) = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + 0.5 \cdot 1 \cdot \left[\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} \right]$$

$$\mathbf{w}_A^1(1) = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + 0.5 \cdot \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1.5 \\ -1.5 \\ 2.5 \end{pmatrix}$$

The input connection weight vector for neuron A is adapted to $\mathbf{w}_A^1(1) = (1.5, -1.5, 2.5)$ which is closer, in terms of Euclidean distance, to the input vector $\mathbf{x} = (1, -2, 2)$ in comparison to the initial connection weight vector of $\mathbf{w}_A^1(0) = (2, -1, 3)$.

5.5 Training Procedure

Kohonen (1996:11) described two methods of training a network: (1) incremental sequential, nonparametric training, (2) batch, parametric training. An unsupervised network such as a self-organizing feature map depends on the layout of input data in the 'R' dimensional space representing the input data, and thus the training is guided by this layout.

In incremental training, the input connection weights are adjusted every time an input is introduced to the network, or every time an output is determined. Incremental training works well for supervised networks, where the target output is compared to a simulated output, every time an output is simulated. This is a continuous procedure in training until the difference between the simulated output and the targeted output is at a minimum.

Batch training involves the adjustment of weights when all the vectors in the input matrix have all been introduced to the network. In supervised networks the outputs of each vector input are still compared to the target output.

The differences for each vector input to its target output are combined for all the input vectors and an averaged difference is determined. This averaged difference determines the amount by which the weights are adjusted. Batch training works well for unsupervised networks with large input data sets. In batch training for unsupervised networks, the amount of change in input connection weight value Δw_j^l for all winning neurons and neighbouring neurons are combined, and an average change in input connection weight value $\text{Avg}(\Delta w_j^l)$ is determined, and used to adjust all the weights. Batch training is the fastest of the two training methods.

The training procedure for both types of training methods comprises five steps:

- The input connection weights are initialised to arbitrary linear or random values.
- The input data is then presented to the self-organising feature map.
- The winning neurons of different clusters are established through competition.
- The connection weights of all winning neurons and their neighbours are adjusted.
- The input data is repeatedly presented until the network stabilises.

5.5.1 Normalisation

Normalisation is input data pre-processing. Normalisation reduces input data values that are much larger than or much smaller than the average input data value. An input space will present very large input data values or very small input data values far away from other data values. These values are outliers from the majority of values in input space. To reduce the number of outliers, normalisation is performed.

A way to normalise data is to find the difference between a data value and the mean of the data set, and then scale each data value to unit standard deviation as follows:

$$\text{Input Data Value} = \frac{[\text{Data Value} - \text{Mean (Data Set)}]}{\text{Standard Deviation (Data Set)}} \quad (5.17)$$

The input data set in table 5.1 was normalised in this manner.

5.5.2 Initialisation

Random initialisation sets input connection weights that are uniformly distributed in the input range between minimum and maximum input data values.

Linear initialisation is done by first calculating the eigenvalues and eigenvectors of the an input data vector \mathbf{x} . Then the input connection weights are set to the middle of the range between the maximum and minimum eigenvectors. The input connection weights may also be set to the middle of the input range between the maximum and minimum input data values.

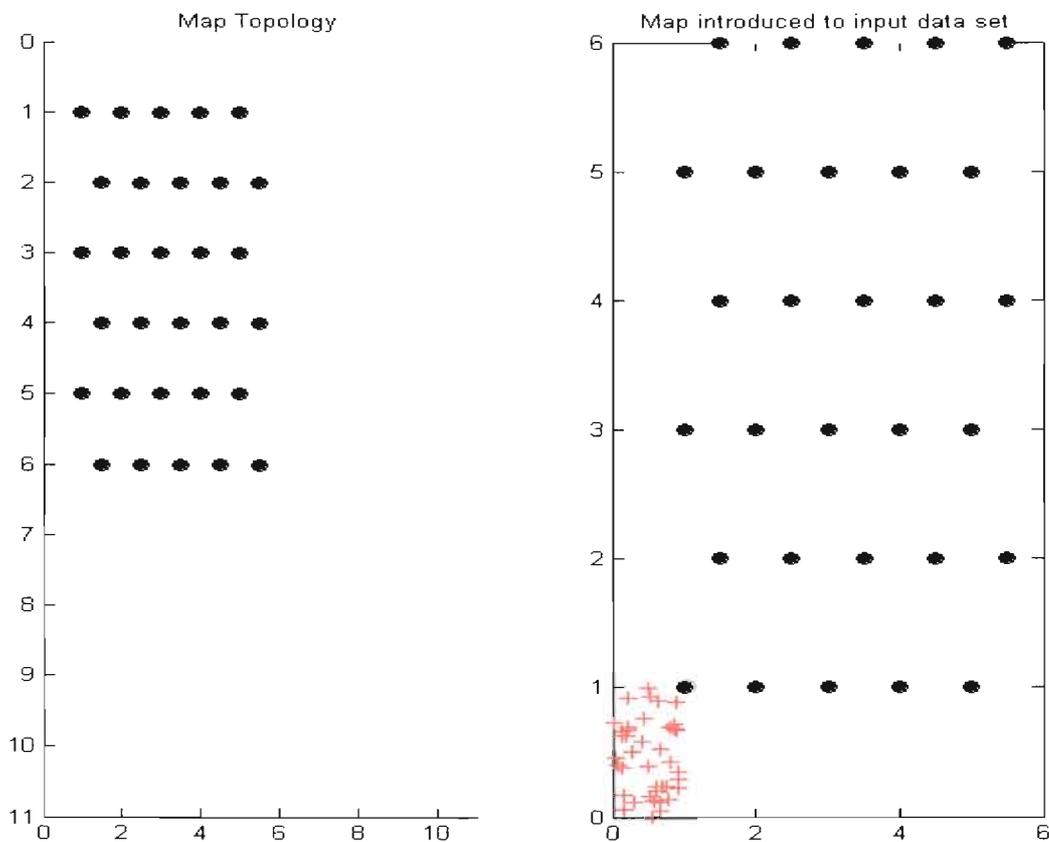


Figure 5.2: Map Topology

For the input data set in table 5.1, the $[6, 5]$ topological map is shown on the left in figure 5.2. The map on the right has been initialised linearly, to the input data set shown by 40 red crosses.

In initialisation, the input data set is introduced to the map. The location of the input data set is controlled by the initial input connection weights. Batch training was used. The stabilised trained map can be seen in figure 5.3. The 40 black crosses are the location of each input element in Euclidean space for all 8 vectors considered.

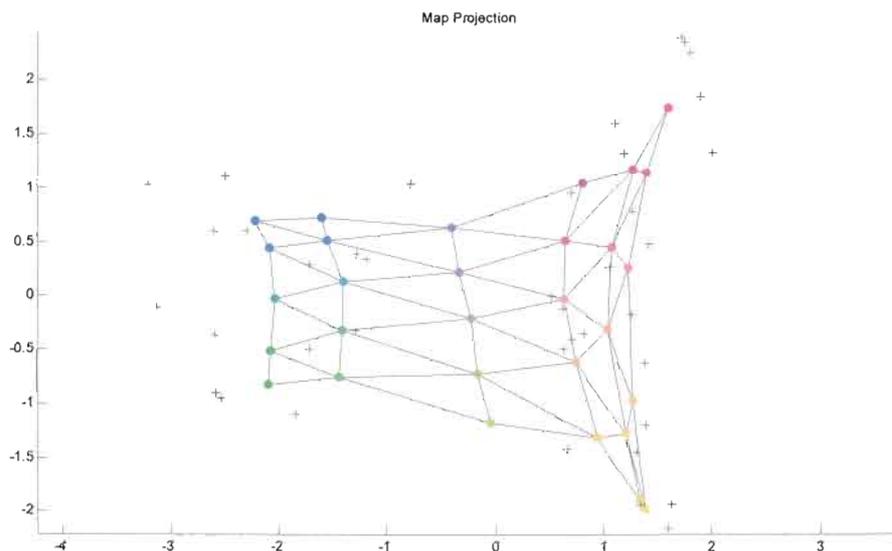


Figure 5.3: Trained Map in Euclidean Space

The 9th vector represents data labels. Each of the 40 input elements has a label according to a bearing defect. The colours on each neuron in figure 5.3 is a classification used to differentiate data clusters. Each cluster is labelled with a bearing condition or defect as can be seen in figure 5.4. These labels follow the layout of the trained map. Blue clusters of neurons represent an outer ring 2mm spall. The purple cluster neurons represent an inner race crack, the green cluster an inner ring 1mm spall and the yellow cluster a new bearing. The tones in colour of the other neurons represent the additional bearing defects considered.

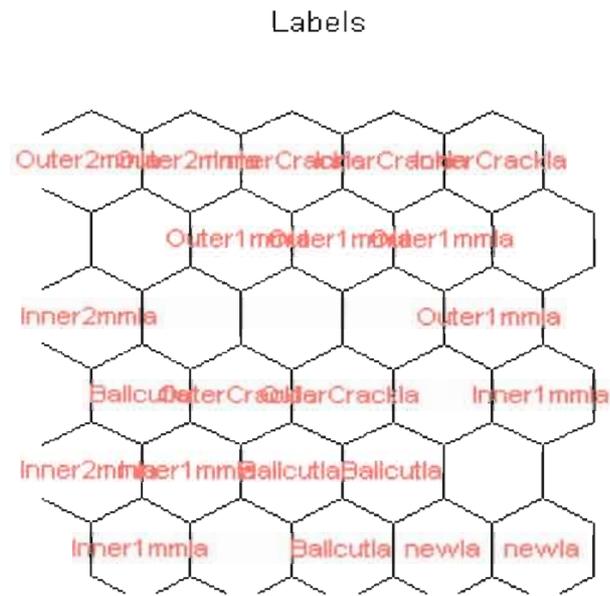


Figure 5.4: Labelled Classification of Trained Map

5.5.3 Adaptive Adjustments

Vesanto *et al.* (2000:26) performed the neural network training procedure in two phases. The first phase is the ordering phase and the second phase is the tuning or convergence phase.

In the ordering phase, the neighbourhood function $h_{ij}(d_{ij})$ includes all the neurons in the output layer, and the distance d_{ij} between the winning neuron and nearest neighbouring neuron j is at a maximum. Through weight adjustments, the distance and number of neurons included in the neighbourhood function decreases. The initial learning rate coefficient η_0 is close to unity, and there after the linear learning rate function $\eta(t)$ controls the learning rate coefficient η , keeping the rate above 0.1.

In the fine tuning phase, the neighbourhood function $h_{ij}(d_{ij})$ includes only the nearest neighbours of the winning neuron i , which may eventually be 1 or 0 neurons. The neighbourhood function ensures that the output layer is not under-utilised in representing the input data. The learning rate coefficient η , is below 0.1 and in the range of 0.01 or less for many iterations, enabling a good accuracy to be attained.

5.5.3.1 Quantisation Error

With the correct number of architectural parameters, training of a self-organising feature map continues until an accuracy limit is reached. Otherwise, architectural parameters, such as the number of neurons, may be changed. The accuracy limits used are the quantisation error and topological error. Villmann *et al.* (1997:4) identified the quantisation error, as the percentage error of the Euclidean distance between the connection weights of only the winning neuron and the represented input data vector. The quality of the self-organising feature map is determined by the mean quantisation error, QE, which is computed as follows:

$$QE = \frac{\sum_1^Q \left[\frac{\sum_{r=1}^R h_{ij}(d_{ij}) \|w_r^1 - p_r\|^2}{R} \right]}{Q} \quad (5.18)$$

Where $h_{ij}(d_{ij})$ is the Gaussian neighbourhood function, w_r^1 is the input connection weight to data r and p_r is input data r . R is the number of elements in each data vector, and Q is the number of different data vectors used in the self-organising map. The lower the value of QE, the more closely the winning connection weight vector matches the input data vector. QE can be analysed in terms of fraction or percentage. The quantisation error is determined after every epoch, and kept under the error limit.

5.5.3.2 Topological error

Topological error is the percentage difference in the shape of the hexagonal output layer before and after the map is trained. Before training the shape is fixed, but after training the shape is irregular. Villmann *et al.* (1997:261) proposed the topographic function Φ_A^m , which evaluates the topology preservation of the mapping in self-organising feature maps. The original untrained map is, m , and the trained map is A .

The neighbourhood preservation of the mappings $\Psi_{m \rightarrow A}$ and $\Psi_{A \rightarrow m}$ are denoted by $f_j(w_n^l)$ and $f_j(-w_n^l)$; respectively, with j being the index of the neuron in the map and $n = 1; \dots; R$. The topographic function Φ_A^m of map M_A is then defined by:

$$\Phi_A^m(w_n^l) = \begin{cases} \frac{1}{R} \sum_{j \in A} f_j(w_n^l) & w_n^l < 0 \\ \Phi_A^m(1) + \Phi_A^m(-1) & w_n^l = 0 \\ \frac{1}{R} \sum_{j \in A} f_j(w_n^l) & w_n^l > 0 \end{cases} \quad (5.19)$$

The topographic function $\Phi_A^m(0) = 0$ if and only if the map topology at neuron j is perfectly preserved. The trained map in figure 5.3 had a quantisation error of 0.783 and a zero topological error, which means the topology of the untrained map was preserved.

5.6 Parameter Classification

The Self-Organising Feature Map is a visualisation neural network for data clusters. This is illustrated for all the input vectors, that is all the statistical parameters considered, in figure 5.5. The parameter maps are also classified with colour coding and defect labels. The first matrix is the U-matrix or unit matrix, which indicates the Euclidean distances between neighbouring neurons and clusters. The parameter matrices indicate the different representation of the statistical value allocated to a cluster or neuron. This enables visualisation of the distribution of the statistical values over the trained map. The speed matrix is constant, the dark clusters indicate neurons that do not represent a speed value. In the load matrix all the neurons represent the load. The label matrix classifies each neuron and each cluster according to bearing defect, for all the parameter maps. The location of a neuron is the same for all matrices displayed.

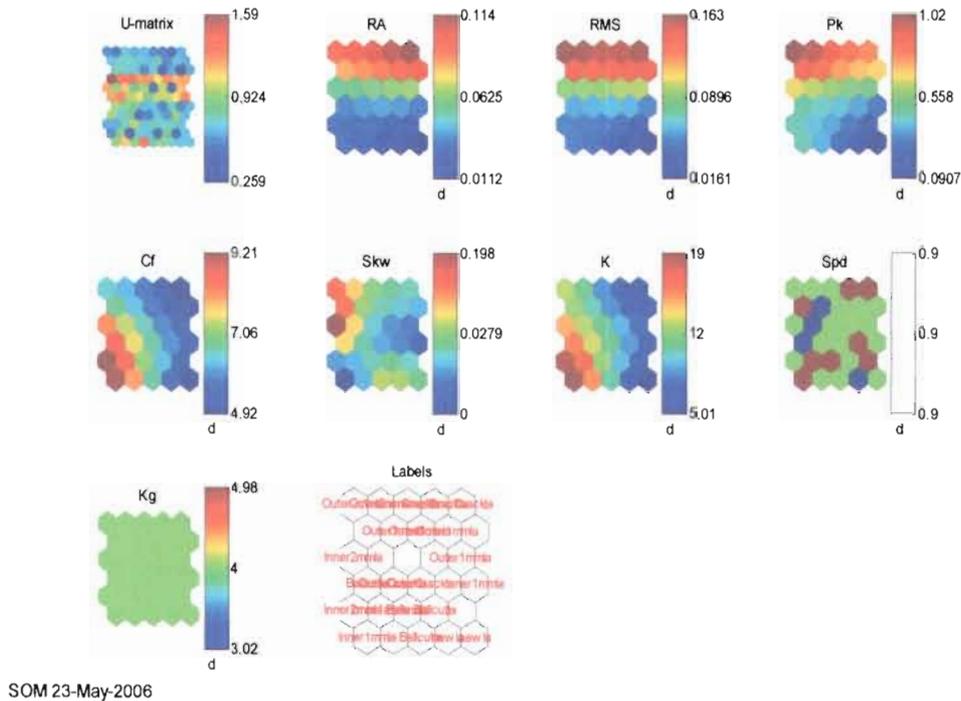


Figure 5.5: Parameter Classification Maps

5.7 Summary

The neuron model and the architecture of the Self-Organising Feature Map were discussed in detail. The algorithms, with which the Self-Organising Feature Map approximates vector quantisation by reducing data dimension in the output layer, were discussed giving an example with one input data set. Large data sets and missing data in the data set are not problematic because of vector quantisation and normalisation. Cluster ordering and visualisation, can be achieved through unsupervised learning, enabling good analysis of the data.

Chapter 6

Experimentation

6.1 Introduction

Different size input data sets are created to compare the performance of the Self-Organising Map to different sized sets of input data. At the same time, the visualisation capabilities of the Self-Organising Map are analysed. A program that automatically detects a bearing defect from one vibration measurement was created.

Defect vibration measurements were taken and a data set of feature parameters of the input data was formed. The data set was large enough to give an option of applying a 400, 2000 or 10000 dimensional data set to Kohonen's self-organising map. Classification of rolling element bearing defects on a Self-Organising Map enabled a guiding prognosis on the type and extent of bearing damage. A detailed analysis of each matrix was performed with reference to cluster representation of statistical parameters.

6.2 Creation of an Inclusive Parameters Input Data Set

The condition monitoring testing conditions were limited to operational bearing loads of five different weights and five different shaft speeds. There were eight bearing defect conditions considered. Fifty vibration measurements of the same bearing defect were taken, under the same shaft speed and bearing load conditions. With fifty vibration measurements, a distinct cluster of the defect can be represented by many neurons, providing better detection of a defect.

For single speed and loading conditions, 400 bearing vibration measurements were taken. Seven statistical parameters were determined from each measurement. These were the: RA - rectified average, RMS - root mean square, Pk - peak, Cf - crest factor, Skw - skewness, K - kurtosis and Std - standard deviation. The program that created the input data through determining statistical parameters is shown in annexure 6 (p142). However, this is a shortened form of the program as it determines the statistical feature parameters for one measurement and the complete program determines the parameters for 10000 measurements.

Combining the statistical parameters determined for all five shaft speeds and five loading conditions considered in the monitoring tests, a total input data set from 10000 vibration measurements was created. The input data set is shown in table 6.1. For each bearing size, a 10000 dimensional input data set of feature parameters was created. Three different bearing sizes were used for testing.

The input data set in table 6.1 gives the format of the input data set, and is not the data set that was used to train a Self-Organising Map. The input data set used was too large to be included in the dissertation.

Testing parameters included in the input data set, were speed in revolutions per minute and load in kilograms. Each parameter is an input vector, therefore there were 10 input vectors including the labels vector. The pattern used to label statistical values that were determined under the same load and speed conditions, was the same pattern used for all the considered load and speed testing conditions. The labels could have been numerical. For densely clustered maps, such as for the 10000 dimensional input data set, the classification label will also be dense and it will not be clear whether the labelling is numerical or alphabetical. Although zooming may be applied, only a display is provided here.

Table 6.1: Statistical Parameters Input Training Data

Row	RA	RMS	Pk	Cf	Skw	K	Std	Spd	Load	Defect
1 - 50	T	T	T	T	T	T	T	T	T	Tested
1	0.5	0.6	0.2	3.6	0.07	3.2	0.6	15Hz	4KG	New1
:	:	:	:	:	:	:	:	15Hz	4KG	:
50	0.5	0.6	0.3	4.0	0.1	3.4	0.6	15Hz	4KG	New50
51	0.2	1	0.6	6	0.08	12	1	15Hz	4KG	Ball 1
:	:	:	:	:	:	:	:	15Hz	4KG	:
100	0.3	2	0.8	7.8	0.2	18	2	15Hz	4KG	Ball 50
101	0.6	2.1	1.1	9.1	0.7	22	2	15Hz	4KG	1mm Outer 1
:	:	:	:	:	:	:	:	15Hz	4KG	:
150	0.7	3.2	0.9	6.6	0.8	30	3	15Hz	4KG	1mm Outer50
151	0.4	0.9	0.8	9.3	-0.1	28	0.9	15Hz	4KG	Inner cut 1
:	:	:	:	:	:	:	:	15Hz	4KG	:
200	0.5	0.6	0.7	8	-0.3	26	0.5	15Hz	4KG	Inner cut 50
201	0.7	0.2	1.3	6.6	-0.4	40	0.2	15Hz	4KG	2mm Outer 1
:	:	:	:	:	:	:	:	15Hz	4KG	:
250	0.8	0.4	1.5	7	-0.6	35	0.3	15Hz	4KG	2mm Outer
251	0.4	0.7	0.8	9.2	-0.7	19	0.7	15Hz	4KG	50
:	:	:	:	:	:	:	:	15Hz	4KG	1mm Inner 1
300	0.5	0.8	1.1	8.8	-0.8	21	0.8	15Hz	4KG	:
301	0.3	0.5	0.5	6.5	-0.2	17	0.6	15Hz	4KG	1mm Inner 50
:	:	:	:	:	:	:	:	15Hz	4KG	Outer cut 1
350	0.4	0.7	0.5	7	-0.1	15	0.7	15Hz	4KG	:
351	0.3	0.1	1.3	9	0.2	25	0.1	15Hz	4KG	Outer cut 50
:	:	:	:	:	:	:	:	15Hz	4KG	2mm Inner 1
400	0.4	0.2	1.6	7	0.4	17	0.2	15Hz	4KG	:
										2mm Inner 50
800	"	"	"	"	"	"	"	20Hz	4KG	"
1200	"	"	"	"	"	"	"	25Hz	4KG	"
1600	"	"	"	"	"	"	"	30Hz	4KG	"
2000	"	"	"	"	"	"	"	40Hz	4KG	"
4000	"	"	"	"	"	"	"	"	5KG	"
6000	"	"	"	"	"	"	"	"	10KG	"
8000	"	"	"	"	"	"	"	"	15KG	"

6.3 Compared Error Performances

The Self-Organising Map was trained with input data sets of different dimensions, and its performances were recorded. This was done in order to recommend specific dimensions of input data sets so that the best analysis and visualisation can be obtained from a Self-Organising Map. Eleven input data sets of different dimensions were used, and are given in table 6.2. The topological map sizes, the errors obtained, and the training times are given for all three bearings tested. The results of bearing 6008ZZ3C, 6010ZZ3C and bearing 6012ZZ3C are in table 6.3, 6.4, and 6.5, respectively.

Table 6.2: Dimension Size of Input Data Sets

Input Data Set	1	2	3	4	5	6	7	8	9	10	11
Dimension	40	100	400	800	1200	1600	2000	4000	6000	8000	10000

Table 6.3: Training Errors for Bearing 6008 Input Data Sets

Input Data Set	Matrix Map Size	Quantisation Error	Topological Error	Training Time (s)
1	[6, 5]	0.783	0	0
2	[12, 4]	0.505	0	0
3	[13, 8]	0.483	0.052	0
4	[14, 10]	0.6	0.026	0
5	[15, 11]	0.653	0.027	0
6	[16, 12]	0.458	0.054	1
7	[17, 13]	0.449	0.046	1
8	[22, 15]	0.569	0.067	2
9	[24, 16]	0.532	0.079	2
10	[26, 17]	0.543	0.048	4
11	[27, 19]	0.557	0.065	5

Figure 6.1 illustrates that for large data sets, data set 6 up to data set 11, the quantisation error is between 0.449 and 0.557 as compared to 0.783 - 0.6 for the smaller data sets. Therefore, the topological error is relatively low for small data sets.

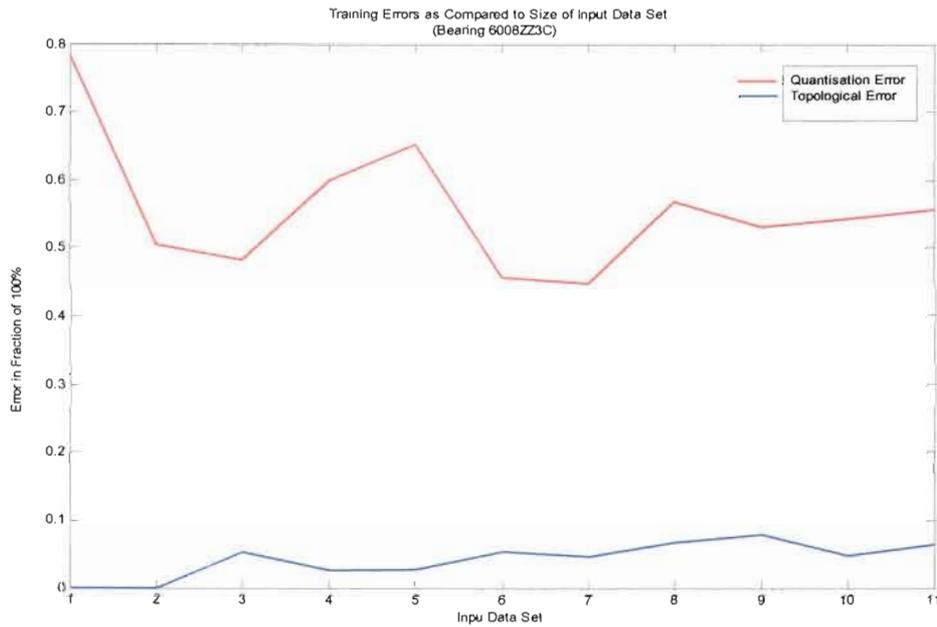


Figure 6.1: Bearing 6008ZZ3C Error vs. Dimension

Table 6.4: Training Errors for Bearing 6010 Input Data Sets

Input Data Set	Matrix Map Size	Quantisation Error	Topological Error	Training Time (s)
1	[8 , 4]	0. 714	0. 05	0
2	[10, 5]	0. 628	0. 01	0
3	[13 , 8]	0. 482	0. 028	0
4	[16 , 9]	0. 591	0. 04	0
5	[17 , 10]	0. 641	0. 048	1
6	[18 , 11]	0. 632	0. 051	1
7	[19 , 12]	0. 622	0. 058	1
8	[23 , 14]	0. 726	0. 05	2
9	[24 , 16]	0. 659	0. 05	2
10	[26 , 17]	0. 683	0. 087	4
11	[28 , 18]	0. 670	0. 062	5

The errors for Bearing 6010 are illustrated in figure 6.2. The statistical values for bearing 6010ZZ3C were slightly larger than those for bearing 6008ZZ3C. For large data sets, data set 6 up to data set 11, the quantisation error is between 0.622 and 0.726 which is high in comparison to that for bearing 6008ZZ3C. The topological error is relatively low and lowest for a dimension of 100.

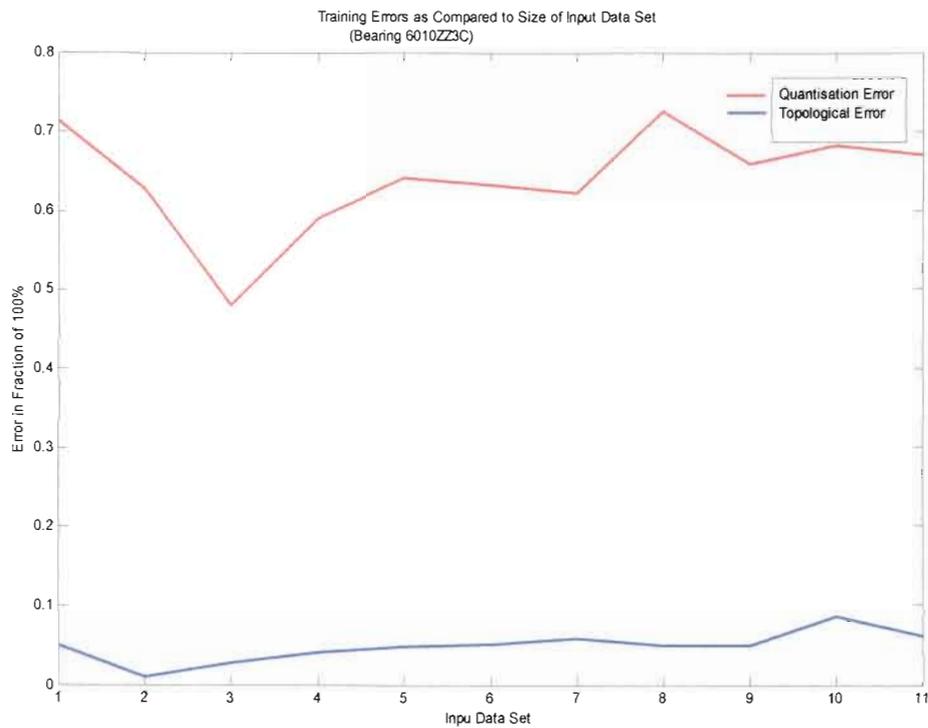


Figure 6.2: Bearing 6010ZZ3C Error vs. Dimension

Table 6.5: Training Errors for Bearing 6012 Input Data Sets

Input Data Set	Matrix Map Size	Quantisation Error	Topological Error	Training Time (s)
1	[8 , 4]	0. 817	0. 05	0
2	[10 , 5]	0. 551	0	0
3	[14 , 7]	0. 593	0. 03	0
4	[16 , 9]	0. 712	0. 04	0
5	[17 , 10]	0. 739	0. 043	1
6	[18 , 11]	0. 735	0. 039	1
7	[19 , 12]	0. 701	0. 035	1
8	[23 , 14]	0. 843	0. 056	2
9	[24 , 16]	0. 573	0. 05	2
10	[26 , 17]	0. 593	0. 059	4
11	[30 , 17]	0. 523	0. 075	5

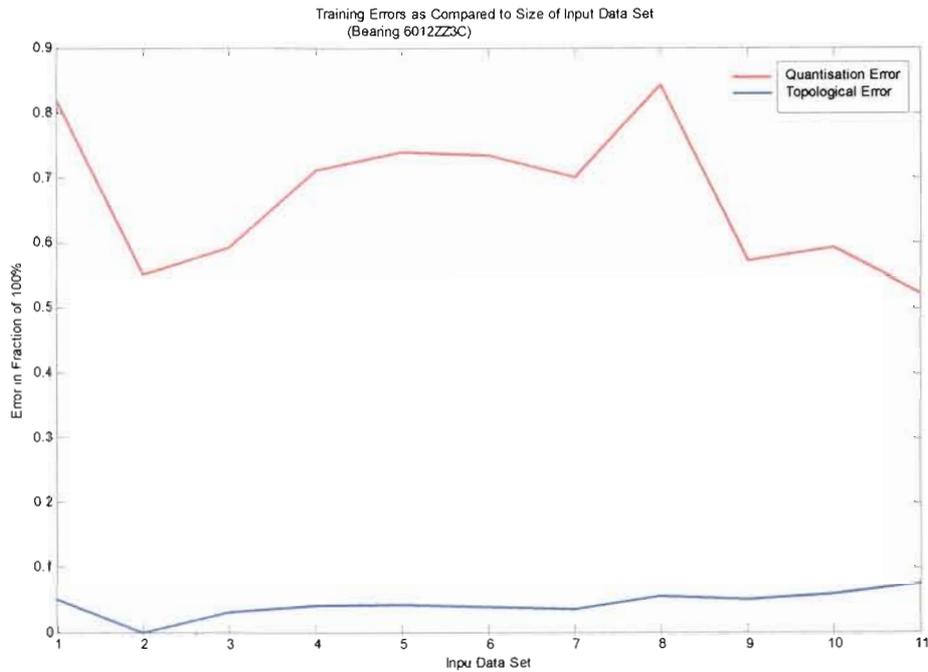


Figure 6.3: Bearing 6012ZZ3C Error vs. Dimension

The statistical values for bearing 6010ZZ3C and bearing 6012ZZ3C were similar, which is the reason why the error patterns are similar. The topological error is relatively low and zero for a dimension of 100. In data set 2 only two defect types are included, whereas in data set 3 all defects are included and it has a low topological error and moderately low quantisation error compared to the others for all bearings considered. Therefore, this all inclusive data set (which is data set 11 with a dimension of 10000) is the recommended data set. It takes the longest to train, and can have a moderate quantisation error and one the highest topological errors. However, for its inclusiveness of all the testing and defect conditions, data set 11 is recommended as the first training data set for any analysis of a bearing defect.

6.4 Kohonen's Feature Map in Parameters Input Data

The input data set of the 6012ZZ3C deep-groove ball bearing were used in this section. A self-organising feature map program was developed, (annexure 7, p143).

Here, two self-organising maps were trained, therefore two input data sets were used. The first input data set was the complete 10000 dimension input data set. The second was the 400 dimension data set representing all the bearing defects at 15Hz and 5kg testing conditions. The program automatically determines the number of neurons and size of the map that best represents the input data by evaluating the dimension of the input data. The map is trained through an evaluation of the quantisation error and topological error.

A training report of the two data sets is shown in annexure 8 (p148). The size of the map for the complete B12 input data is a 30 by 17 hexagonal map, this means that the map has 510 neurons. The rough training phase was performed in two epochs and took two seconds. The fine tuning phase was performed in twelve epochs and took 16 seconds. The final quantisation error was reduced to 0.523 and the final topographic error was reduced to 0.075.

For the 400 dimension, B12_15Hz_5kg input data, the size of the map is a 13 by 8 hexagonal map, with 104 neurons. The rough training phase was achieved within the first epoch and so was the fine tuning phase. The final quantisation error was reduced to 0.486 and the final topographic error was reduced to 0.035. The errors can also be viewed in terms of percentages. The quantisation error may also be used as an indicator of data outliers. Figure 6.4 shows a trained map of the B12 input data.

The black crosses are the input data, and the green dots are the neurons. The tone of green indicates the dimension differences between clusters. Green is the majority colour which indicates that the clusters are very close to each other. There is an outlier cluster shown in yellow. The quantisation error also indicates this. The clusters are labelled according to the bearing defects shown in figure 6.5.

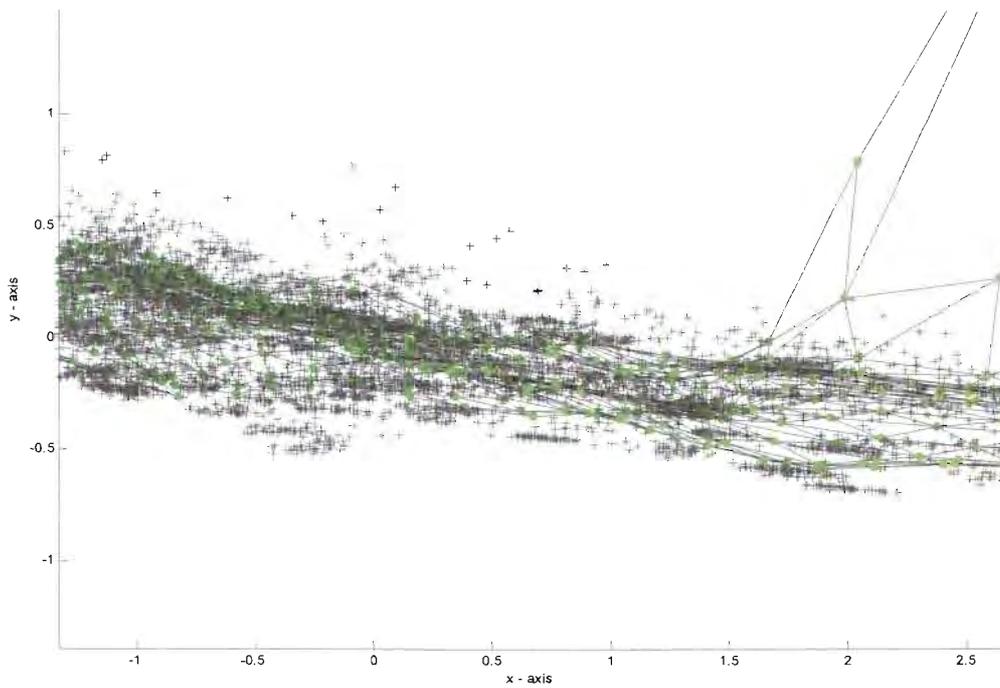


Figure 6.4: Trained Map Projection in B12 Input Data Euclidean Space



Figure 6.5: B12 Trained Map Cluster Distributions and Defect Labelling

The trained map of the B12_15Hz_5kg input data is illustrated in figure 6.41. The clusters are more distinct as indicated by the colouring. The green coloured clusters are outliers with a quantisation error of 0.486. The topology of the map is more conserved as opposed to the B12 topology, as indicated by the topological error. The representation of the map clusters is shown in figure 6.7.

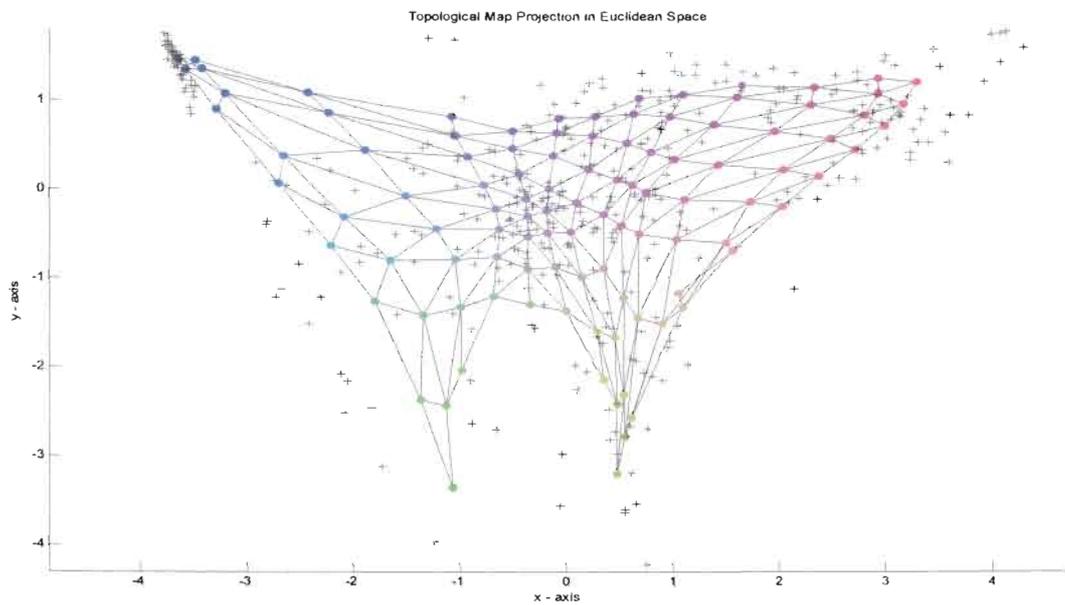


Figure 6.6: Trained Map Projection in B12_15Hz_5kg Input Data Euclidean Space

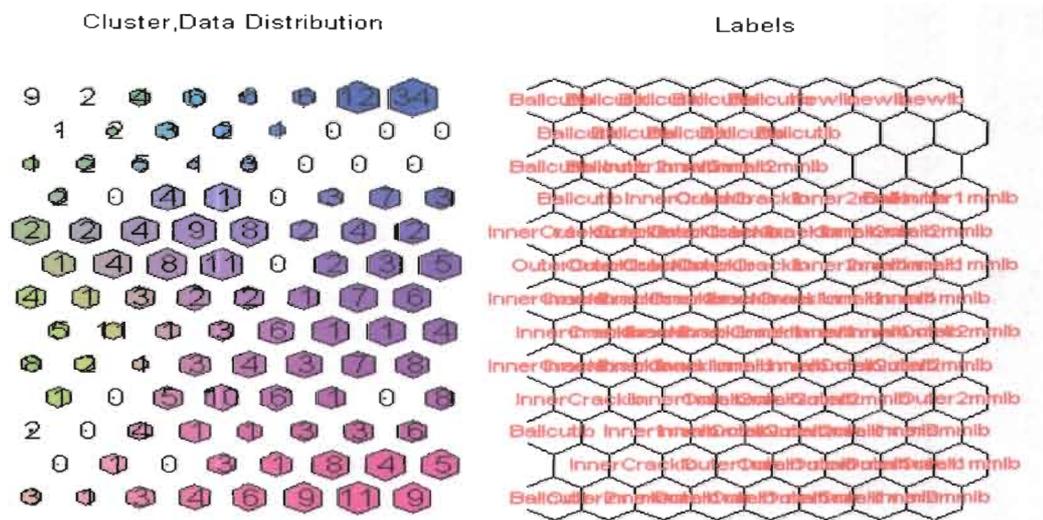
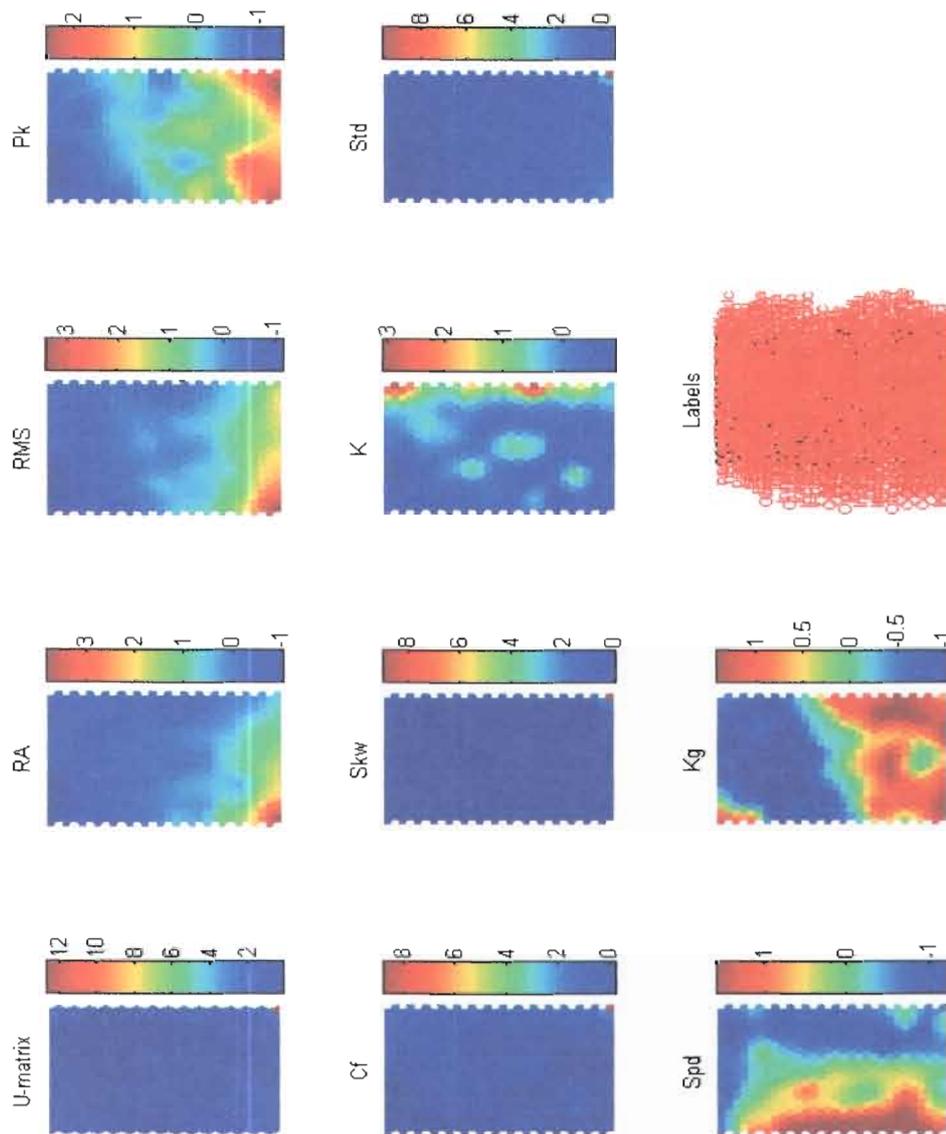


Figure 6.7: B12_15Hz_5kg Trained Map Cluster Distributions and Defect Labelling

6.5 Prognosis and Visualisation of Rolling Element Bearing Defects



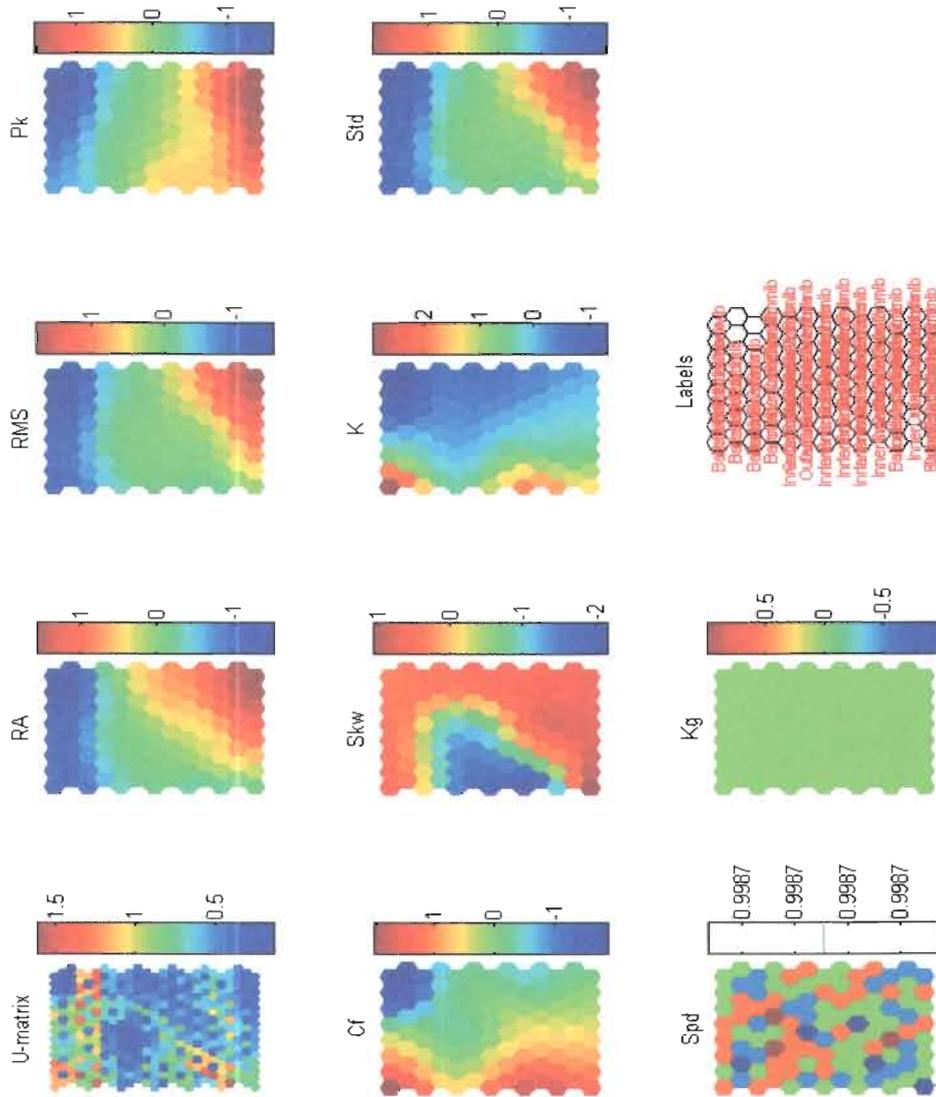
SOM 31-Jan-2006

Figure 6.8: B12 Nine Parameters Self-Organising Feature Maps

Figure 6.8 illustrates nine parameters on the self-organising feature map for the B12 input data. Each map is referred to as a matrix. The U-matrix shows how far each neuron is from its neighbours in terms of Euclidean distance, reflected by colour coding. The neurons are closely placed since the majority of the map is toned in blue. The matrices are very dense because of the large 10000 data dimensionality, so identifying the exact bearing defects is difficult. The speed and load matrix indicates the conditions under which the bearing defects were measured. Once the conditions under which a defect is located are known, then a small dimension input data can be selected and analysed separately.

This is the case with the B12_15Hz_5kg data set. The nine parameter self-organising map of the B12_15Hz_5kg input data is illustrated in figure 6.9. The B12_15Hz_5kg data set has a dimension of 400, and has larger hexagons. The variation of the parameter values are now clearly visible. Notice the single colour for the load matrix of 5kg. The speed matrix shows definite clusters for one speed condition, and neurons that are non representative because of the separation in clusters. In the U-matrix the red indicates clusters and neighbouring neurons that are far apart thus indicating a separation in clusters. The blue indicates a dense data cluster. The label matrix is to be used in conjunction with the parameter matrices, so that the classification of the cluster can be identified.

In the B12_15Hz_5kg Self-Organising map, the patterns of the RA, RMS and Std matrices are similar. With the top blue cluster classified as ball defect, and the darker blue on the top right classified as new bearings. The ball defect can resemble other defects and its clusters are spread on the left and towards the bottom of the matrix. Green represent a majority of inner crack clusters. Yellow is classified as inner ring 1mm spall, and the red towards brown as outer ring 1mm and 2mm spalls. Clusters can be differentiated. The patterns of the crest factor and kurtosis reflect the locations of the ball defect, on the right of the matrix, and a new bearing classified with low crest factors and kurtosis values. Visual analysis of bearing defects is now possible.



SOM 31-Jan-2006

Figure 6.9: B12_15Hz_5kg Nine Parameters Self-Organising Feature Maps

The pattern of the skewness matrix is the opposite to that of the crest and kurtosis matrices. The type and severity of the defect can also be analysed, by combining an analysis on different sizes of input data sets. The problem now is to locate or detect the severity and defect type of an unknown bearing vibration measurement.

The solution to locating an unknown bearing vibration measurement is to project the neuron or unit that best represents the parameters of the measured vibration. This means that the measured vibration parameters are included in the input data, as is the case in figures 6.10 and 6.11 which project a $\pm 1\text{mm}$ inner race spall.

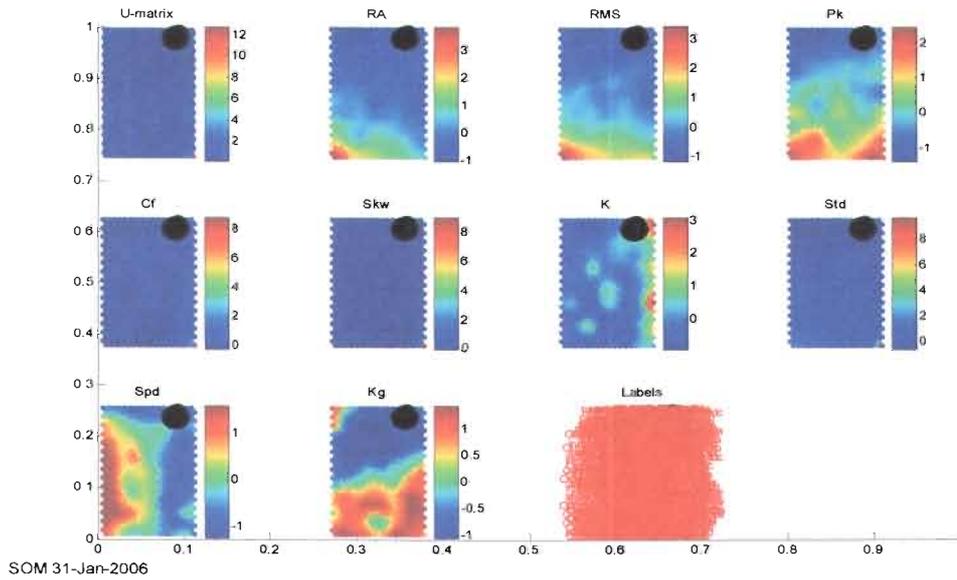


Figure 6.10: B12 Nine Parameters Defect Projections, SOM

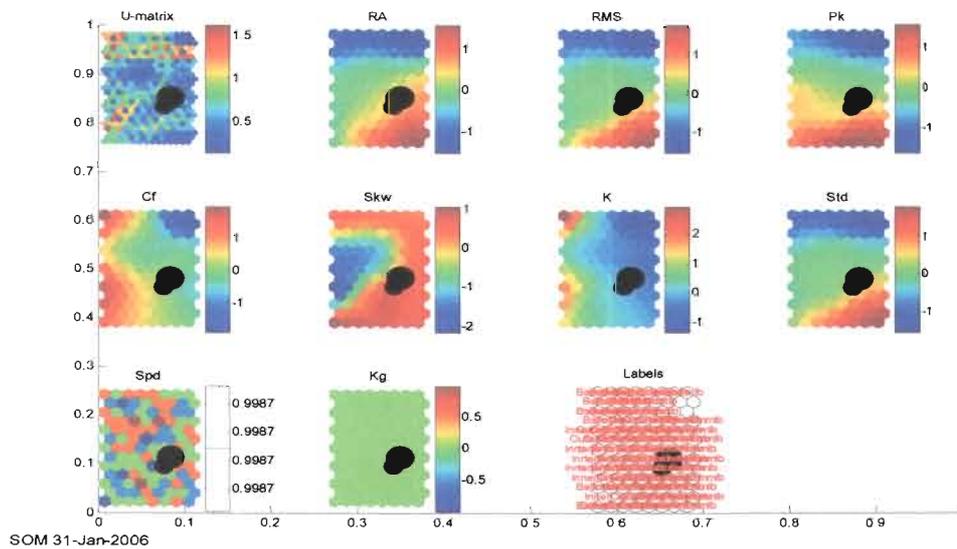


Figure 6.11: B12_15Hz_5kg Nine Parameters Defect Projections, SOM

Here, figures 6.12 and 6.13 were created from the B12 data input. On the speed matrix, the projection appears in the blue colour zone indicating that the measurement was taken at a low speed testing condition of 15Hz. On the load matrix the projection also appears in the blue colour zone indicating that the measurement was taken at a low load testing condition of 5kg. This is done to establish the conditions under which the tested bearing is operating.

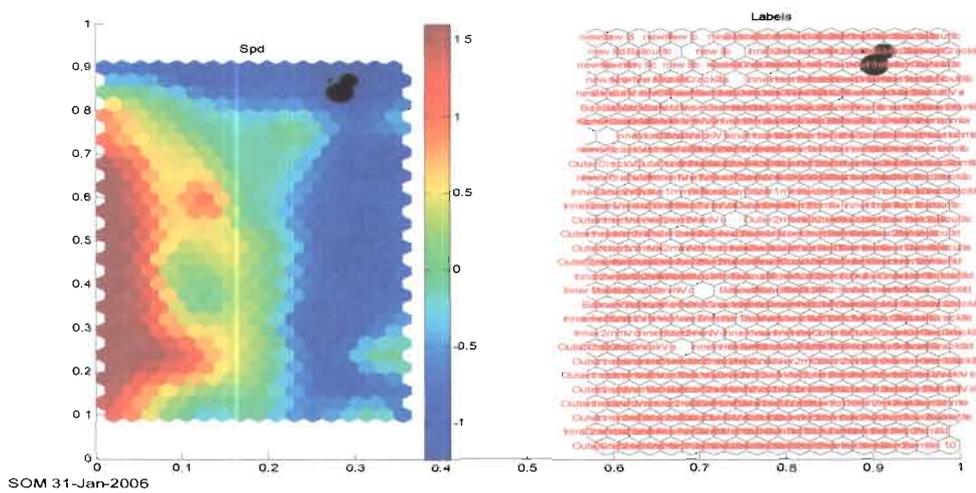


Figure 6.12: B12 Speed Clustering and Defect Projection

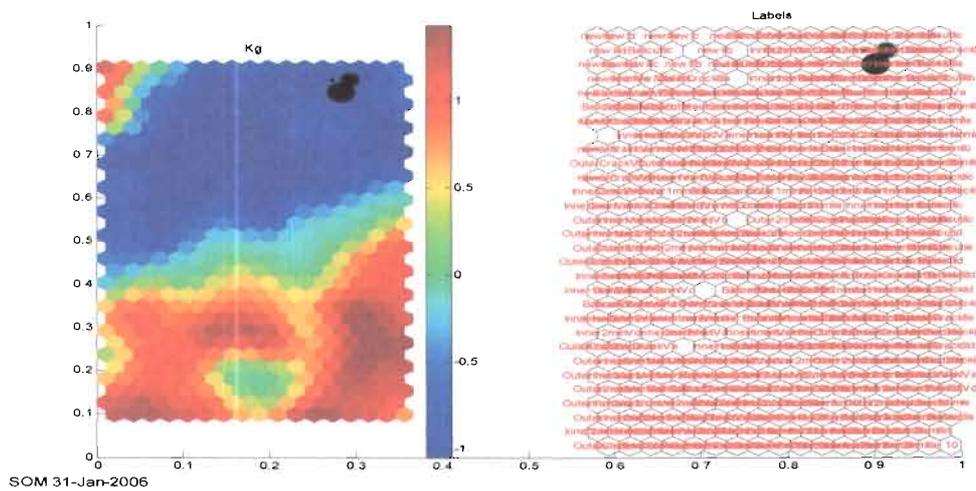


Figure 6.13: B12 Load Clustering Defect Projection

The defect projection utilises three circular black projections. The largest circle indicates the neuron that best represents the defect. The medium and smallest circles indicate the degree of defect representation by the respective projected neurons. Since the input data is represented by neurons on a hexagonal topological map, there are neurons that best resemble the different clusters. These are the winning neurons, and are shown with projected white hexagons in the U-matrix of figures 6.14 and 6.15. The largest projected hexagon is the neuron that wins the most representation.

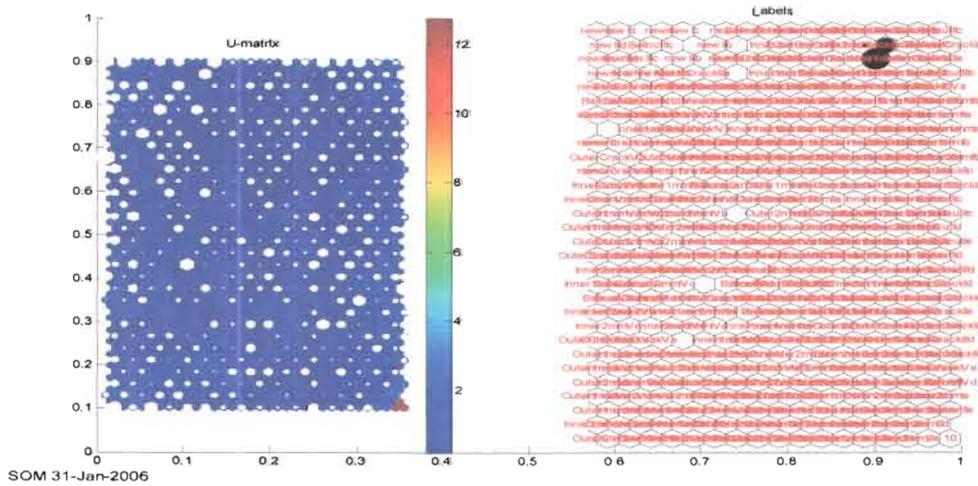


Figure 6.14: B12 Unit Matrix Clusters with Best Matching Units

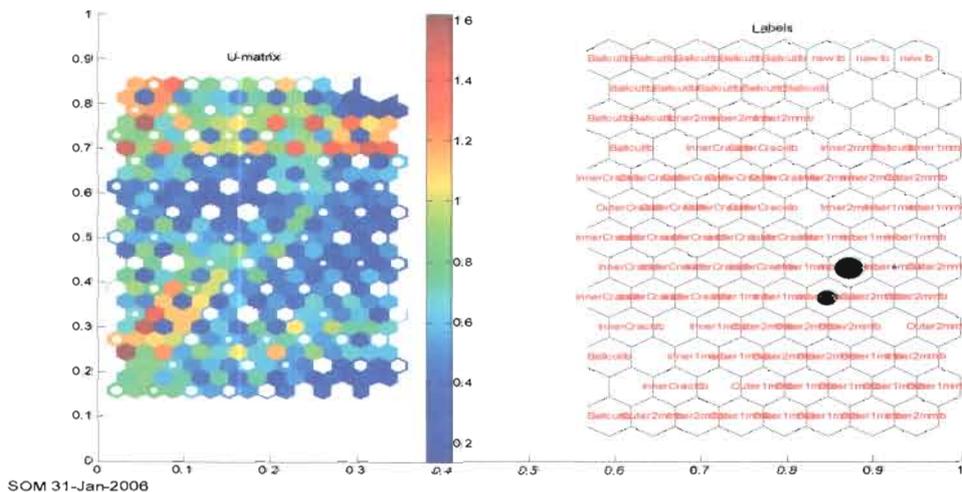


Figure 6.15: B12_15Hz_5kg Unit-Matrix Clusters with Best Matching Units

The k-means cluster using the Davies Bouldin index are displayed in figure 6.16. For the B12 input data, 30 k clusters were generated, and figure 6.17 for the B12_15Hz_5kg input data, 17 k clusters were generated. These are the major clusters that represents the input data. A cluster is indicated by a distinct colour that does not tone into neighbouring clusters. The k-means clustering assists in relating defect representation on the self-organising map.

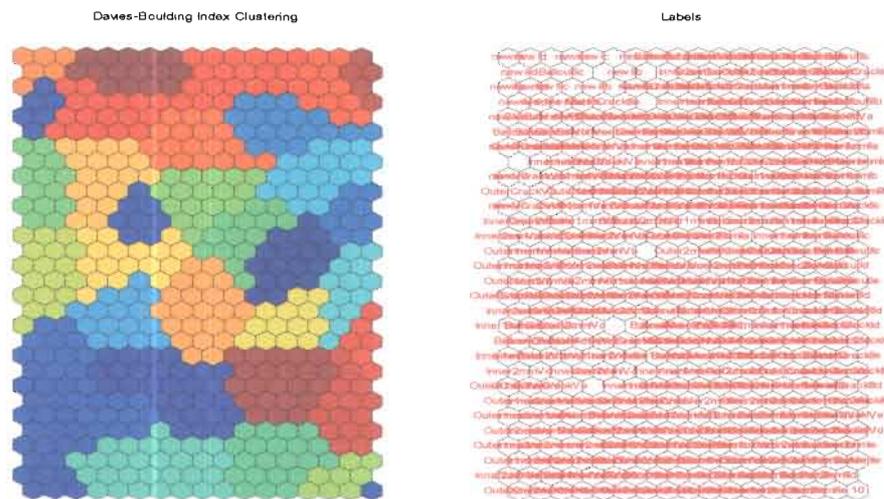


Figure 6.16: B12 Davies Bouldin Clustering

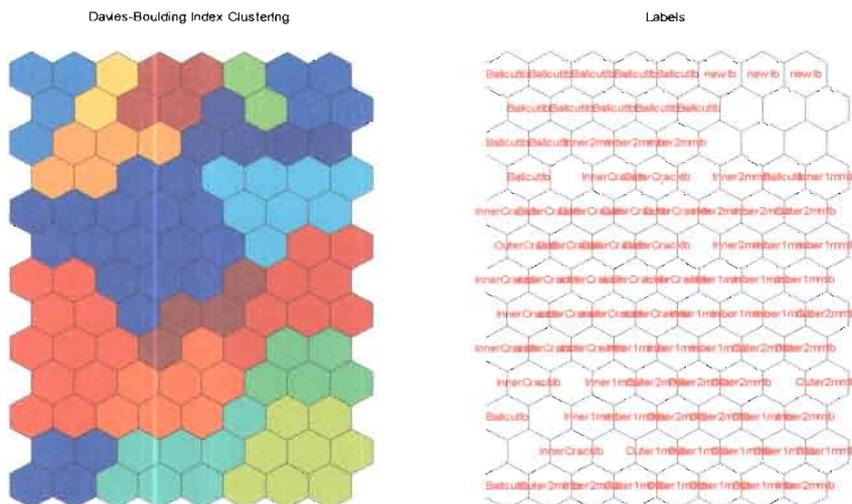


Figure 6.17: B12_15Hz_5kg Davies Bouldin Clustering

Figures 6.18 and 6.19 depict the learning vector quantisation classification maps. Clusters are automatically classified according to bearing defect. The inner race spall is automatically projected into the definite defect cluster. The colour coding indicates cluster separation as in the U-matrix. The defect projection in the learning vector quantisation map is used to support the analysis in the self-organising map.

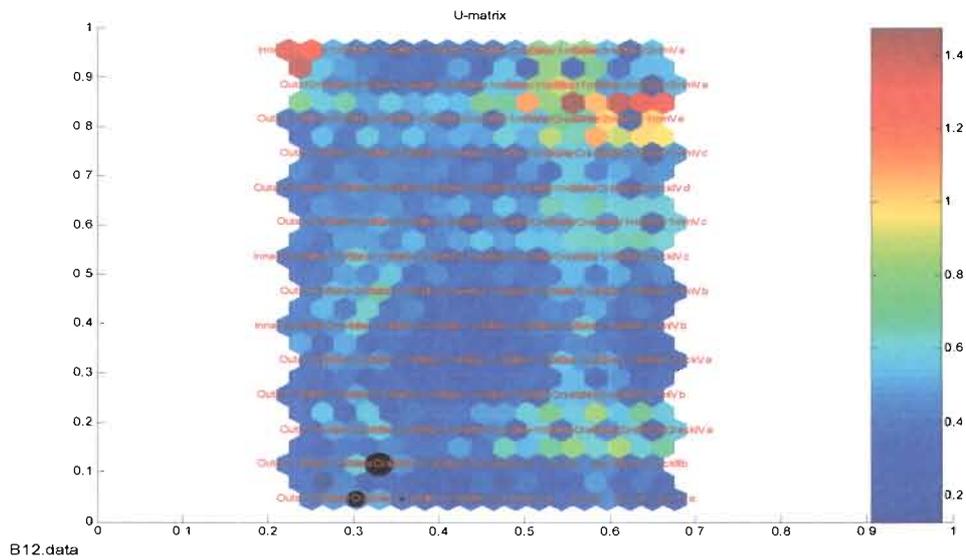


Figure 6.18: B12 Learning Vector Quantisation Defect Clustering

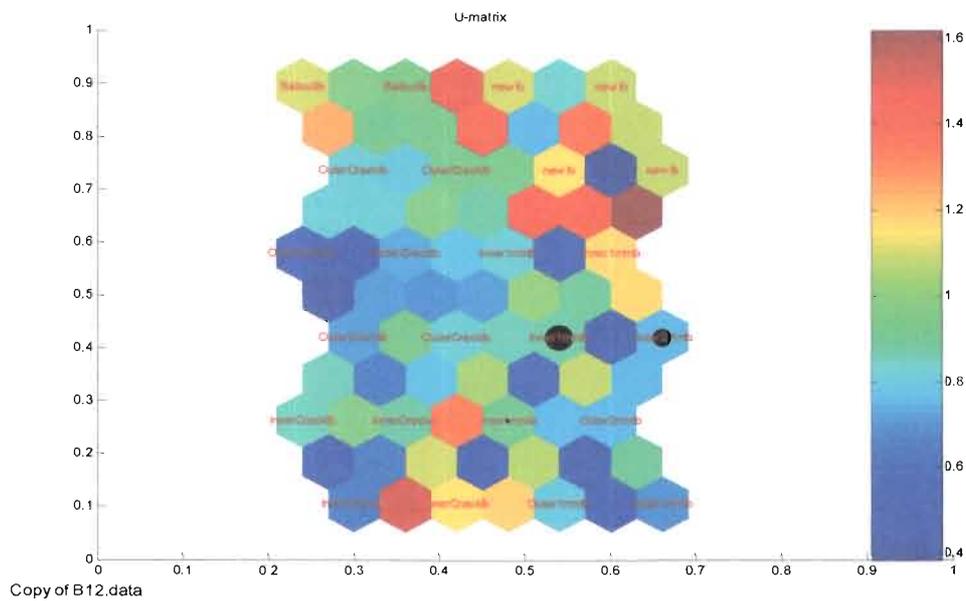


Figure 6.19: B12_15Hz_5kg Learning Vector Quantisation Defect Clustering

Taylor *et al.* (2004:83) stated that diagnosis and prognosis of rolling element bearing damage has received less attention and is considerably more involved than detection. The self-organising feature maps provide assistance in effective prognosis of rolling element bearing defects, because the defect is not only detected according to its location in the bearing but also its severity is indicated.

The nature of the bearing damage, for example spalling, fluting ,etching etc. can be detected as a spall defect in the self-organising map. Visual inspection gives more information on the nature of bearing damage. Therefore, a bearing needs to be replaced if a Self-Organising Map projects a severely damaged bearing and any of the following types of damage are visible: fatigue spalling, heat discoloration, water stains or rust, severe wear from contamination, damaged cage, brinelling.

6.6 Summary

A new practical dimension to condition monitoring of rolling element bearings has been developed. The use of signal processing time domain and frequency domain analysis of bearing vibration has been combined with a visual analysis of distinct bearing defects through the application of the Self-Organising Feature Map.

The significance of the time domain and frequency domain analysis was illustrated with a display of signal patterns of different bearing defects. It was clear that each defect provides its own unique statistical parameters. The input data set was structured in a format that enabled adequate classification of the self organising map. The analysis of a $\pm 1\text{mm}$ inner race spall was performed using both high and low dimension input data sets, for better analysis of the parameters. Other visualisation methods were used, such as the k-means clustering with the Davies Bouldin index, the winning neuron projections, and the Learning Vector Quantisation Map.

Chapter 7

Conclusions and Recommendations

7.1 Summary and Conclusions

Accuracy and reliability of vibration analysis on rolling element bearings was found to be lacking in industry. To provide assistance to the highly skilled personnel in performing condition assessments on rolling element bearings, a condition monitoring system based on artificial intelligence was proposed. To develop the system, an understanding of the characteristics and operating fundamentals of rolling element bearings was required. The deep-groove ball bearing was considered for the research. So a study on the characteristics and behaviour of rolling element bearings, without and with damage, was the initial task of the research in the form of a literature review. The literature review pointed to a neural network analysis using Kohonen's Self-Organising Maps.

The next task was to obtain a vibration measuring system in the form of a data acquisition and signal processing system using commercially available software. The software used was Lab View 7. A rig was designed and constructed to test the condition monitoring system. The rig was designed to be rigid with the aim of acquiring vibrations from bearing faults only, as these occur at frequencies much lower than the natural frequency of the rig. The various types of damage that may occur were artificially created by a corrosive technique on the raceways as well as cuts to the ball element and rings of the bearings to simulate spalls and cracks. The identification of incipient bearing faults were done using Kohonen's Self-Organising Map neural network. The network could display bearing faults on a two-dimensional map. A program was developed in Matlab 7 that automatically applies Kohonen's Self-Organising Maps to identify bearing faults after acquiring the bearing's vibration

measurements. For a new bearing, a projection of neurons is made in the 'new' bearing classified cluster. Kohonen's Self-Organising Map neural network could achieve the detection of a bearing fault with a low topological error of about ± 0.04 , but a moderate quantisation error of ± 0.5 .

Kohonen's Self-Organising Map was used together with the learning vector quantisation map, and the k-means clustering to achieve good visualisation and classification when analysing rolling element bearing faults. A prognosis, on the severity of bearing damage and the stage of failure of the bearing, could be made. The main achievement was that the high frequency detection technique by means of time domain and frequency domain analysis can be combined with the bearing fault visualisation analysis of Kohonen's Self-Organising Map, to make informed decisions possible. Adequate maintenance schedules can then be planned resulting in lower maintenance costs.

7.2 Recommendations

When a fast Fourier transformation is performed, the signal representation is moved from the time domain to the frequency domain and this change of domain can lead to loss of information and to interpretation difficulties. This disadvantage is overcome through wavelet analysis, which enables the examination of the frequency information of the signal as it evolves with time, by using shifted and scaled versions of a wavelet.

Wavelet transform makes it possible to work in the time frequency domain and to perform good time resolution at high frequencies allowing the identification of temporal instants at which transient phenomena take place. Here the presence of impulses in the vibration signals can be assessed by the high frequency part of the wavelet transform. Therefore, it is recommended that wavelets could be used in conjunction with the high frequency detection technique. Another technique that could be considered is spectrogram, that tracks the order of the frequencies in time.

In addition, statistical trending of vibration data in addition to the application of the self-organising map, could also be performed.

A recommendation is to use more sensitive techniques in acquiring vibration measurements. Thus, a more sensitive analysis of incipient bearing damage would be achieved by the self-organising map. For example, measurements in the ultrasonic ranges of the bearing frequencies could be used to set up parameters that can be incorporated with the artificial intelligence neural network to detect incipient damage within the range of the chosen ultrasonic technique.

Kohonen's neural networks could also be used together with a feed-forward multi-layer perceptron neural network for severe fault detection. The multi-layer perceptron indicates only two conditions, namely, that the bearing is either damaged or not damaged.

7.3 Fields for Further Study

The fields for further study reflect opportunities of improving the self-organising map in order to better monitor the conditions of rolling element bearings.

The bearing faults of interest in the research focused on spalls, and cracks, which were artificially created. The reason is that fatigue spalling is the predominant bearing failure mode, and a cracked bearing reflects a severely damaged condition. However further studies need to be conducted to include other bearing failure modes such as brinnelling, electric arch fluting, lubrication, etc. The severity of ball defects were not acquired, since only one severe spall on the ball element was considered. This could also be studied further.

Although the developed monitoring system recognises bearing conditions accurately, the results were obtained using experimental data obtained in a laboratory. In reality, bearing vibration measurements are acquired from bearings that are mounted on

critical industrial machinery such as turbines and electric generators in power stations. Vibration signals obtained from this environment are expected to have different characteristics than those obtained from a testing rig in the laboratory. Future work will be directed towards investigating the reliability of the developed monitoring system method in diagnosis on bearings operating in an industrial environment.

The Kohonen algorithm could also be modified in order to improve the generalisation capabilities, by introducing a pre-processor for a hidden Markov model classifier. This improvement would enable continuous classification that can be used for online monitoring by the self-organising map. Thus an online condition monitoring for rolling element bearings could be developed.

Pseudomodal energies combined with multi-layer perceptron network to identify multiple fault conditions could be used. However, instead of labelled classification on a topological map, a numeric output classification can be performed. The advantage here is that the quantisation error could be lower than the quantisation error achieved by the self-organising map. The limitation is that a fixed output would be achieved.

Fuzzy logic in addition to neural networks could be used to improve the uncertainties or errors adapted by neural networks.

Derivation of Bearing Characteristic Equations

Annexure 1: Derivation of Bearing Characteristic Equations

The symbols and variables are tabulated in table 3.1 (p44) in chapter 3. The bearing frequencies are derived from determining the angular speed of the cage.

Angular speed is:

$$\omega = \frac{V}{r} \quad (\text{A1.1})$$

Angular speed of the cage is:

$$\omega_c = \frac{V_c}{r_c} \quad (\text{A1.2})$$

Assuming that there is no slip between the balls and roller race:

$$V_c = \frac{V_i + V_o}{2} \quad (\text{A1.3})$$

Linear velocities of the outer and inner races can be expressed as follows:

$$\omega_i = \frac{V_i}{r_i} \Rightarrow V_i = \omega_i \cdot r_i \quad (\text{A1.4})$$

$$\omega_o = \frac{V_o}{r_o} \Rightarrow V_o = \omega_o \cdot r_o \quad (\text{A1.5})$$

The radius of rotation is measured from the centre of rotation to the point of roller contact on each race:

$$r_c = \frac{P_d}{2} \quad (\text{A1.6})$$

Derivation of Bearing Characteristic Equations

$$r_i = \left(\frac{P_d}{2} - \frac{B_d \cdot \cos(\alpha)}{2} \right) \quad (A1.7)$$

$$r_o = \left(\frac{P_d}{2} + \frac{B_d \cdot \cos(\alpha)}{2} \right) \quad (A1.8)$$

Substituting equations A1.3, A1.4, A1.5, A1.6, A1.7, and A1.8 into equation A1.2:

$$\omega_c = \frac{\left[\omega_i \left(\frac{P_d}{2} - \frac{B_d \cdot \cos(\alpha)}{2} \right) + \omega_o \left(\frac{P_d}{2} + \frac{B_d \cdot \cos(\alpha)}{2} \right) \right]}{\frac{P_d}{2}} \quad (A1.9)$$

$$\omega_c = \text{FTF} \quad (A1.10)$$

Fundamental train frequency with reference to the number of ball elements:

$$\text{FTF} = \frac{1}{2} \left[\omega_i \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right) + \omega_o \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \right] \quad (A1.11)$$

$$\text{FTF} = \frac{1}{2} \cdot \frac{N_s}{60} \cdot \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \quad (A1.12)$$

$$\text{FTF} = \frac{1}{2} \cdot \frac{N_s}{60} \cdot \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \quad (A1.13)$$

The ball pass frequency of the outer race is the number of ball multiplied by the rotational difference between the cage and the outer race.

$$\text{BPFO} = Z |\omega_c - \omega_o| \quad (A1.14)$$

Derivation of Bearing Characteristic Equations

$$\begin{aligned} \text{BPFO} &= \left| Z \left\{ \frac{1}{2} \left[\omega_i \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right) + \omega_o \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \right] - \omega_o \right\} \right| \\ \text{BPFO} &= \left| Z \left[\left(\frac{\omega_i}{2} - \frac{\omega_i \cdot B_d \cdot \cos(\alpha)}{2P_d} \right) + \left(\frac{\omega_o}{2} + \frac{\omega_o \cdot B_d \cdot \cos(\alpha)}{2P_d} \right) - \omega_o \right] \right| \\ \text{BPFO} &= \left| Z \left[\left(\frac{\omega_i - \omega_o}{2} \right) - \left(\frac{\omega_i - \omega_o}{2} \right) \cdot \frac{B_d \cdot \cos(\alpha)}{P_d} \right] \right| \\ \text{BPFO} &= \left| \frac{Z}{2} (\omega_i - \omega_o) \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \right| \end{aligned} \quad (\text{A1.15})$$

$$\text{BPFO} = \frac{Z}{2} \cdot \frac{N_s}{60} \cdot \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \quad (\text{A1.16})$$

The ball pass frequency of the inner race is the number of ball elements multiplied by the speed difference between the cage and the inner race.

$$\text{BPFI} = Z |\omega_i - \omega_c| \quad (\text{A1.17})$$

$$\begin{aligned} \text{BPFI} &= \left| Z \left\{ \omega_i - \frac{1}{2} \left[\omega_i \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right) + \omega_o \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \right] \right\} \right| \\ \text{BPFI} &= \left| Z \left(\omega_i - \frac{\omega_i}{2} + \frac{\omega_i \cdot B_d \cdot \cos(\alpha)}{2P_d} - \frac{\omega_o}{2} - \frac{\omega_o \cdot B_d \cdot \cos(\alpha)}{2P_d} \right) \right| \\ \text{BPFI} &= \left| Z \left[\left(\frac{\omega_i - \omega_o}{2} \right) + \left(\frac{\omega_i - \omega_o}{2} \right) \cdot \frac{B_d \cdot \cos(\alpha)}{P_d} \right] \right| \end{aligned}$$

Derivation of Bearing Characteristic Equations

$$\text{BPFI} = \left| \frac{Z}{2} (\omega_i - \omega_o) \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \right| \quad (\text{A1.18})$$

$$\text{BPFI} = \frac{Z}{2} \cdot \frac{N_s}{60} \cdot \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \quad (\text{A1.19})$$

The linear speed of the ball surface may be described as at the point of contact with either the inner or outer race.

$$\text{BSF} = \omega_r = \frac{V_r}{r_r} \quad (\text{A1.20})$$

$$V_r = (\omega_i - \omega_c) r_i \quad (\text{A1.21})$$

The linear speed of the ball surface can be defined in terms of the inner race and cage.

$$\text{BSF} = \left| \frac{(\omega_i - \omega_c) r_i}{r_r} \right| \quad (\text{A1.22})$$

$$\text{BSF} = \left| \frac{(\omega_i - \omega_c) r_i}{\frac{B_d}{2}} \right|$$

$$\text{BSF} = \left| \frac{\left\{ \omega_i - \frac{1}{2} \left[\omega_i \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right) + \omega_o \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \right] \right\} \left\{ \frac{P_d}{2} - \frac{B_d \cdot \cos(\alpha)}{2} \right\}}{\frac{B_d}{2}} \right|$$

$$\text{BSF} = \left| \frac{\omega_i - \frac{\omega_i}{2} - \frac{\omega_i \cdot B_d \cdot \cos(\alpha)}{2P_d} + \frac{\omega_o}{2} + \frac{\omega_o \cdot B_d \cdot \cos(\alpha)}{2P_d} \left(\frac{P_d}{2} - \frac{B_d \cdot \cos(\alpha)}{2} \right)}{\frac{B_d}{2}} \right|$$

Derivation of Bearing Characteristic Equations

$$\text{BSF} = \left| \omega_i - \frac{\omega_i}{2} - \frac{\omega_i \cdot B_d \cdot \cos(\alpha)}{2P_d} + \frac{\omega_o}{2} + \frac{\omega_o \cdot B_d \cdot \cos(\alpha)}{2P_d} \left(\frac{P_d - B_d \cdot \cos(\alpha)}{B_d} \right) \right|$$

$$\text{BSF} = \left| \frac{\omega_i - \omega_o}{2} \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right) \left(\frac{P_d - B_d \cdot \cos(\alpha)}{B_d} \right) \right|$$

$$\text{BSF} = \left| \frac{\omega_i - \omega_o}{2} \left(\frac{P_d}{B_d} - \cos(\alpha) + \cos(\alpha) - \frac{B_d \cdot (\cos(\alpha))^2}{P_d} \right) \right|$$

$$\text{BSF} = \left| \frac{P_d}{2B_d} (\omega_i - \omega_o) \left(1 - \left(\frac{B_d \cdot \cos(\alpha)}{P_d} \right)^2 \right) \right| \quad (\text{A1.23})$$

$$2 \times \text{BSF} = \left| \frac{P_d}{B_d} (\omega_i - \omega_o) \left(1 - \left(\frac{B_d \cdot \cos(\alpha)}{P_d} \right)^2 \right) \right|$$

$$2 \times \text{BSF} = \frac{P_d}{B_d} \cdot \frac{N_s}{60} \cdot \left[1 - \left(\frac{B_d \cdot \cos(\alpha)}{P_d} \right)^2 \right] \quad (\text{A1.24})$$

The bearing frequencies are not synchronous to the shaft rotational speed, because the shaft speed is multiplied by a fraction, due to changes of the angle of contact. The bearing frequency equations are tabulated in annexure 2.

Induced Bearing Frequencies Equations with Measurement and Applications

Annexure 2: Induced Bearing Frequency Equations with Measurement and Applications

Table A2.1: Bearing Characteristic Equations

(Layout courtesy of Taylor et al.,

2004:26)

Label	Formulas	Race Rotation	Measurement	Application
BPFO	$\frac{Z}{2} \cdot \frac{N_s}{60} \cdot \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right)$	Inner or Outer race Rotation	Discrete frequency with or without harmonics or sidebands of FTF or shaft speed.	Indicates outer race defect.
BPFI	$\frac{Z}{2} \cdot \frac{N_s}{60} \cdot \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right)$	Inner or Outer race Rotation	Discrete frequency with or without harmonics or sidebands of FTF or shaft speed.	Indicates Inner race defect.
2×BSF	$\frac{P_d}{B_d} \cdot \frac{N_s}{60} \cdot \left[1 - \left(\frac{B_d \cdot \cos(\alpha)}{P_d} \right)^2 \right]$	Inner or Outer race Rotation	Discrete frequency with or without harmonics or sidebands of FTF or shaft speed.	Indicates defects on the balls.
FTF	$\frac{1}{2} \cdot \frac{N_s}{60} \cdot \left(1 - \frac{B_d \cdot \cos(\alpha)}{P_d} \right)$ <p style="text-align: center;">or</p> $\frac{1}{2} \cdot \frac{N_s}{60} \cdot \left(1 + \frac{B_d \cdot \cos(\alpha)}{P_d} \right)$	Inner race Rotation or Outer race Rotation	Discrete frequency, sidebands at BPFO, BPFI, 2×BSF, or a difference frequency with wide band noise.	Indicates severe looseness or severe bearing malfunction.

Induced Bearing Frequencies

Annexure 3: Induced Bearing Frequencies

Table A3.1: Bearing Dimensions

Bearing	Bore	Outer Diameter	Thickness	Internal Clearance		PCD	Ball Diameter	Number of
	mm	mm	mm	Min μ	Max	mm	mm	
6008ZZ3C	40	68	15	18	36	54	7.94	12
6010ZZ3C	50	80	16	23	43	65	8.73	14
6012ZZ3C	60	95	18	23	43	77.5	10.32	14

PCD = Pitch Circle Diameter

```
clear
clc
```

```
% BEARING FREQUENCIES PROGRAM
%*****
```

```
%=====
% Bearing Type      6012  (Dimensions in millimetres)
%=====
```

```
Ns    = [900;1200;1500;1800;2400]; % Testing shaft speeds in rpm
D     = 95; % Outer ring diameter
d     = 60; % Bore diameter
Pcd   = 77.5; % Pitch circle diameter
Bd    = 10.32; % Rolling element diameter
Z     = 14; % Number of rolling elements
alpha = 0; % Contact angle in degrees
```

```
% BPFO = Outer Race Ball Pass Frequency (Hz)
% BPFI = Inner Race Ball Pass Frequency (Hz)
% 2BSF = Ball Spin Frequency Double (Hz)
% FTF  = Fundamental Train Frequency (Hz)
% Bspeed= Ball speed (Hz)
```

```
BPFO = (Z/2)* (Ns/60)*(1 - ((Bd*cos(alpha))/Pcd));
BPFI = (Z/2)* (Ns/60)*(1 + ((Bd*cos(alpha))/Pcd));
zxBSF = (Pcd/Bd)*(Ns/60)*(1 - (((Bd*cos(alpha))/Pcd)^2));
FTF = (1/2)* (Ns/60)*(1 - ((Bd*cos(alpha))/Pcd));
Bspeed= zxBSF/2;
```

```
Freq12= [BPFO BPFI FTF zxBSF Bspeed]
```

```
Freq12 =
```

```
[ N(rpm)   BPFO      BPFI      FTF      2xBSF      Bspeed   ]
  900      91.0181   118.9819   6.5013   110.6479   55.3240
 1200     121.3574   158.6426   8.6684   147.5306   73.7653
 1500     151.6968   198.3032   10.8355  184.4132   92.2066
 1800     182.0361   237.9639   13.0026  221.2959   110.6479
 2400     242.7148   317.2852   17.3368  295.0611   147.5306
```

Loading Configurations

Annexure 4: Loading Configurations

Table A4.1: Test Rig Loading Adjustments

Load cell mass	1/3 of 248.5 g = 82.83 g		
Spring masses 2 off	1/3 of 149.65 g = 49.88 g		
Case 1:	Bearing slide mass + Magnetic accelerometer mass + fittings mass (Without the bearing seat)	1.92 Kg	<i>Mass of cylinder arm and the top components is automatically measured and does not need to be added as a constant.</i>
Case 2:	Bearing slide mass + Threaded accelerometer mass + fittings mass (Without the bearing seat)	1.92 Kg	
Case 3:	Bearing slide mass + Magnetic accelerometer mass + ... Threaded accelerometer mass + fittings mass (Without bearing seat)	2.2 Kg	<i>Mass of cylinder arm and the top components equals to a rounded off mass of 1.2 Kg</i>
Total (Kg) mass to be added for case 1: 2.05271			
Total (Kg) mass to be added for case 2: 2.05271			
Total (Kg) mass to be added for case 3: 2.33271			

Bearing Type	Bearing seat mass (Kg)	VI constant change for every bearing seat in case 1 (Kg)	VI constant change for every bearing seat in case 2 (Kg)	VI constant change for every bearing seat in case 3 (Kg)
6008ZZ3C	0.8865	2.93921	2.93921	3.0865
6010ZZ3C	0.8905	2.94321	2.94321	3.0905
6012ZZ3C	0.8755	2.92821	2.92821	3.0755

Table A4.2: Bearing Design Load Ratings
(catalogue)

(from table page B-18 pg A-63 NTN)

Bearing	Rated Static Load 'Co' (KG)	Maximum Operating Static Load (KG)	1/15 of Maximum Operating Safe Static Load (KG)
6008ZZ3C	1009.17	917.4	61.16
6010ZZ3C	1437.3	1306.6	87.10
6012ZZ3C	1997.96	1816.3	121.08

```

C = [12900;16800;22700];
N = 2400;
Swor_king_hrs = 24*365*1;
Sfh = Swor_king_hrs/(500*3);
Sfn = (33.3/N)^(1/3);
SP_kg = ((Sfn/Sfh)*C)/9.81;
Mwor_king_hrs = 24*365*2;
Mfh = Mwor_king_hrs/(500*3);
Mfn = (33.3/N)^(1/3);
MP_kg = ((Mfn/Mfh)*C)/9.81;
Lwor_king_hrs = 24*365*3;
Lfh = Lwor_king_hrs/(500*3);
Lfn = (33.3/N)^(1/3);
LP_kg = ((Lfn/Lfh)*C)/9.81;
LP_kg is the equivalent dynamic bearing load for non-stop operation until failure after 3yrs
    
```

```

% Rated basic dynamic load (N)
% Maximum shaft speed (rpm)
% FOR 1 YEAR OF OPERATION
% life factor
% speed factor
% Dynamic load until Failure
% FOR 2 YEARS OF OPERATION
% life factor
% speed factor
% Dynamic load until Failure
% FOR 3 YEARS OF OPERATION
% life factor
% speed factor
    
```

```

Load_Matrix_kg = [LP_kg MP_kg SP_kg]
Load_Matrix_kg =
    
```

Bearing	3-years	2-years	1-year
6008ZZ3C	18.0356	27.0534	54.1068
6010ZZ3C	23.4882	35.2324	70.4647
6012ZZ3C	31.7371	47.6056	95.2113

Natural Responses

Annexure 5: Natural Responses

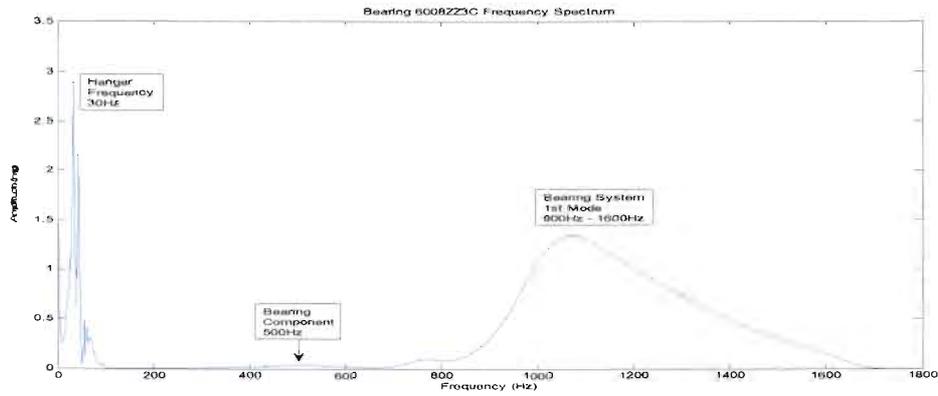


Figure A5.1: Bearing 6008ZZ3C Frequency Response

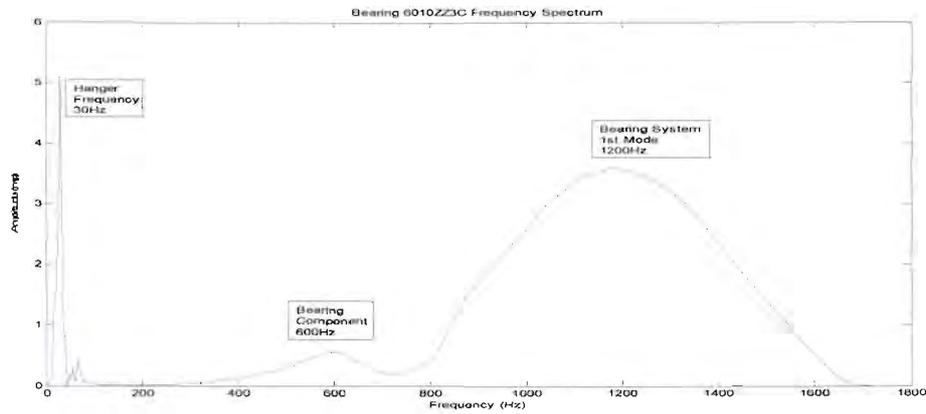


Figure A5.2: Bearing 6010ZZ3C Frequency Response

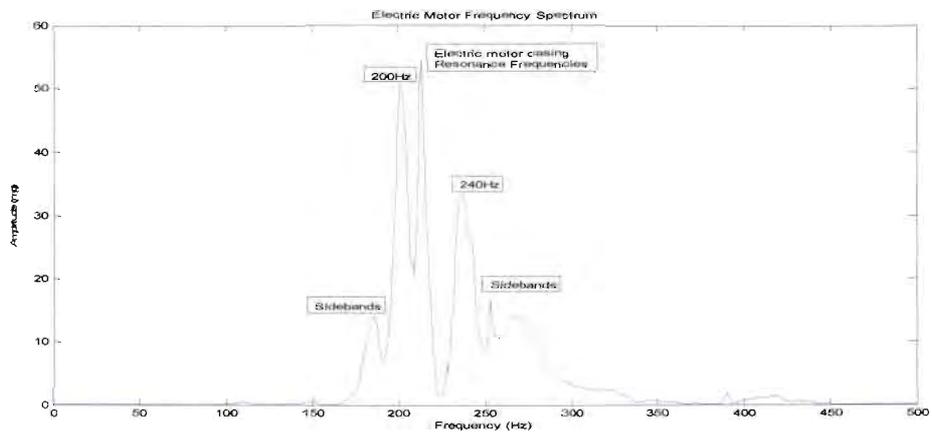


Figure A5.3: Electric Motor Casing Frequency Response

Natural Responses

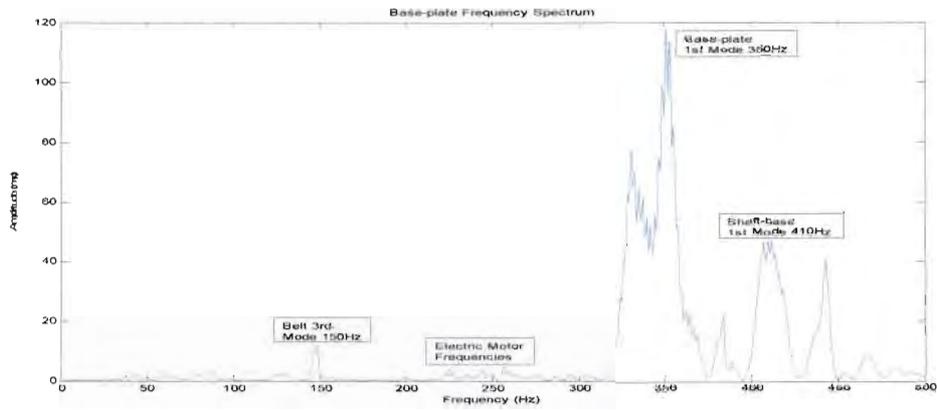


Figure A5.4: Base-plate Frequency Response

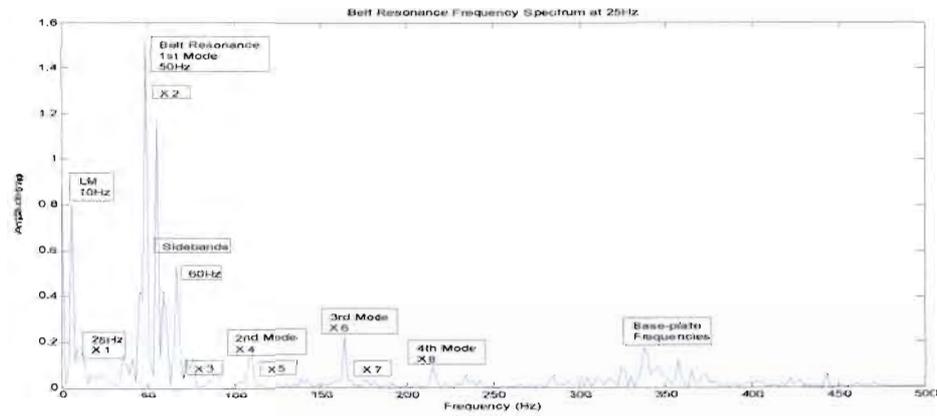


Figure A5.5: Belt Resonance Frequency Response

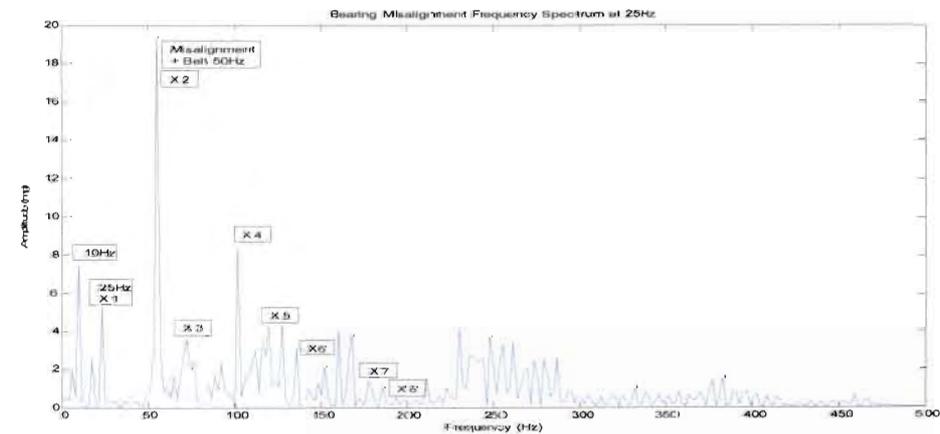


Figure A5.6: Bearing Misalignment Frequency Response

Input Data Feature Parameters

Annexure 6: Input Data Feature Parameters

```
% SINGLE SIGNAL FEATURE PARAMETERS PROGRAM
%*****

%This program determines statistical parameters of a time domain
vibration waveform. The program uses the digital time series of the
waveform to calculate statistical parameters. This program calculates
for a data set of one time waveform.

clear
clc
echo on

% Time Domain digital Signal:
%*****
data = xlsread('outer_race_1mm2');
y = data(:,2);
t = data(:,1);

%FILTER DESIGN
%*****
% The signal is filtered with a high-pass filter at a sampling
% frequency of 8192, a cut-off frequency of 600Hz up to 800Hz.
% 3dB decrease in passing magnitude, 20dB decrease to stop-band.

Wp = 600/4096; Ws = 800/4096;
[n,Wn] = buttord(Wp,Ws,3,20);
[d,c] = butter(n,Wn,'high');
SA = filter(d,c,y);

y = SA;
N = length(t);

% Statistical parameters:
%*****
mu = mean(y); % MEAN, mu
rng = max(y) - min(y); % RANGE, rng
med = median(y); % MEDIAN, med
va = var(y); % VARIANCE, va
sigma = std(y); % STANDARD DEVIATION, sigma
RA = (1/N)*(sum(abs(y))); % RECTIFIED AVERAGE, RA
RMS = sqrt((1/N)*sum((y-mu).^2)); % ROOT MEAN SQUARE, RMS
Pk = max(abs(y)); % MAXIMUM PEAK, Pk
Cf = Pk/RMS; % CREST FACTOR, Cf
Skw = ((1/N)*sum((y-mu).^3))./(sigma.^3); % SKEWNESS, Skw
K = ((1/N)*sum((y-mu).^4))./(sigma.^4); % KURTOSIS, K
P = [mu rng med va sigma] % Statistical Matrix

PF = [RA RMS Pk Cf Skw K sigma] % Parameter Matrix

echo off
```

Automatic Bearing Defect Classification

Annexure 7: Automatic Bearing Defect Classification

```
% SINGLE SIGNAL FEATURE PARAMETERS PROGRAM
%*****

%This program determines statistical parameters of one time domain
vibration waveform. The parameters are linked to the Self-Organising
Feature Map program. The tested waveform is projected on the map and
analysed.

clear
clc
clf reset;
figure(gcf)
echo on

% TIME DOMAIN DIGITAL SIGNAL:
%*****
data = xlsread('outer_race_1mm2');
y = data(:,2);
t = data(:,1);

%FILTER DESIGN
%*****

Wp = 600/4096; Ws = 800/4096;
[n,Wn] = buttord(Wp,Ws,3,20);
[d,c] = butter(n,Wn,'high');
SA = filter(d,c,y);

y = SA;
N = length(t);

% STATISTICAL PARAMETERS:
%*****
mu = mean(y);
rng = max(y) - min(y);
med = median(y);
va = var(y);
sigma = std(y);
RA = (1/N)*(sum(abs(y)));
RMS = sqrt((1/N)*sum((y-mu).^2));
Pk = max(abs(y));
Cf = Pk/RMS;
Skw = ((1/N)*sum((y-mu).^3))./(sigma.^3);
K = ((1/N)*sum((y-mu).^4))./(sigma.^4);
P = [mu rng med va sigma]

PTEST = [RA RMS Pk Cf Skw K sigma]

pause
clc
```

```
% MEAN, mu
% RANGE, rng
% MEDIAN, med
% VARIANCE, va
% STANDARD DEVIATION, sigma
% RECTIFIED AVERAGE, RA
% ROOT MEAN SQUARE, RMS
% MAXIMUM PEAK, Pk
% CREST FACTOR, Cf
% SKEWNESS, Skw
% KURTOSIS, K
% Statistical Matrix
% Parameter Matrix
```

Self-Organising Feature Map Program

```
%TESTED BEARING DEFECT SELF-ORGANISING FEATURE MAP
%*****

sDP = som_data_struct([PTEST]);
sDP = som_label(sD, 'add', [1:1]', 'Test');

%For fifty vibration parameters;
%sD = som_label(sD, 'add', [1:50]', 'Test');

Dt      = som_read_data('B12_5kg_15Hz.data');
Dtest   = [sDP;Dt.data]; % [NB. Correct and check data set labels!!]
sDtest  = som_normalize(Dtest, 'var');
sMtest  = som_make(sDtest);
sMtest  = som_autolabel(sMtest, sDtest, 'vote');

%TRAINING THE DIAGNOSING SOM
%*****

D = som_read_data('B12_5kg_15Hz.data');
sD = som_normalize(D, 'var');
sM = som_make(sD);
sM = som_autolabel(sM, sD, 'vote');
figure(1)
som_show(sM, 'umat', 'all', 'comp', [1:9], 'empty', 'Labels', 'norm', 'd');
som_show_add('label', sM.labels, 'textsize', 8, 'textcolor', 'r', ...
    'subplot', 11);

pause
clc

%DEFECT PROJECTIONS
%*****

Feat = sMtest.codebook(73, :);
Dtraj = [Feat];

bm = som_bmus(sM, Dtraj, [1:3]);
T = bm';

figure(2)
som_show(sM, 'umat', 'all', 'comp', [1:9], 'empty', 'Labels', 'norm', 'd');
som_show_add('comet', T, 'MarkerColor', 'k', 'subplot', 'all');
som_show_add('label', sM.labels, 'textsize', 8, 'textcolor', 'r', ...
    'subplot', 11);

pause
clc
```

Self-Organising Feature Map Program

```
%COMPONENT DEFECT PROJECTIONS
%*****

figure(3)
som_show(sM, 'comp', [2], 'empty', 'Labels', 'norm', 'd');
som_show_add('comet', T, 'MarkerColor', 'k', 'subplot', 'all');
som_show_add('label', sM.labels, 'textsize', 8, 'textcolor', 'r', ...
    'subplot', 2);

pause
clc

%COMPONENT DEFECT PROJECTIONS
%*****

figure(4)
som_show(sM, 'comp', [8], 'empty', 'Labels', 'norm', 'd');
som_show_add('comet', T, 'MarkerColor', 'k', 'subplot', 'all');
som_show_add('label', sM.labels, 'textsize', 8, 'textcolor', 'r', ...
    'subplot', 2);

pause
clc

%COMPONENT DEFECT PROJECTIONS
%*****

figure(5)
som_show(sM, 'comp', [9], 'empty', 'Labels', 'norm', 'd');
som_show_add('comet', T, 'MarkerColor', 'k', 'subplot', 'all');
som_show_add('label', sM.labels, 'textsize', 8, 'textcolor', 'r', ...
    'subplot', 2);

pause
clc

%CLUSTERS
%*****

% The Davies-Boulding index clustering
figure(6)
subplot(1,2,1)
[c,p,err,ind] = kmeans_clusters(sM, 40); % find at most 8 clusters
[dummy,i] = min(ind); % most probable is 7 or 6
cl = p{i}; % clusters
som_cplane(sM, cl);
title('Davies-Boulding Index Clustering')

subplot(1,2,2)
som_cplane(sM, 'none')
hold on
som_grid(sM, 'Label', sM.labels, 'Labelsize', 8, ...
    'Line', 'none', 'Marker', 'none', 'Labelcolor', 'r');
hold off
title('Labels')
```

Self-Organising Feature Map Program

```
pause
clc
```

```
%EUCLIDEAN PLANE PROJECTIONS
%*****
```

```
figure(7)
[Pd,V,me,l] = pcaproj(sD,2); Pm = pcaproj(sM,V,me); % PC-projection
Code = som_colorcode(Pm); % colour coding
hits = som_hits(sM,sD); % hits
U = som_umat(sM); % U-matrix
Dm = U(1:2:size(U,1),1:2:size(U,2)); % distance matrix
Dm = 1-Dm(:)/max(Dm(:)); Dm(find(hits==0)) = 0; % clustering info

subplot(1,3,1)
som_cplane(sM,Code,Dm);
hold on
som_grid(sM,'Label',cellstr(int2str(hits)),...
         'Line','none','Marker','none','Labelcolor','k');
hold off
title('Cluster,Data Distribution')

subplot(1,3,2)
som_grid(sM,'Coord',Pm,'MarkerColor',Code,'Linecolor','k');
hold on, plot(Pd(:,1),Pd(:,2),'k+'), hold off, axis tight, axis equal
title('Map Projection')
rotate3d on %Rotates axis three dimensionally

subplot(1,3,3)
som_cplane(sM,'none')
hold on
som_grid(sM,'Label',sM.labels,'Labelsize',8,...
         'Line','none','Marker','none','Labelcolor','r');
hold off
title('Labels')
```

```
pause
clc
```

```
figure(8)
[Pd,V,me,l] = pcaproj(sD,2); Pm = pcaproj(sM,V,me);
som_grid(sM,'Coord',Pm,'MarkerColor',Code,'Linecolor','k');
hold on, plot(Pd(:,1),Pd(:,2),'k+'), hold off, axis fill, axis equal
title('Map Projection')
rotate3d on %Rotates axis three dimensionally
```

```
pause
clc
```

Self-Organising Feature Map Program

```
%BEST MATCHING UNIT,WINNING NEURON,U-MATRIX OR NEURON MATRIX
%*****
figure(9)
som_show(sM,'umat','all','empty','Labels');
som_show_add('comet',T,'MarkerColor','k','subplot',2);
som_show_add('label',sM.labels,'Textsize',8,'TextColor','r',...
    'Subplot',2);
h = som_hits(sM,sD);
som_show_add('hit',h,'MarkerColor','w','Subplot',1);

% Here are the distances to BMU, 2-BMU and WMU:
[qe] = som_bmus(sM,Dtraj,'all');
qe(1,[1,2,end]) % first test vector

pause
clc

%COMPONENTS BEST MATCHING UNIT,WINNING NEURON
%*****
figure(10)
som_show(sM,'comp',[2],'empty','Labels','norm','d');
som_show_add('comet',T,'MarkerColor','k','subplot',2);
som_show_add('label',sM.labels,'Textsize',8,'TextColor','r',...
    'Subplot',2);
h = som_hits(sM,sD);
som_show_add('hit',h,'MarkerColor','w','Subplot',1);

pause
clc

% LEARNING VECTOR QUANTISATION
%*****
%Clustering using the LVQ algorithm

sM = som_supervised(sD,'small');
bm = som_bmus(sM,Dtraj,[1:3]);
T = bm';

figure(11)
som_show(sM,'umat','all');
som_show_add('comet',T,'MarkerColor','k','subplot','all');
som_show_add('label',sM.labels,'TextSize',8,'TextColor','r');

sDLv = som_label(sD,'clear','all');
sDLv = som_autolabel(sDLv,sM); % classification
ok = strcmp(sDLv.labels,sD.labels); % errors
100*(1-sum(ok)/length(ok)) % error percentage (%)

echo off
```

SOM Training Report

Annexure 8: SOM Training Report

```
clf reset;
figure(gcf)
echo on
```

Complete Input Data Set

```
%TESTED BEARING B12.Data SOM
%*****
```

```
Dtest = som_read_data('B12.data');
```

```
data read ok
sDtest = som_normalize(Dtest,'var');
sMtest = som_make(sDtest);
Determining map size...
  map size [30, 17]
Initialization...
Training using batch algorithm...
Rough training phase...
```

```
Training:  2/  2 s
Finetuning phase...
```

```
Training:  2/  5 s
Training:  4/  5 s
Training:  6/  6 s
Final quantization error: 0.523
Final topographic error: 0.075
```

```
%TRAINING THE DIAGNOSING SOM
%*****
```

```
D = som_read_data('B12.data');
```

```
data read ok
sD = som_normalize(D,'var');
sM = som_make(sD);
Determining map size...
  map size [30, 17]
Initialization...
Training using batch algorithm...
Rough training phase...
```

```
Training:  2/  2 s
Finetuning phase...
```

```
Training:  2/  5 s
Training:  4/  6 s
Training:  6/  6 s
Final quantization error: 0.523
Final topographic error: 0.075
```

SOM Training Report

15Hz - 5KG Input Data Set

```
%TESTED BEARING SOM
%*****
Dtest = som_read_data('B12_15Hz_5KG.data');

data read ok
sDtest = som_normalize(Dtest,'var');
sMtest = som_make(sDtest);
Determining map size...
  map size [13, 8]
Initialization...
Training using batch algorithm...
Rough training phase...

Training:   0/   0 s
Finetuning phase...

Training:   0/   0 s
Training:   0/   0 s
Final quantization error: 0.486
Final topographic error:  0.035
```

```
%TRAINING THE DIAGNOSING SOM
%*****
D = som_read_data('B12_15Hz_5KG.data');

data read ok
sD = som_normalize(D,'var');
sM = som_make(sD);
Determining map size...
  map size [13, 8]
Initialization...
Training using batch algorithm...
Rough training phase...

Training:   0/   0 s
Finetuning phase...

Training:   0/   0 s
Training:   0/   0 s
Final quantization error: 0.486
Final topographic error:  0.035
```

References

- Akedson, J, McMahon, D. 1985. A direct assessment of bearing conditions with minimal knowledge of history. *Conditions monitoring and diagnostic technology*, 1(2): pp263-273, Jul.
- Al-Gahmd, A.M, Mba, D. 2004. A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size. *Mechanical systems and signal processing*, [online]ws: pp1-35. Oct. Available at:
<<http://www.elsevier.com/locate/jnlabr/ymssp>>. Accessed: 4/5/2005.
- Alguindigue, I.E, Loskiewicz-Buczak, A, Uhrig, R.E. 1993. Monitoring and diagnosis of rolling element bearings using artificial neural networks. *IEEE transactions on industrial electronics*, 40(2): pp209-217, Apr.
- Ambros-Ingerson, J, Granger, R, Lynch, G. 1990. Simulation of paleo-cortex performs hierarchical clustering. *Science*, 247: pp1344-1348.
- An, J.L, Han, T, Yang, B.S. 2003. ART-KOHONEN neural network for fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 18: pp645-657.
- Avallone, E, Marks, L.S, Baumeister, T. 1989. *Marks standard handbook for mechanical engineers*. Eighth edition. New York: McGraw-Hill. pp738-745.
- Bat Services. 2004. *Bearing Analysis*. Johannesburg. pp3-5
- Barkov, A.V, Barkova, N.A. (1995a) Condition assessment and life prediction of rolling element bearings, part 1. *Vibro-acoustical systems and technologies*.

[online]ws: pp1-5. Jun. Available at:

<http://www.vibrotek.com/articles_bearings.php>. Accessed: 27/07/2005.

Barkov, A.V, Barkova, N.A. (1995b) Condition assessment and life prediction of rolling element bearings, part 2. *Vibro-acoustical systems and technologies*.

[online]ws: pp1-5. Sept. Available at:

<http://www.vibrotek.com/articles_bearings.php>. Accessed: 27/07/2005.

Barthel, K.. 1977. The shock pulse method for measuring the condition of antifriction bearings. *Tappi*, 60(8): pp111-113, Aug.

Bermington, S, Harrison, P. 2001. *Great inventions, the illustrated science encyclopaedia*, pp194-196.

Birchon, D. 1975. Non destructive testing. *Oxford university press*, 9(1): pp55-75, Febr.

Bouldin, D.L. 1979. A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell*, PAMI-I.pp224–227.

Brown P.J. 1977. Condition monitoring of rolling element bearings. *Noise control vibration insulation*, 8(2): pp41-44, Febr.

Bunch, B.H, Draper, E.T, Wilson I. 1995. *The world book of maths power*. Volume 1. London: World Book, Inc. pp393-416.

Chen, P. 2000. *Bearing Condition Monitoring and Fault Diagnosis*. M.Sc. Dissertation. Alberta. University of Calgary. pp2–64.

Collins, J.A. 2003. *Mechanical design of machine elements and machine - a failure prevention perspective*. Ohio: John Wiley & Sons, Inc. pp54.

Demuth, H, Beale, M. 2001. *Neural network toolbox for use with Matlab*. Version 4. Natick: The math works, Inc. pp1–200.

De Backer, N, Naud, A, Scheunders, P. 1998. Non-linear dimensionality techniques for unsupervised feature extraction. *Pattern Recognition Letters*, 19: pp711–720, Jan.

Dorf R.C, Bavarian B. 1993. *The Electrical Engineering Hand Book*, London: Prentice Hall. pp420-429.

Durbin, R, Rumelhart, D.E. 1989. A computationally powerful and biologically plausible extension to back propagation network. *Neural computation*, 1: pp133–142.

Dyer, D, Stewart, R. 1978. Detection of rolling element bearing damage by statistical vibration analysis. *Journal of mechanical design*, 100: pp229-235, Apr.

Euliano, N.R, Principe, J.C, Lefebvre, W.C. 2000. *Neural and adaptive systems*. U.S.A: John Wiley & Sons, Inc. pp300-320.

Fukushima, K. 1975. Cognitron: a self-organising multi-layered neural network. *Biological cybernetics*, 20: pp121-126.

Gani A, Salami M.J.E. 2004. Vibration faults simulation system: a lab equipment to aid teaching of macaronis courses. *International journal of engineering education*, 20(1): pp61-69, Sept.

Glenn, S. 1996. *Probabilistic expert systems*. New York: Society for industrial and applied mathematics. pp1- 20.

Godin, N, Huguet, S, Gaertner, R. 2004. Integration of the Kohonen's self-organising map and k-means algorithm for the segmentation of the AE data collected during tensile tests on cross-ply composites. *NDT&E International*, 38: pp299-309, Sept.

Grossberg, S. 1972. Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or un-learnable pattern classes. *Kybernetik*, 10: pp49-57.

Grossberg, S. 1976a. Adaptive pattern classification and universal recording: I. parallel development and coding of neural detectors. *Biological Cybernetics*, 23: pp121-134.

Grossberg, S. 1976b. Adaptive pattern classification and universal recording: II. feedback, expectation, olfaction, illusions. *Biological Cybernetics*, 23: pp187-202.

Hansford, C. 2001. Vibration - an effective indicator of process efficiency . *The South African mechanical engineer*, 51: p39, Sept.

Harris, T.A. 2001. *Rolling Bearing Analysis*. Fourth edition. U.S.A: Vibration consultants, Inc. John Wiley & Sons, Inc. pp1-67

Halsey, D, Bernard, J. 1989. Bearings. In: *Collier's encyclopaedia*, B:738-739.

Haykin, S. 1994. *Neural networks a comprehensive foundation*. U.S.A: Macmillan. pp2-443.

Hebb, D.O. 1949. *The organisation of behaviour: a neuropsychological theory*. NewYork: John Wiley & Sons, Inc. p5.

Irwin, G.W, Warwick K. 1995. Neural network applications in control. *IEEE transactions on neural networks*, U.K.p2.

Kangas, J, Kohonen, T, Laaksonen, J. 1990. Variants of self-organising maps. *IEEE transactions on neural networks*. 1(1): pp93-99, Mar.

Kardar, M. 2001. The effect of friction, wear and lubrication on bearing performance. *Mechanical Technology*: p17, Sept.

Kohonen, T. 1982a. Self-organised formation of topologically correct feature maps. *Biological Cybernetics*, 43: pp59-69.

Kohonen, T. 1982b. Clustering, taxonomy, and topological maps of patterns. *Proceedings of the 6th international conference on pattern recognition*. Munich. pp114-128.

Kohonen, T. 1990a. Self-organising map. *Proceedings of the IEEE*, 78: pp1464-1480.

Kohonen, T. 1990b. Improved versions of learning vector quantisation. *International joint conference on neural networks*. San Diego, 1: pp545-550.

Kohonen, T. 1996. Developments and applications of the self-organising map and related algorithms. *Mathematics and computers in simulation*, 41: pp3-12.

Kohonen, T. 1998. The self-organising map. *Neurocomputing*, 21(1): pp1-6, Dec.

Kreyszig, E. 1999. *Advanced engineering mathematics*. eighth edition. New York: John Wiley and sons. Pp559-1077.

Li, Y, Shiroishi, J. 1997. Bearing condition diagnostics via vibration and acoustics emission measurements. *Mechanical systems and signal processing*, 11(5): pp693-705, Apr

Li, B, Chow, M, Tipsuwan, Y, Hung J.C. 2000. Neural-network-based motor rolling bearing fault diagnosis. *IEEE transactions on industrial electronics*, 47(5): pp1060-1069. Oct.

- Macaully, D 1989. *The way things work*. London: D.K Limited. p10.
- MacQueen, J. 1967. Some methods for classification and analysis of multivariate observation. *In proceedings of the 5th Berkeley symposium on mathematical statistics and probability*. California: University of California press, 1: pp281-297.
- Mao, J, Jain A.K. 1995. Artificial neural networks for feature extraction and multivariate data projection. *IEEE transactions on neural networks*, 6: pp296-317.
- Martin, H.R, Honarvar, F. 1994. Application of statistical moments to bearing failure detection. *Applied acoustics*, 44: pp67-77, Apr.
- Martin-del-Brio, B, Serrano-Cinca, C. 1993. Self-organising neural networks for the analysis and representation of data: some financial cases. *Neural computing and applications*, 1: pp193-206.
- Mathew, J, Alfredson, R.J. 1984. Condition monitoring of rolling element bearings. *Vibration acoustics*, 106: pp447-453, Jul.
- McCulloch, W.S, Pitts, W.H. 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biophysics*, 5: pp115-133.
- Naber, G. 1980. *Topological methods in Euclidean spaces*. NewYork: Cambridge University press. pp397-443.
- Oja, E. 1982. A simplified neuron model as a principal component analyser. *Journal of mathematical biology*, 15(1): pp267-273.
- Ono, K, Okada, Y. 1998. Analysis of ball bearing vibrations caused by outer race waviness. *ASME*, 120(1): pp901-908, Oct.

Oppenheim, A.V, Willsky, A.S. 1983. *Signals and systems*. London: Prentice-Hall international, Inc. pp85-100.

Piersol, A.G, Bendat, J.S. 1986. *Random data analysis and measurement procedures*. Second edition. New York: John wiley and sons. pp370-397.

Proakis, J.G, Manolakis, D.G. 1996. *Digital signal processing, principles, algorithms, and applications*. Third edition. New Jersey: Prentice-Hall, Inc. pp200-210.

Proakis, J.G, Ingle, V.K. 2004. *A self-study guide for digital signal processing*. New Jersey: Prentice-Hall, Inc. pp110-130.

Ramirez, R.W. 1985. *The FFT fundamentals and concepts*. U.S.A: Tekronix, Inc. pp1-50.

Rao, S.S. 1995. *Mechanical Vibrations*. Third edition. U.S.A: Addison-Wesley, Inc. pp15-821.

Robson, E. 1989. *Predictive maintenance through the monitoring and diagnostics of rolling element bearings*. U.S.A: Bently Nevada .p47.

Rosenblatt, F. 1958. The perceptron: a probabilistic model for information storage and organisation in the brain. *Psychological Review*, 65: pp386-408.

Ross, J. 1997. Materials and manufacturing impact bearing performance. *Power transmission design*, 39(2): pp45-47, Aug.

Rubini, R, Meneghetti, U. 2000. Application of the envelope and wavelet transform analysis for the diagnosis of incipient faults in ball bearings. *Mechanical systems and signal processing*, 15(2): pp287-302. Jul.

Ruiyun, Z. 1984. Friction experiment. *International journal of mechanical engineering education*, 13(4): pp1-3, Aug.

Samanta, B, Al-Balushi, K.R. 2001. Artificial neural network based fault diagnostics of rolling element bearings using time domain features. *Mechanical systems and signal processing*, 17(2): pp317-328, Sept.

Scheffer, C, Heyns, P.S. 2001. Wear monitoring in turning operations using vibration and strain measurements. *Mechanical systems and signal processing*, 15(6): pp1185-1202, Apr.

Scheffer, C, Heyns, P.S, Kratz, H, Klocke, F. 2003. Development of a tool wear-monitoring system for hard turning. *International journal of machine tools and manufacture design research and application*, 43: pp973-985, Apr.

Shao, Y, Nezu, K. 1995. An on-line monitoring and diagnostic method of rolling element bearing with artificial intelligence. *Sapporo*, 3(1): pp1543-1548, Jul.

Sharkels, P. 2004. Murphy's Law watching for whatever can go wrong, because it will. *Mechanical engineering*, 126(5): pp36-41. May.

Shiroishi, J, Li, Y. 1997. Bearing condition diagnostics via vibration and acoustics emission measurements. *Mechanical systems and signal processing*, 11(5): pp693-705, Apr.

Shigley, J.E. 1986. *Mechanical Engineering Design*. 1st metric edition. New York: McGraw-Hill. pp389-413.

Singer, C. 1954. *A history of technology*. Volume I-V. New York: AB Nordbok. pp1-30.

- SKF. 1992. *An introduction to vibration monitoring manual*. Denmark. pp2-8.
- SKF. 1996. *Bearing maintenance handbook*. Denmark. pp1-62.
- SKF. 2004. *Bearing maintenance handbook*. Denmark. p31.
- SKF, South Africa. 2004. Custom-designed bearings for conveyor idlers. *Materials handling and logistics today*: pp31-32, Jan.
- Sohoel, E. 1994. A complete set of tools for evaluating the running condition of operating rolling element bearings. *The SPM Method*: p2. Aug.
- SPM Instrument. 2000. *Basic facts on machine vibration*. Strangnas. p12.
- SPM Instrument. 1998. *Shock pulse method*. Strangnas. p8.
- StatSoft, Inc. 1984-2003. Neural networks. [Online]. Available at: <http://www.Documents and Settings\Default\Desktop\Networks\Application of Neural Networks.htm>>. Accessed: 08/03/2004.
- Stolarski, T.A.1990. *Tribology in machine design*. London: Heinemann Newnes. pp1-50.
- Strandh, S. 1982. *Machines an illustrated history*. Sweden: AB Nordbok. pp1-20.
- Stratonovitch, R.L. 1957. Dokl. Acad. Nauk. S.S.S.R. *Sov.phys.Dokl*, 2: pp416.
- Taylor, J.I., Wyndell Kirkland, D. 2004. *The Bearing Analysis Handbook* First edition. U.S.A: Vibration consultants, Inc. pp1-76.

Technical Associates of Charlotte, Inc. 1994. *Role of spike energy, high frequency detection (HFD), shock pulse method (SPM) and specification of their alarm levels at various speeds*. New York. pp1-23.

The Vibration Institute of South Africa. 2003. *Bearings*. Jhb. pp35-40.

Van der Merwe, N.T, Hoffman, A.J. 2002. The application of neural networks to vibrational diagnostics for multiple fault conditions. *Computer standards and interfaces*, 24: pp139-149, Apr.

Van Wyk, E.M.P. 1998. The application of neural networks and digital signal processing techniques for condition monitoring in a power station. M.Eng. Dissertation. Potchefstroom: PU for CHE. pp6-42.

Villmann, T, Der, R, Herrmann, M, Martinetz, T. (1997). Topology preservation in self-organising feature maps: exact definition and measurement. *IEEE Transaction Neural Networks*, 8(2): pp256-266. Mar.

Vinh, T.T. (1997:269) A new approach to impact testing instrumentation using an electronic slave hammer. *Mechanical Systems and Signal Processing*, 11(2): pp269-285, Oct.

Von der Malsburg, C. 1973. Self-organisation of orientation sensitive cells in the straight cortex. *Kybernetik*, 14: pp85-100.

Widrow, B, Lehr, M.A. 1990. 30 years of adaptive neural networks: perceptron madaline, and back propagation. *Proceedings of the IEEE*, 78(9): pp1415-1442.

Willshaw, D, von der Malsburg, C. 1976. How patterned neural connections can be set up by self-organisation. *Proceedings of the royal society of London*, series B 194: pp431-445.

Wong, M.L.D, Jack, L.B, Nandi A.K. 2005. Modified self-organising map for automated novelty detection applied to vibration signal monitoring. *Mechanical systems and signal processing*. [online]ws: pp. January 24. Available at: <http://www.elsevier.com/locate/jnlabr/ymssp>. Accessed: 17/03/2005.

Zhang Jin, Tatematsu Ryuji, Kanada Noboru, Kitamura Toshiya, Yamada Shinji.
Vesanto, J, Himberg, J, Alhoniemi, E, Parhankangas, J. 2000. *SOM toolbox for Matlab 5*. Helsinki University of Technology. Finland.pp1-58.

Zhang, S, Ganesan, R, Xistris, D .1995. Self-organising neural networks for automated machinery monitoring systems. *Mechanical systems and signal processing*, 10(5): pp517-532, Sept.