Intelligent Fault Diagnosis of Gearboxes and its Applications on Wind Turbines

by

Sajid Hussain

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy in Electrical and Computer Engineering

Faculty of Engineering and Applied Sciences
University of Ontario Institute of Technology
Oshawa, Ontario, Canada
February 2013

Dedicated to my parents and my wife
They mean a lot to me, here and hereafter.

Abstract

The development of condition monitoring and fault diagnosis systems for wind turbines has received considerable attention in recent years. With wind playing an increasing part in Canada's electricity demand from renewable resources, installations of new wind turbines are experiencing significant growth in the region. Hence, there is a need for efficient condition monitoring and fault diagnosis systems for wind turbines. Gearbox, as one of the highest risk elements in wind turbines, is responsible for smooth operation of wind turbines. Moreover, the availability of the whole system depends on the serviceability of the gearbox.

This work presents signal processing and soft computing techniques to increase the detection and diagnosis capabilities of wind turbine gearbox monitoring systems based on vibration signal analysis. Although various vibration based fault detection and diagnosis techniques for gearboxes exist in the literature, it is still a difficult task especially because of huge background noise and a large solution search space in real world applications. The objective of this work is to develop a novel, intelligent system for reliable and real time monitoring of wind turbine gearboxes. The developed system incorporates three major processes that include detecting the faults, extracting the features, and making the decisions. The fault detection process uses intelligent filtering techniques to extract faulty information buried in huge background noise. The feature extraction process extracts fault-sensitive and vibration based transient features that best describe the health of the gearboxes. The decision making module implements probabilistic decision theory based on Bayesian inference. This module also devises an intelligent decision theory based on fuzzy logic and fault semantic network.

Experimental data from a gearbox test rig and real world data from wind turbines are used to verify the viability, reliability, and robustness of the methods developed in this thesis. The experimental test rig operates at various speeds and allows the implementation of different faults in gearboxes such as gear tooth crack, tooth breakage, bearing faults,

and shaft misalignment. The application of hybrid conventional and evolutionary optimization techniques to enhance the performance of the existing filtering and fault detection methods in this domain is demonstrated. Efforts have been made to decrease the processing time in the fault detection process and to make it suitable for the real world applications. As compared to classic evolutionary optimization framework, considerable improvement in speed has been achieved with no degradation in the quality of results. The novel features extraction methods developed in this thesis recognize the different faulty signatures in the vibration signals and estimate their severity under different operating conditions. Finally, this work also demonstrates the application of intelligent decision support methods for fault diagnosis in gearboxes.

Declaration

No part of this thesis has been submitted elsewhere for any other degree or qualification.

The content of this thesis is all my own work unless referenced to the contrary in the text.

Acknowledgements

I would like to thank and offer my sincerest gratitude to my thesis supervisor, Dr. Hossam A. Gabbar, for his continuous support and motivation throughout this thesis. I cannot imagine having any better advisor for my doctoral study regarding his knowledge, perceptiveness, and encouragement.

I must express my sincere gratitude and indebtedness to Dr. Sarbast Rasheed at University of Ontario Institute of Technology for his valuable feedback and kind guidance in writing this thesis.

I am indebted, forever, to my parents and my lovely wife for their patience, understanding, and endless encouragement during the course of this study.

Sajid Hussain
Oshawa, Ontario, Canada
February, 2013

List of Figures

List of Figures

List of Tables

List of Abbreviations and Acronyms

MW	Mega Watts
GHG	Greenhouse Gas
Mt CO_2 eq	Mega Tonnes of Carbon Dioxide Equivalent
OECD	Organization for Economic Cooperation and Development
LWK	LandWirtschaftsKammer: Chamber of Agriculture of North Rhine-Westphalia, Germany
WMEP	Wissenschaftliches Mess-und Evaluierungsprogramm (in German): Scientific Measurement and Evaluation Programme
ISET	Institut für Solare Energieversorgungstechnik (in German)
CM-FDS	Condition Monitoring and Fault Diagnosis System
FSN	Fault Semantic Network
FES	Fuzzy Expert System
BBN	Bayesian Belief Network
NDT	Non-Destructive Testing
TSA	Time Synchronous Averaging
FFT	Fast Fourier Transform
SNR	Signal-to-Noise Ratio
ED	Enhanced Diagnostic
NN	Neural Network
ANFIS	Adaptive Neuro-Fuzzy Inference System
NREL	National Renewable Energy Laboratory
GRC	Gearbox Reliability Collaborative
DFT	Dynamometer Test Facility
kW	Kilo Watts
RPM	Revolution per Minute
HSS	High Speed Shaft
PHM	Prognostics and Health Management
Hz	Hertz

List of Abbreviations and Acronyms

KHz	Killo Hertz
SCADA	Supervisory Control and Data Acquisition
STFT	Short Time Fourier Transform
WA	Wavelet Analysis
WVD	Wigner-Ville Distribution
GA	Genetic Algorithms
SA	Simulated Annealing
GP	Genetic Programming
NMF	Non-Negative Matrix Factorization
EGA	Enhanced Genetic Algorithm
RCGA	Real Coded Genetic Algorithm
PSO	Particle Swarm Optimization
SVM	Support Vector Machine
DAG	Directed Acyclic Graph
CBM	Condition Based Maintenance
GPM	General Path Model
RUL	Remaining Useful Life
TOF	Time of Failure
CM	Condition Monitoring
SHM	Structural Health Monitoring
PSD	Power Spectral Density
GB	Giga Bytes
RAM	Random Access Memory
Ghz	Giga Hertz
SRS	Shock Response Spectrum
R&D	Research and Development
SDOF	Single Degree of Freedom
GMF	Gear Mesh Frequency

List of Abbreviations and Acronyms

ERB	Equivalent Rectangular Bandwidth
NARX	Nonlinear Autoregressive Model with Exogenous Inputs
TDNN	Time Delay Neural Network
MF	Membership Function
MAE	Mean Absolute Error
MSE	Mean Square Error
POOM	Process Object Oriented Methodology
FM	Failure Mode
PCA	Principal Component Analysis
MLP	Multi-Layer Perception

Nomenclature

F_s	Sampling frequency.
n	Total number of samples in a signal.
μ_4	Fourth central moment of a distribution.
$x_t(i)$	i^{th} sample of a signal x_t.
μ	First central moment of a distribution or mean of a distribution.
σ	Standard deviation of a distribution.
$c1, c2$	Constraints on objective function.
η	Positive penalty multiplier.
p_i	Penalty functions.
δ	Multiplier.
F_c	Centre Frequency of band pass filter.
Q	Quality factor of band pass filter.
R_p	Ripples in pass band of a band pass filter.
N	Order of a band pass filter.
Fs_1, Fs_2	Stop frequencies for a band pass filter.
Fp_1, Fp_2	Pass frequencies for a band pass filter.
x^*	Optimal point in one-dimensional search space.
$x^{(lo)}$	Lower bound one-dimensional search space.
$x^{(hi)}$	Upper bound one-dimensional search space.
$x^{(1)}, x^{(2)}$	Intermediate points in one-dimensional search space.
$x^{(mid)}$	Middle points in one-dimensional search space.
α	Golden section ratio.
$x^{(qu)}$	Unique optimum of the quadratic function.
$F^{(lo)}$	Evaluation of function $F(.)$ at point $x^{(lo)}$.

$F^{(mid)}$	Evaluation of function $F(.)$ at point $x^{(mid)}$.
$F^{(hi)}$	Evaluation of function $F(.)$ at point $x^{(hi)}$.
ε	Stopping criteria.
$\psi(t)$	Mother wavelet.
a	Scale parameter of a wavelet.
β	Shape parameter of a wavelet.
b	Time translation of a wavelet.
$\psi(f)$	Fourier transform of $\psi(t)$.
$x(t)$	Time signal.
$X(f)$	Fourier transform of $x(t)$.
$F^{-1}(.)$	Inverse Fourier transform.
$C_{i,j}^{Gen}$	Chromosome i or j in generation Gen.
γ	Weighting factor for arithmetic crossover or arithmetic mutation.
ξ	Maximum allowable ripples in pass band of a band pass filter.
$\mathbf{x}_i$	Position vector in particle swarm optimization.
$\mathbf{v}_i$	Velocity vector in particle swarm optimization.
d_1, d_2	Self-confidence and swarm-confidence factors.
r_1, r_2	Randomly generated, uniformly distributed parameters in the range [0 1].
w	Inertia weight.
T	Current iteration time or sampling time.
M	Mass of single degree of freedom system.
k	Spring stiffness.
c	Damping coefficient.
f_N	Resonance frequency.
ζ	Damping ratio.
ω_d	Damped natural frequency.
ω_n	Natural frequency.

q	Number of calculation frequencies.
f_l	Lower frequency.
f_u	Upper frequency.
$G_{\max}$	Maximum shock value.
G_{ref}	Reference shock value.
N_B	Number of blocks.
N_C	Number of channels.
S_r	Rise slope.
S_f	Fall slope.
F_{SR}	Slope rise feature.
F_{SF}	Slope fall feature.
F_A	Slope amplitude feature.
$\hat{x}(t)$	Hilber transform of $x(t)$.
$u(t)$	Complex signal of $x(t)$.
$A(t)$	Amplitude of the complex signal.
$\theta(t)$	Angle of the complex signal.
$f(t)$	Force acting on spring.
K_s	DC gain.
δ	Logarithmic decrement.
S_p	Peak position on time axis.
p	Index representing current peak.
P	Total number of peaks.
T_p	Average pulse repetition time.
f_c	Bearing cage frequency.
f_{bd}	Bearing ball defect frequency.
f_{id}	Bearing inner defect frequency.

f_{od}	Bearing outer defect frequency.
α	Contact angle.
n_b	Number of balls in bearing.
d	Roller diameter.
D	Pitch diameter.
$h(t)$	Impulse response of a filter.
R	Normalization constant.
m	Duration of impulse response.
ϕ	Phase of the tone or impulse response signal.
Vb_S	Vibration steepness index.
S_{max}	Maximum steepness value.
Vb_A	Vibration amplitude index.
A_{max}	Maximum amplitude value.
a_o, a_i, b_i	Coefficients of approximation equation to approximate $\hat{f}(x)$ from $f(x)$.
κ	Constant in calculating child wavelet.
χ	Predictor function for predicting time series.
R_q	q^{th} fuzzy rule.
A_q	q^{th} antecedent.
C_q	q^{th} consequent class.
μ_{ik}	Membership values of kth data point in the ith cluster space.
m'	Weighting parameter for fuzzy clustering.
U	Fuzzy partition matrix.
v	Clusters centres matrix.
d	Similarity matrix.
m	Number of features (Chapter 7).
a_o, a_1	Regression coefficients (Chapter 7)

$s_{y/x}$	Standard deviation for regression line.
S_t	Sum of squares around the mean.
r^2	Coefficient of determination.

CHAPTER 1

Introduction

There is a growing interest in renewable energy systems with increased concerns over climate change. Wind energy has an attractive share in renewable energy because it diversifies a resource portfolio and improves overall reliability of the power system. However, the engineering challenge for the wind industry is to design an efficient wind turbine to harness wind energy and turn it into electricity. Wind turbine technology has proved itself over the last 20 years and is developing rapidly. The electricity generation has become more affordable and cost-effective with the advent of wind turbines as large as 7.5 megawatts (MW) in capacity as shown in Figure 1.1 [1]. Despite all technological advancements in wind turbine design and installation, there is a price to pay in terms of harsh operating environments and reduced accessibility. In this chapter, we briefly introduce the fundamental concerns about climate; discuss the growing interests in renewable energy resources, especially the wind and provide a short discussion about the wind turbines and their reliability. Finally, an outline of the thesis is given with experimental and real world data collection setup.

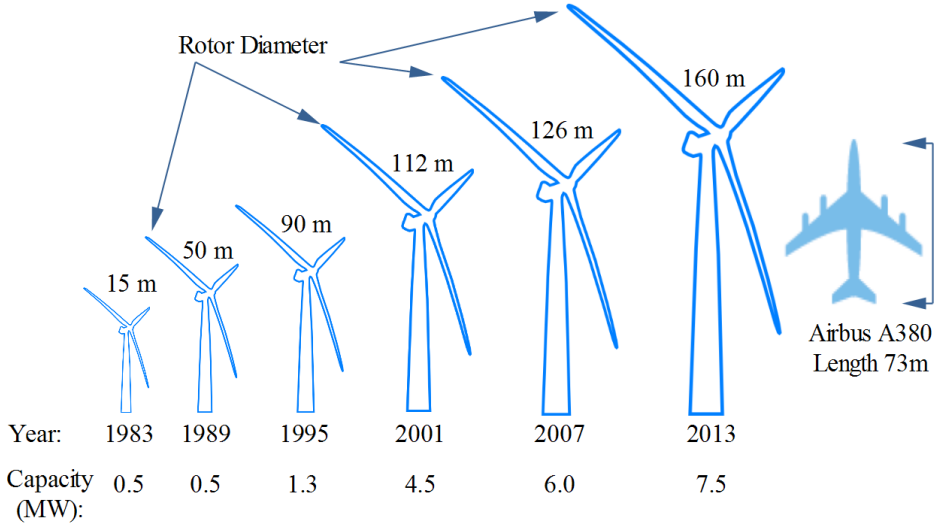


Figure 1.1: Increase in Wind Turbine Size and Capacity [1]

1.1 Climate Concerns

There have been increasing concerns about the global warming, greenhouse gases and climate change over the past couple of decades. One of the major culprits in global warming is the generation of coal-fired electricity. According to environment Canada report [2], Canada's total greenhouse gas (GHG) emissions were estimated to be 692 mega tonnes of carbon dioxide equivalent (Mt CO_2 eq) in 2010. This represents an increase of approximately 2 Mt (0.25%) from the 2009 level of 690 Mt. Although, the GHG emissions have been decreased by 48 Mt (6.5%) since 2005, this is still 17% higher than the target values of 607 Mt set by OECD, the Organisation for Economic Co-operation and Development. In 2010, Canada ranked 15th out of 17 OECD countries on GHG emissions per capita and scored "D" grade on the scale of "A to D". The electricity production sector contributes 14% in GHG emissions as shown in Figure 1.2 [2].

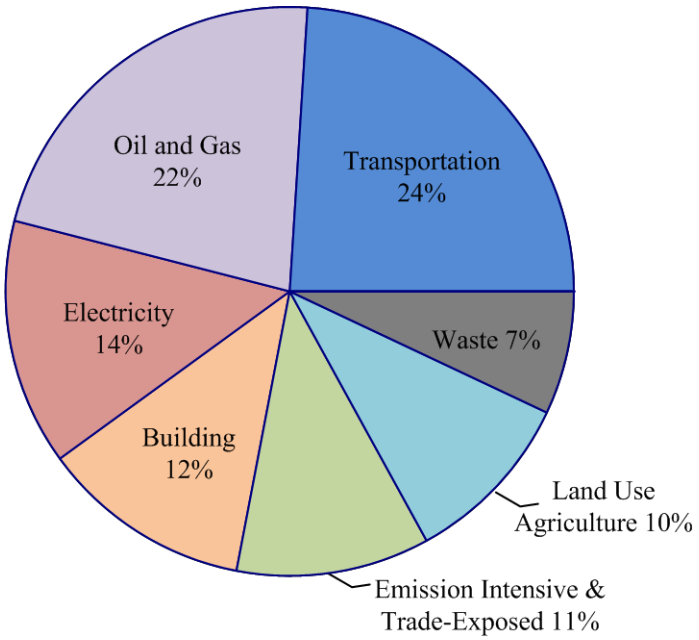


Figure 1.2: Green House Gases Emissions by Sector – Canada [2]

1.2 Growth of Wind Energy in Canada

Canada is committed to tackle the climate change by building a low-carbon economy and reducing the GHG emission levels by 17% till 2020. Canada's electricity production sector is one of the cleanest in the world. However, Canada has taken further steps

towards developing an even cleaner electricity grid. Canada has moved forward on tough rules for coal-fired electricity sector. In Canada, coal-fired electricity generation is responsible for 77% of the total GHG emissions from the electricity and heat sector.

Renewable energy resources are alternatives of the coal-fired electricity generation. Wind energy is one of the cleanest and renewable energy resources available throughout the world. The wind energy contributes only 1.3% of the total energy demand of Canada as shown in Figure 1.3 [3].

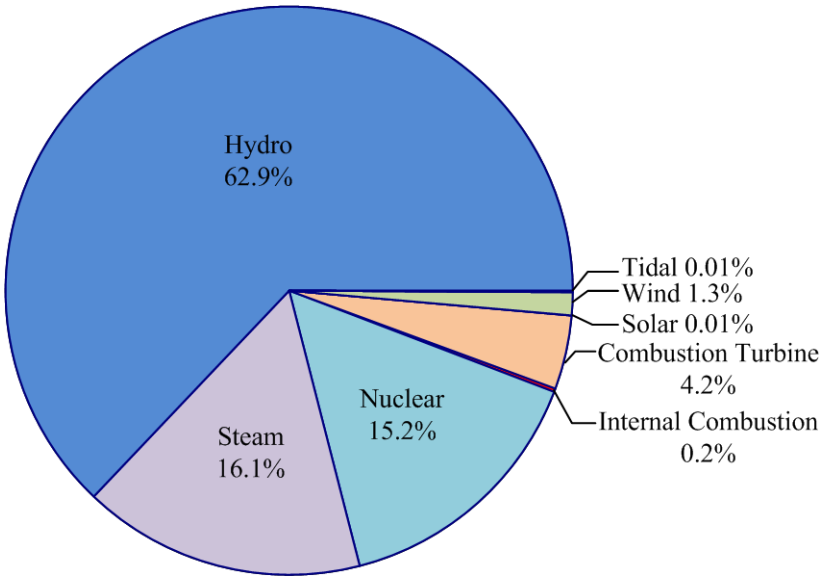


Figure 1.3: Electricity Generation Resources – Canada [3]

According to Canadian Wind Energy Association, Canada is expecting to fulfill 20% of its total energy demand through wind resources by the year 2025. The wind power generation in Canada has increased exponentially for the last decade as shown in Figure 1.4. In the Ontario province alone, the total installed wind power capacity is around 2000 MW and it is expected to reach 7600 MW by the year 2018. In the year 2011, Ontario built 8 wind farms with 251 wind turbines with a total capacity of 522 MW [4].

1.3 Reliability of Wind Turbines

As shown in Figure 1.1, the wind turbines are increasing in size and capacity with advancement in technology. Also, due to the fact that rich wind resources available

offshore, there is a significant move towards offshore wind energy installation. This poses significant technological challenges regarding wind turbine's reliability and maintenance.

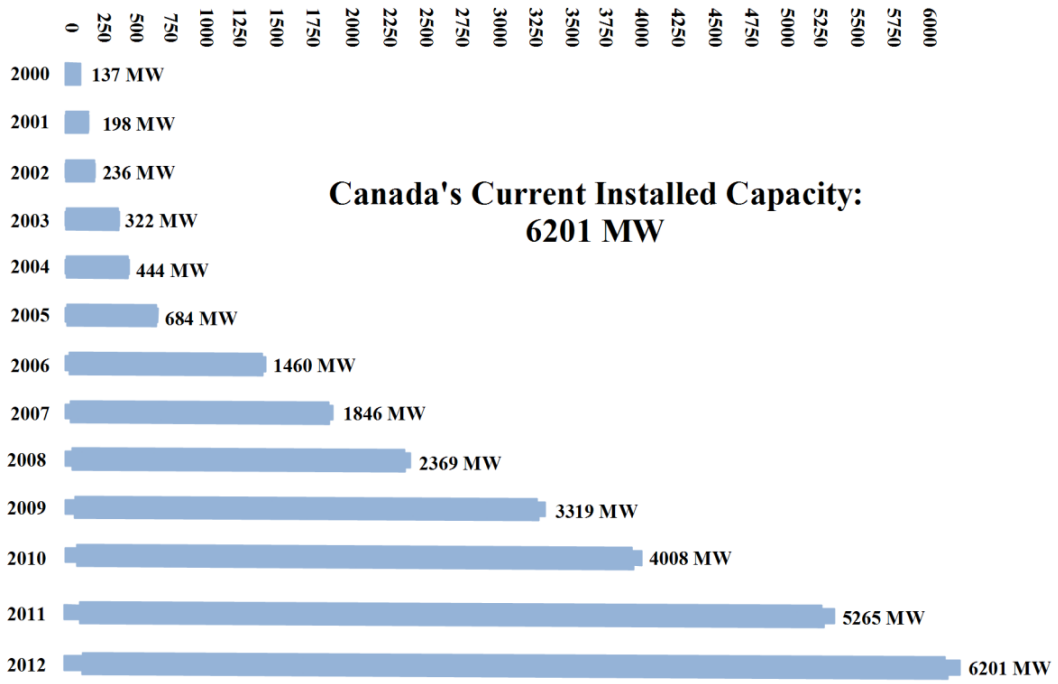


Figure 1.4: Canada's Current Installed Wind Energy Capacity [4]

The first challenge is the increase in wind speeds offshore that will lead to much higher mechanical loadings on wind turbine's components. The second challenge is the limitation of maintenance accessibility and the difficulties encountered due to severe weather conditions. Maintenance of a small component in a wind turbine is relatively easy if the turbine is installed onshore but it carries a very high cost if the turbine is operated offshore. Also the weather conditions impose restrictions on accessibility of an offshore wind turbine and a relatively simple problem can cost several days of downtime and associated tangible and intangible costs of maintenance. There are several studies carried out in recent years about the reliability of the wind turbines. In [5], authors report onshore wind turbines failure rate of around 1-3 failures per turbine per year using data from Germany and Denmark. The average failure rate of a wind turbine has a tendency to increase with an increase in its size and rating as shown in Figure 1.5. Figure 1.5 is a result of a survey conducted by LandWirtschaftsKammer (LWK) [1]. The LWK survey consists of a population of 158-643 (over the period) wind turbines of average age up to

15 years. The population includes both geared and direct drive configuration wind turbines.

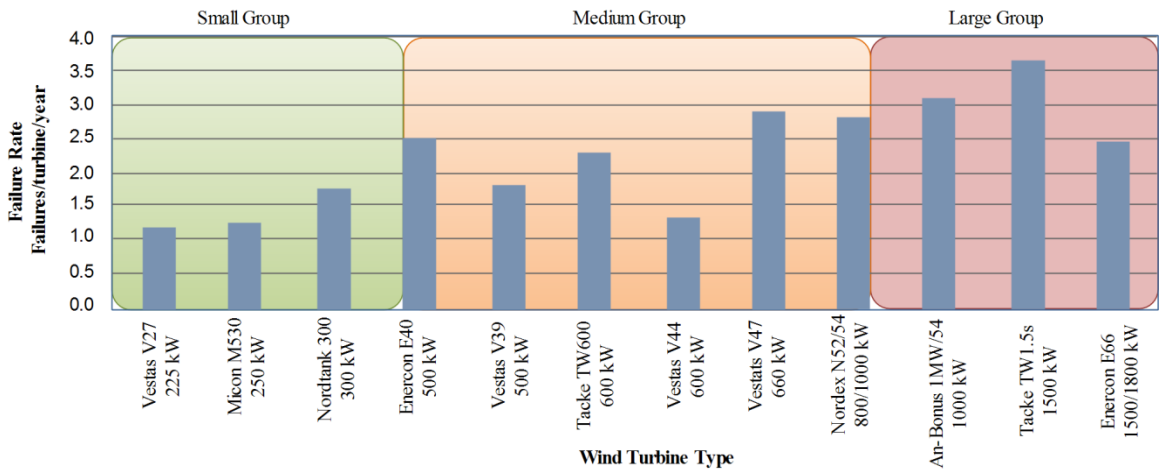


Figure 1.5: Failure Rates of Wind Turbine Models by Capacity [1]

Another German wind power survey called Scientific Measurement and Evaluation Programme (WMEP) [6] conducted by Institut für Solare Energieversorgungstechnik (ISET) lists different subassemblies in a wind turbine and their failure probabilities. In Figure 1.6 [1], failure rates and downtimes for each failure type are listed from two large surveys of European wind turbines over 13 years (LWK and WMEP). Figure 1.6 shows higher failure rates for electrical systems in wind turbines and large downtime for gearbox failures. Using a simple relation between risk, failure rate and consequences, the gearbox failure is one of the highest risk events of wind turbines.

1.4 Fault Diagnosis

In any mechanical process, the detection of faults is the first and foremost priority of a maintenance function before diagnosis, prognosis, and cause or consequence analysis. Detection and diagnosis is inevitable to solve problems in maintenance and operations. Many detection techniques are able to detect different faults. However, most of these techniques suffer from low detection quality and/or slow response time, which restrict their use in online or real time fault detection applications. Diagnosis is not an easy task in a complex rotating machinery because there could be many factors involved. These factors include type, frequency, number of alarms, effectiveness of the fault data presentation, and the time delay in fault detection. All these factors make the fault

diagnosis a difficult and challenging task. Hence, there is a requirement for a real time automated condition monitoring and fault diagnosis system (CM-FDS) to aid in decision making.

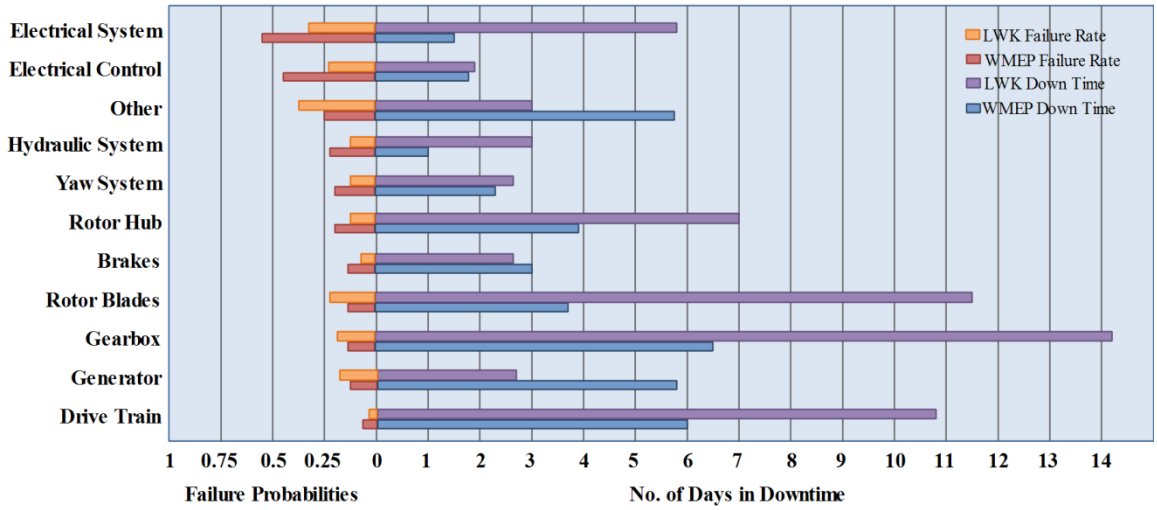


Figure 1.6: Wind Turbine Sub-assemblies Failure Rate and Down Time [1]

Since, gearbox is a critical component in any mechanical process involving rotating machines such as wind turbines; it should be monitored effectively and efficiently. In a wind turbine system, typically the gearbox is responsible for around 15-20% of its maintenance and downtime costs [7]. The average preventive maintenance and repair cost of a gearbox in a wind turbine costs around $50,000 CAD in every 7-8 years of its operational life [8]. The cost of unexpected gearbox failure is even higher. As a specific example, according to the World Wind Energy Association, a gearbox of a wind turbine developed a severe fault in one of the wind farms in Canada in 2004 because of no condition based maintenance program in place. The repair and replacement cost was around $426000 CAD. This catastrophe could have been avoided with proper predictive and condition monitoring systems in place [9].

1.5 Condition Monitoring and Fault Diagnosis System (CM-FDS)

This section presents a condition monitoring and fault diagnosis system (CM-FDS) for wind turbine gearboxes. The system incorporates modules such as fault detection, diagnosis, prognosis, and root-cause and consequence analysis along with maintenance

work and design change recommendations as shown in Figure 1.7. The purpose of this research is to present vibration based, real time, fault detection and diagnosis methods for wind turbine gearboxes. In Figure 1.7, the fault detection module uses vibration signals emanating from mechanical drive train of a wind turbine and intelligently searches for any possible fault buried under huge background noise. Adaptive techniques combined with optimization methods are used in this module for tuning the detection filters and satisfying an objective function. The fault detection step is a critical component in an overall diagnosis and decision making process as a poor detection quality may affect the results. Fault diagnosis module in Figure 1.7 incorporates extraction of representative health features, classification of different faults, and identification of faulty components. This module classifies different types of faults and makes the fault diagnosis decisions either comparing different clusters with information available in a database or estimating the faulty component's rotation frequency from the filtered signals. Fault prognosis module uses de-noising and soft computing techniques to model the dynamics of the system and predicts the system's behaviour in future. Decision making module uses Bayesian belief network (BBN), fuzzy expert system (FES), and fault semantic networks (FSN) to reason causes behind a particular fault. This module either learns the BBN, the FES or the FSN from historical data or the parameters are defined by an expert in this field.

1.6 Problem Definition

Typically, faults in wind turbine gearboxes arise while they are in operation. Therefore, it is vital to detect, diagnose and analyse these faults as early as possible. The process should be non-destructive in nature to avoid wind turbine's disassembly. It is pivotal to use non-destructive testing (NDT) techniques such as vibration analysis but vibration analysis generally suffers from various problems.

1.6.1 Background Noise

Vibration signals coming from wind turbines contain a huge amount of background noise because of the dynamic environment and severe weather conditions. Time domain averaging is normally performed to mediate the effects of the background noise like time

synchronise averaging (TSA) [10;11] or an inherent averaging present in fast Fourier transform (FFT) but at the expense of losing the important information present in the signal, especially, low energy faults. So, there is a trade-off between noise filtering through the averaging process and quality of results.

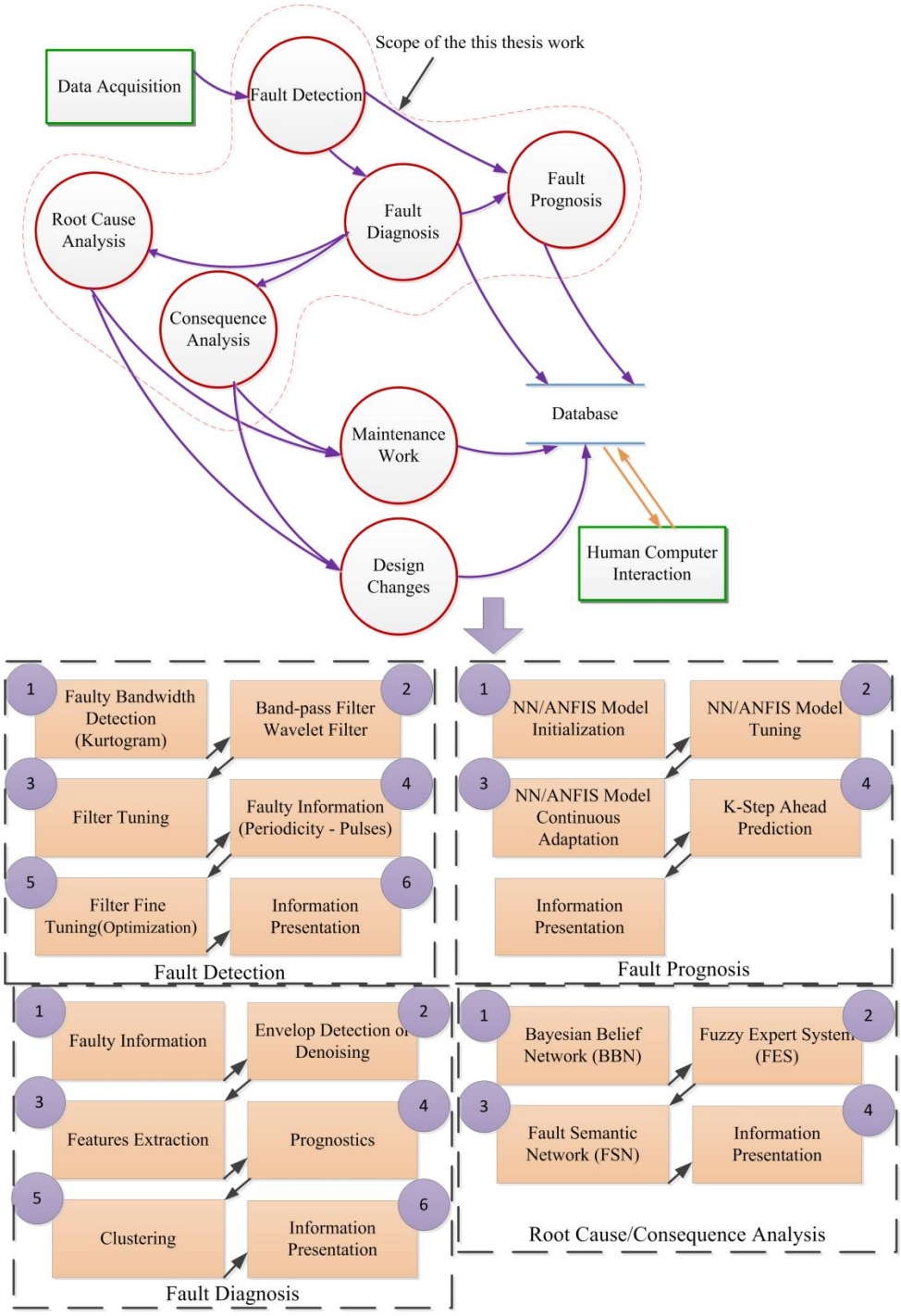


Figure 1.7: The Condition Monitoring and Fault Diagnosis System (CM-FDS)

1.6.2 Real Time Fault Detection

Time domain filtering is performed to avoid the averaging process, and evolutionary algorithms are used for tuning the filters [12;13]. The process generally improves the detection capability and quality of results but at the expense of computation time. Therefore, this technique is normally not suitable in real time fault detection applications.

1.6.3 Early Fault Detection and Diagnosis

Many critical mechanical processes like wind turbines require an early detection of faults. A satisfactory operation of these processes requires reliable detection and monitoring techniques for sub-systems like rotors, gearboxes and generators. The early fault detection enables timely repair of the system and avoids further deterioration. Methods based on evolutionary algorithms generally perform well in detection of early faults but at the expense of computation time [14]. Hence, their use is difficult for online real time diagnosis systems.

1.7 Research Objectives

To date, there is no such approach to address all the issues discussed in the previous section simultaneously. If efforts are made to improve the detection quality, the time complexity is compromised and vice versa. Also, in the diagnosis domain, no such features extraction technique exists in open research where an early indication of faults development can be seen. The primary objective of this thesis is to develop an integrated and intelligent fault diagnosis system that incorporates a trade-off between detection quality and time complexity. The system should be able to detect faults early in time, extracts faulty features from the vibration signals that best represent the health of the system under observation, predicts the faulty conditions and performs cause and consequence analysis. As shown in Figure 1.7, the proposed intelligent system consists of four main modules, fault detection through intelligent adaptive filtering, feature extraction and clustering through different signal processing techniques, prognostics and decision making through BBN, FES or FSN. The strategy is to develop advanced and

more robust techniques at each processing stage to improve the reliability of wind turbine condition monitoring systems. The specific objective's summary is as follows.

(1) **Intelligent Filtering**: To develop optimized and intelligent filtering techniques for reliable, vibration based fault detection in wind turbines. The methods should be able to increase the signal to noise ratio (SNR).

(2) **Features Extraction:** To extract representative features those best describe the health of the system under test. The extracted features should be sensitive to fault conditions and robust to random variations because of noise.

(3) **Clustering and Prediction:** To perform clustering and prediction on the extracted features for diagnosis of different types of faults and predicting the system response in future.

(4) **Fault Diagnosis:** To develop an enhanced diagnostic (ED) scheme for automatic decision-making. The suggested ED scheme consists of the FSN module along with BBN and FES for cause and consequence analysis.

1.8 Research Methodology

The framework of the research methodology consists of different filtering techniques combined with conventional and evolutionary optimization algorithms for fault detection. Different novel features extraction algorithms are also proposed that best represent the healthy conditions of the system under observation. The BBN, FES and FSN are investigated for root-cause and consequence analysis or decision making. Figure 1.8 shows a summary of the proposed research methodology. Wavelet analysis and band-pass filtering combined with evolutionary and conventional optimization algorithms form the basis of the background noise removal from vibration signals. A psychoacoustic filtering phenomenon is also included for this purpose. However, psychoacoustic filters are static in nature. Three different novel features extraction techniques are presented in Figure 1.8. Under the clustering and prediction framework, fuzzy c-means clustering, neural

networks (NN) and adaptive neuro-fuzzy inference system (ANFIS) are implemented. FSN, FES and BBN form the basis of cause and consequence analysis module.

1.9 Expected Value of the Proposed Research

The proposed research is closely related to the maintenance industry and market requirement. Mechanical processes must be monitored through reliable condition monitoring technologies in order to reduce operation and maintenance costs and at the same time avoid catastrophic failures that could lead to assets damage, accidents and loss of human lives. It is expected that the developed methods can be used for identifying different vibration based faults in wind turbine gearboxes, where they provide early warnings for incipient faults and formulate a framework for fault diagnosis. The developed methods can also be applied to detect vibration based faults in other mechanical processes. The original contribution of this research work is enhancement of existing condition monitoring and fault diagnosis systems (CM-FDS) for wind turbine gearboxes. The development process is based on the application of advanced signal processing techniques and the non-stationary operating conditions of the wind turbines.

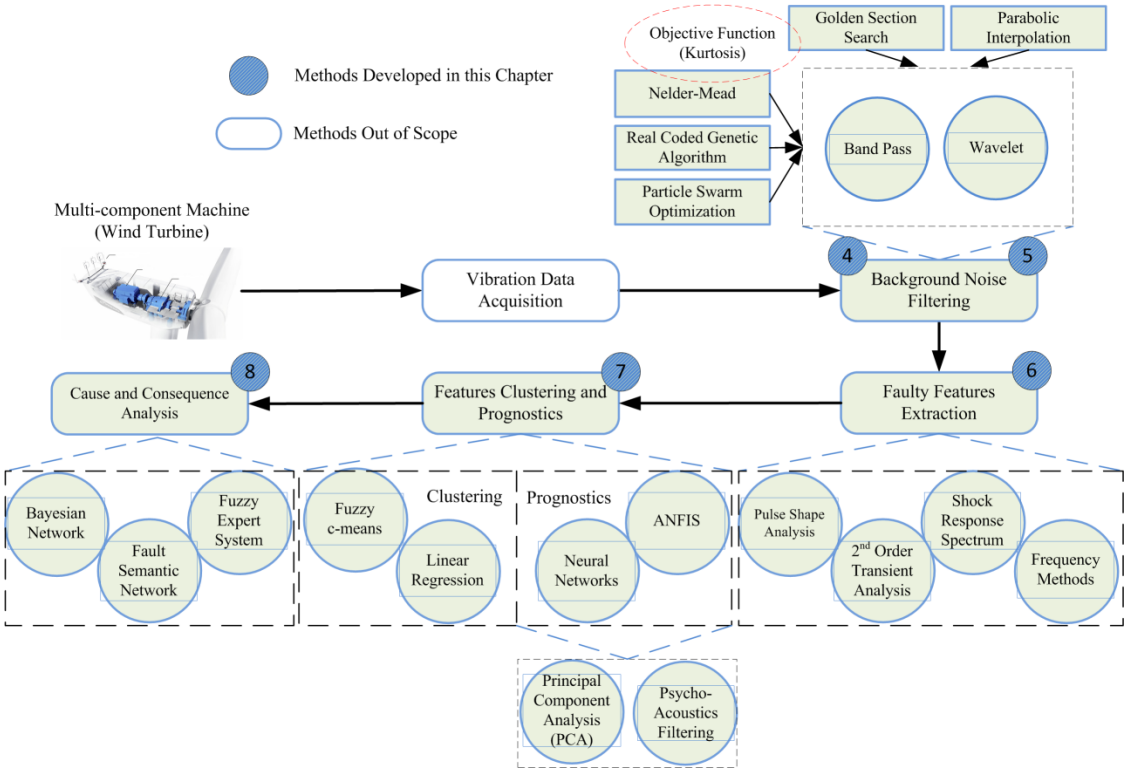


Figure 1.8: The Proposed Research Methodology

1.10 Vibration Data Collection

In this thesis we use the following data sets from experimental test rigs as well as industry.

1.10.1 Dataset: National Renewable Energy Laboratory

The data used in this section emanate from a planetary gearbox inside a wind turbine. The data are provided by the National Renewable Energy Laboratory (NREL), through a consortium called the Gearbox Reliability Collaborative (GRC) [15]. The gearbox under test is one of the two units taken from the field and redesigned, rebuilt and instrumented with over 125 sensors. The gearbox first finished its run-in in the NREL dynamometer test facility (DTF) and later was sent to a wind plant close to NREL for field test, where two oil losses occurred. The test turbine in the field is a stall-controlled, three-bladed, upwind turbine with a rated power of 750kW. The turbine generator operates at 1200 RPM and 1800 RPM nominal on two different sets of windings depending on the power. The planetary gearbox has an overall ratio of 1:81.491. It is composed of one low speed (LS) planetary stage and two parallel stages as shown in Figure 1.9 [16]. This work uses data from test case CM_2a with main shaft speed of 14.72 RPM and high speed shaft (HSS) speed of 1200 RPM as listed in Table 1.1. The data were collected for the duration of 10 minutes at the sampling frequency of $F_s = 40\text{KHz}$.

Table 1.1: Vibration Data Sets NREL

Test Case	Main Shaft Speed (rpm)	Nominal HSS Speed (rpm)	Electric Power (% of rated)	Record Duration (min)
CM_2a	14.72	1200	25%	10
CM_2b	22.09	1800	25%	10
CM_2c	22.09	1800	50%	10

The data files are provided in a Matlab® packed binary format for direct import into Matlab® (*.mat). Header information is included that identifies signals (variables). The data are divided into ten one minute data sets for each test case listed and described in

Table 1.1. Files are labelled, for example, as CM_2a_01.mat for the first one minute of test case CM_2a.

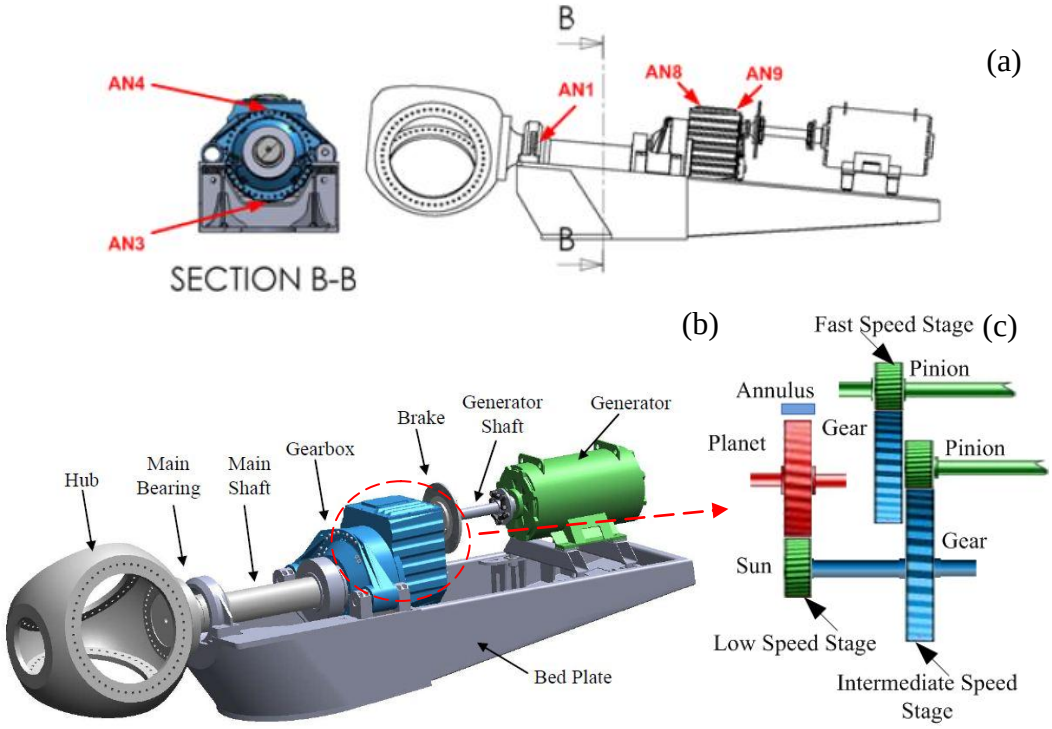


Figure 1.9: Data Collection Setup (a) Sensor Locations (b) GRC Drive train Configuration (c) Planetary Gearbox [15]

1.10.2 Dataset: Prognostics and Health Management Society

Data collected in this section come from public data sets distributed by Prognostics and Health Management (PHM) Society under 2009 PHM challenge competition [17]. The data are representative of a generic industrial gearbox shown in Figure 1.10. Data were sampled synchronously from accelerometers mounted on both the input and output shaft retaining plates of the gearbox. An attached tachometer generates 10 pulses per revolution providing very accurate zero crossing information. Data were collected at 30, 35, 40, 45 and 50 Hz shaft speed under high and low loading. Additionally, different repeated runs are included in the data, although the run time and load were not sufficient to induce significant fault progression. There are a total of 560 samples to be classified. Data are provided in *.csv file format and collected with different faults and their combinations, as listed in Table 1.2.

Table 1.2: Vibration Data Sets PHM Society [17]

IS: Input Shaft, ISd: Input Side, ID: Idler Shaft, OS: Output Shaft, OSd: Output Side
Bearing elements: 8, R.E. diameter: 0.3125, Pitch diameter: 1.319, Contact angle: 0
RPM: 1800, 2100, 2400, 2700, 3000, Load: Low, High
Sampling frequency: 66.6667KHz, Unit: m/s^2, Each acquisition time: 4 sec
No. of acquisitions: Two acquisitions per RPM per Load

	Gears				**Bearings**						**Shafts**	
Case	**32T**	**96T**	**48T**	**80T**	**IS:ISd**	**ID:ISd**	**OS:ISd**	**IS:OSd**	**ID:OSd**	**OS:OSd**	**Input**	**Output**
Spur1	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Spur2	Chipped	Good	Eccentric	Good	Good	Good	Good	Good	Good	Good	Good	Good
Spur3	Good	Good	Eccentric	Good	Good	Good	Good	Good	Good	Good	Good	Good
Spur4	Good	Good	Eccentric	Broken	Ball	Good	Good	Good	Good	Good	Good	Good
Spur5	Chipped	Good	Eccentric	Broken	Inner	Ball	Outer	Good	Good	Good	Good	Good
Spur6	Good	Good	Good	Broken	Inner	Ball	Outer	Good	Good	Good	Imbalance	Good
Spur7	Good	Good	Good	Good	Inner	Good	Good	Good	Good	Good	Good	Keyshear
Spur8	Good	Good	Good	Good	Good	Ball	Outer	Good	Good	Good	Imbalance	Good
Case	**16T**	**48T**	**24T**	**40T**	**IS:ISd**	**ID:ISd**	**OS:ISd**	**IS:OSd**	**ID:OSd**	**OS:OSd**	**Input**	**Output**
Helical1	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
Helical2	Good	Good	Chipped	Good	Good	Good	Good	Good	Good	Good	Good	Good
Helical3	Good	Good	Broken	Good	Good	Good	Good	Combine	Inner	Good	Bent	Good
Helical4	Good	Good	Good	Good	Good	Good	Good	Combine	Ball	Good	Imbalance	Good
Helical5	Good	Good	Broken	Good	Good	Good	Good	Good	Inner	Good	Bent	Good
Helical6	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good

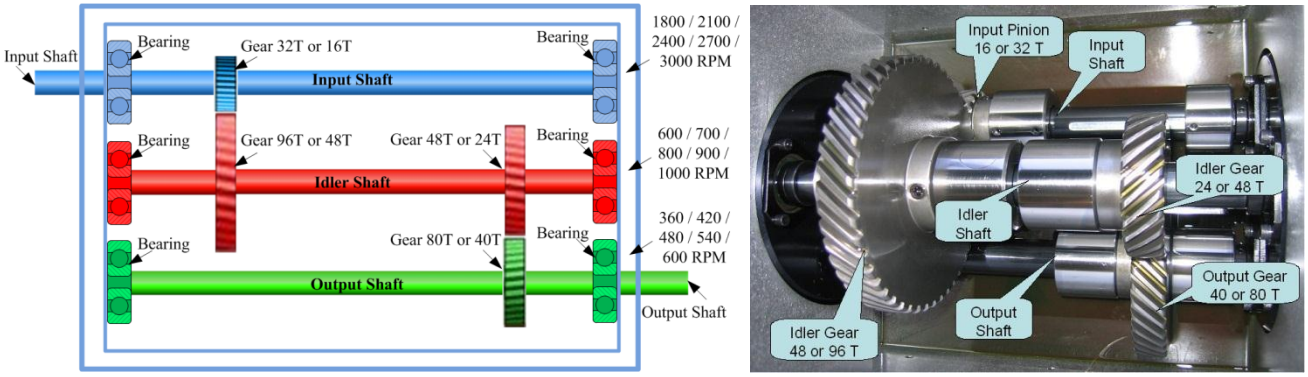


Figure 1.10: PHM Society Gearbox Data Collection Setup [17]

1.10.3 Datasets: Real World Wind Turbine

The real world datasets used in this work are collected from supervisory control and data acquisition (SCADA) module from two wind turbines. One dataset are spanned over two years (May, 2010 – April, 2012) and another dataset are spanned over a period of three

months (June, 2012 – September, 2012). The data were collected at minimum of 10-minute intervals. However, the data collection is triggered, and the interval can increase depending on the state of the wind turbine. Different types of data were collected that include:

(a) **Wind parameters:** Wind parameters include the direct measurements of the wind including wind speed, wind direction etc.

(b) **Energy conversion parameters**: Energy conversion parameters include, torque, blade pitch angle, rotor speed and power output etc.

(c) **Vibration parameters**: Vibration parameters include drive train accelerations and tower accelerations etc.

(d) **Temperature parameters**: Temperatures measured at different turbine's components include in this category. Parameters like bearings' temperatures and nacelle interior temperatures are some examples.

The vibration data were sampled at 23.67 KHz. In dataset1, the vibration data belong to Channel1: Generator Drive End (Bearing) and in dataset2, the vibration data belong to Channel7: Planetary Stage (Gearbox) as shown in Figure 1.11(b). In dataset1, maintenance was performed on bearing at generator's drive end and a new bearing was replaced afterwards. In dataset2, there was a failure of planetary gearbox inside the wind turbine. Planetary gearbox is typically used in applications requiring a large reduction or increase in speed at high loads, such as helicopters and wind turbines. Generally, a planetary gearbox has three or more planet gears each meshing with a sun and a ring gear. Drive is provided via the sun gear and the ring gear is stationary. Figure 1.11(a) is an illustration of a single stage planetary gear set which consists of a sun gear, a ring gear and four planet gears. In Figure 1.11(b), the sensors and their corresponding vibration structures are as follows:

Channel1. Generator Drive End (Bearing)
Channel2. Generator Non Drive End (Bearing)

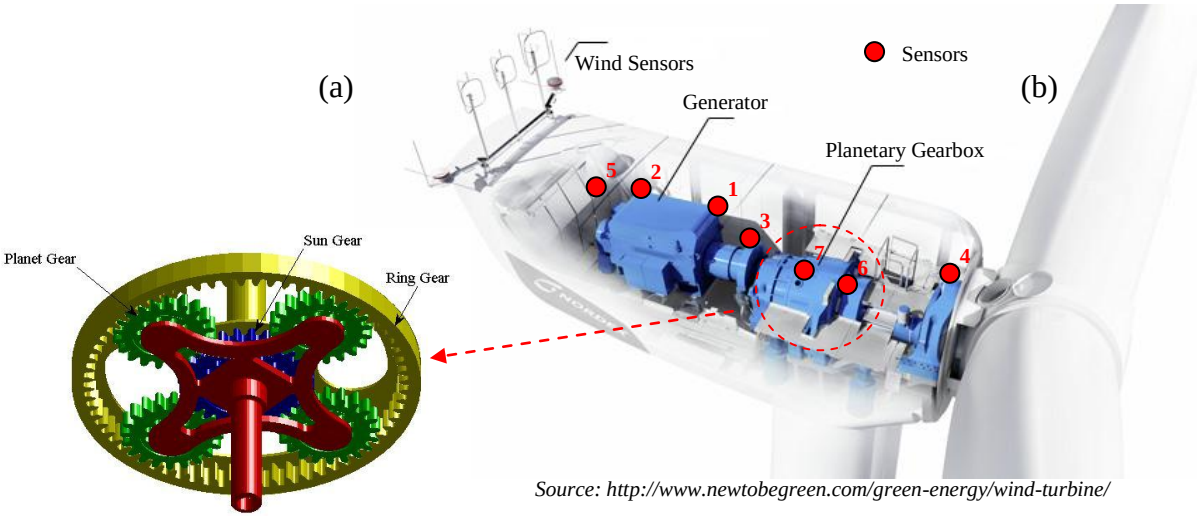
Channel3. High Speed Tachometer

Channel4. Main Bearing Front

Channel5. Main Bearing Rear

Channel6. Low Speed Tachometer

Channel7. Planetary Stage (Gearbox)



Source: http://www.newtobegreen.com/green-energy/wind-turbine/

Figure 1.11: (a) Planetary Gearbox System (b) Wind Turbine

The rest of the thesis is organised as follows:

Chapter 2: Chapter 2 discusses the related work in vibration based fault detection of machines. The use of evolutionary and conventional approaches for noise removal in vibration signals in the research literature along with techniques for decision support and prognostics are also discussed.

Chapter 3: Chapter 3 introduces the concepts behind condition monitoring of wind turbines. Different parts of a wind turbine structure, its systems and sub-systems are introduced. Also, some commercially available condition monitoring systems for wind turbines are listed.

Chapter 4: Chapter 4 applies adaptive band pass and wavelet filtering with conventional optimization techniques for background noise removal from vibration signals.

Chapter 5: Chapter 5 applies adaptive band pass and wavelet filtering with evolutionary optimization techniques for background noise removal from vibration signals.

Chapter 6: Chapter 6 develops novel vibration based features extraction techniques for gears and bearings in time domain. It also introduces some frequency based techniques used in industry and research.

Chapter 7: Chapter 7 proposes prognostics and clustering techniques for fault diagnosis. Two types of prognostics methods and two types of clustering methods are discussed in this chapter.

Chapter 8: Cause and consequence analysis methods are proposed in chapter 8. Three methods including fault semantic networks, fuzzy expert systems, and Bayesian belief network are discussed in this chapter.

Chapter 9: Chapter 9 discusses two real world case studies for wind turbines. Data emanating from wind turbine's supervisory control and data acquisition module are analysed in this chapter.

Chapter 10: Finally, Chapter 10 concludes the thesis along with proposals for future research.

1.11 Conclusion

This chapter highlights the increased concerns about GHG emissions and the importance of green and renewable energy. Wind energy is from one of the renewable energy resources present today but there are technical challenges faced by the wind industry to harness the wind energy efficiently. One of the main challenges is the reliability of the wind turbines and gearbox plays an important role in wind turbine's reliability. This chapter lists some statistical reliability data for wind turbines and identifies gearbox among one of the highest risk elements in a wind turbine. The chapter further analyses the problem deeply and suggests CM-FDS system to tackle it efficiently. Experimental

and real world datasets used in this work are also described. In the following chapter, we will explore some existing methods of fault diagnosis for gearboxes along with their strengths and weaknesses when applied to monitor critical systems in real time environments. We also justify the need for an integrated framework for fault diagnosis of gearboxes.

CHAPTER 2

Literature Review

2.1 Motivation

Gearboxes are complex and inseparable rotating parts of machines today. They are used to step-up or step-down the rotating speeds of the shafts and to transfer the power. Hence, the reliability of the whole system depends on smooth operation of the gearboxes. Development of faults in the gearboxes should be monitored efficiently and at early stages of occurrences in order to avoid costly failures. In this chapter, we explore some existing methods for fault detection in gearboxes. We focus on fault diagnosis methods based on vibration analysis.

2.2 Fault Diagnosis of Gearboxes Based on Vibration Analysis

Vibration signal analysis is a widely used technique to detect faults in rotating machines, especially bearings and gearboxes. Many techniques have been developed in the past to extract machine health related features from vibration signatures. These techniques are performed in conjunction with non-destructive testing (NDT) methods in order to identify faulty components in multi-component machines. NDT is widely used in different disciplines of life. NDT plays an important role in determining the health of the system without physically tearing it apart. In structural mechanics, key properties related to a structure such as displacement and acceleration are measured to determine the health of the structure and NDT is applied to identify faulty components. Because NDT is used to analyze a system's performance without causing any physical damage, it is a highly valuable technique to save both money and time in performing maintenance activities.

NDT works very well in the case of detecting faults in gears and bearings as the kinematics information in gears and bearings are directly related with governing

frequencies in their vibration spectrums. However, a correct identification of the faulty component becomes a difficult task especially for multi-component electromechanical equipment where a majority of components share similar characteristics and interact with each other such as gearboxes. In gearboxes, gears mesh frequencies in the vibration spectrum, governs the faulty information present. The detection task becomes more difficult if one has to identify a faulty gear from a set of gears meshing together.

2.2.1 Fault Detection

Vibration signals emanating from the rotating gearboxes are analysed to ascertain the current condition of the gearboxes. The vibration signals can be classified into stationary and non-stationary signals and based on this classification, the nature of their analysis methods differ. For stationary signals, vibration analysis methods are divided into two domains, namely time and frequency. Time-domain methods include statistical, model based, and signal processing based methods. Frequency-domain methods include spectrum and cepstrum based methods. For non-stationary signals, joint time-frequency vibration analysis methods such as short time Fourier transform (STFT) and wavelet analysis (WA) [18; 19] are used for fault detection in gearboxes. Other methods of non-stationary analysis for gearboxes include Wigner-Ville distribution (WVD) [20], Hilbert-Haung transform [21] and kurtogram analysis [22]. The joint time-frequency analysis methods detect time localized transient features present in the vibration signals. The methods visualize the signals in three dimensions including time, frequency and amplitude variations as in STFT. The WA method provides a powerful multi-resolution analysis and the vibration signals are analysed through wavelet signals of different frequencies and shapes from the wavelet family. A widely applied WA method to detect time localized transient faults in gearboxes is Morlet wavelet analysis [23; 24].

2.2.2 Features Extraction

The features extraction system extracts characteristic signatures from incoming raw vibration signals emanating from gearboxes. The extracted features should be sensitive to the gearbox's health conditions and should reflect the changes in the gearbox's condition or possible fault development over time. As discussed earlier, the vibration based signal

processing analysis is the most common technique because of its advantage of being non-destructive. Features extraction process extracts the information that best represents the faulty conditions present in the signal. Different methods for vibration based features extraction in gearbox fault diagnosis framework have been proposed in research [25-27]. Different techniques such as genetic algorithm, wavelet analysis, and fuzzy inference have been successfully used for features extraction and fault diagnosis of gearboxes [28-31]. In time-domain vibration based features such as kurtosis and spectral kurtosis are extensively used [32-34]. Other studies including statistical-based and transient-based features detection are performed in the past [35-37]. A comprehensive list of time-domain and frequency-domain features for fault detection and diagnosis of gearboxes is discussed in [38].

2.2.3 Evolutionary Algorithms

Evolutionary algorithms have been successfully applied in different fields of scientific computing and became a formal area of study in computer science known as soft computing. The extraordinary complexity of the natural world provides us with remarkably robust and well-designed optimization frameworks [39]. Many organisms in nature have a natural tendency to form swarms, for instance, birds and fish. Swarming behaviour is so prevalent in nature [40]. The global optimization problem has been flourished by nature-inspired techniques, such as ant colony optimization, genetic algorithms (GA), simulated annealing (SA), genetic programming (GP) [41] and others.

There exist different studies of fault detection in dynamic systems using evolutionary algorithms. A fault detection and isolation system is presented in [42]. The system uses wavelet feature-extraction and self-organizing neural networks for fault classification. A vibration based features extraction method is used in [43]. The authors have used genetic programming for features extraction and artificial neural networks with support vector machines for classification of bearing faults. Another acoustics based method is used in [44] where an acoustic signal emanating from a printer is used for features extraction and fault detection. Combinations of wavelet analysis and other artificially intelligent techniques like neural networks, genetic algorithms (GA) and genetic programming (GP) have been successfully used for fault detection and diagnosis of machines and gearboxes.

2.2.3.1 Genetic Algorithms

Genetic algorithms have been extensively used in fault detection in combination with different other algorithms to reduce the solution search space and to speed up the convergence rate. A combination of GA and fast kurtogram is discussed in [45], where rolling element bearing faults are explored. GA is also used for fault diagnosis of gearbox in combination with S-transform and non-negative matrix factorization (NMF) in [46]. A novel enhanced genetic algorithm (EGA) technique is developed in [47] for gearbox's health prognostics. The EGA overcomes problems in classical GA methods and provides a more efficient technique for system training and optimization. A real coded genetic algorithm (RCGA) is used in conjunction with the ant colony optimization technique for real life optimization problems with many local extrema in [48; 49]. Although, a combination of different techniques has been proposed to avoid local extrema, there hardly exists any methodology that combines RCGA with other accelerated search methods to speed up the fault diagnosis process and at the same time preserve the result's quality. Hence, time consumption and slow convergence are still considered major shortcomings of GA.

2.2.3.2 Particle Swarm Optimization

Particle swarm optimization (PSO) is another type of nature-inspired technique that works on population based stochastic optimization principle. PSO was first proposed by Kennedy and Eberhart in 1995 [50]. Natural swarming behaviour of bird flocking and fish schooling is prominent and well-known in PSO implementation. Also, the biological principle of survival of fittest together with evolution of natural behaviour finds its traces in PSO. The algorithm finds an optimal solution in a global solution search space by sharing cognitive and social information among individuals or particles. As compared to other evolutionary population based algorithms such as GA, PSO has an advantage of being simple in implementation with faster convergence and fewer parameters [51; 52]. Many researchers have proposed a modified adaptive PSO with random inertia weights and genetic mutation operators in order to increase the convergence performance and to avoid local extrema [53-56]. Implementation of PSO for fault diagnosis of machines has been investigated by different researchers. In [57], PSO in conjunction with exact WA

and support vector machine (SVM) classifier has been used for fault detection in gearboxes. In [58], PSO is used to calculate an optimal placement of vibration sensor for the fault detection in a gearbox. Another interesting research on the fault diagnosis of the gearbox based on PSO optimization is presented in [59], where PSO is used to train back propagation neural networks (NN). The method increases the convergence speed of the NN and avoids getting stuck in local extrema. Another wavelet based NN algorithm has been proposed with modified PSO for gearbox fault detection in [60].

2.2.4 Fault Diagnosis

Fault diagnosis is a sequential step involving detection, feature-extraction, classification, and decision making as shown in Figure 1.8. Traditionally, human experts are responsible to map extracted, clustered features to respective faults, causes and consequences. This approach is usually tedious, time intensive and unreliable, particularly when the feature space is large or decision has to be made with multiple features under consideration [61].

2.2.4.1 Fuzzy Expert Systems

The fuzzy logic mimics the human reasoning abilities in linguistic form and uses different fuzzy operations to perform the fuzzy diagnosis [62]. An expert system with fuzzy classification has successfully been applied for bearing fault detection in the past [63]. The Fuzzy logic techniques do not have learning capabilities. However, they can be used with evolutionary algorithms to learn the mapping space as described in [64], where an adaptive, neural fuzzy inference system has been used to learn from data. Another fuzzy logic inference system has been used in [65] to detect faults in gearboxes. Here, an energy spectrum of wavelet analysis is used for feature-extraction and the fault conditions have been classified for different rotation frequencies with different levels of Daubechies wavelets. When a condition monitoring system is employed in real-world industrial applications, the critical issue is its reliability. Unreasonably missed alarms (i.e., the monitoring system cannot pick up existing faults) and false alarms (i.e., the monitoring system triggers an alarm because of noise, i.e., not real faults) can seriously mitigate the system’s validity. Hence, a more advanced diagnostic system is in demand for an accurate assessment of gear's health conditions.

2.2.4.2 Bayesian Belief Networks

Bayesian belief networks (BBN) were introduced in 1988. Since then, researchers explored the possibilities to use Bayesian networks for medical decision analysis [66]. BBN is a probabilistic inference network that implies the Bayesian probability theory for the decision making process. BBN is a set of nodes connected through arcs. The arcs are directional, and the network is called directed acyclic graph (DAG). The application of BBN for fault diagnosis in gearboxes is relatively new and there is not much work done in this domain in the past. A BBN is used for fault diagnosis of flexible rotors in [67] where probability tables for each node in the BBN are chosen by an expert. Other approaches for fault diagnosis in gearboxes using BBN is presented in [68; 69], where researchers have used different time-domain statistical features to classify different faults in gearboxes.

2.2.5 Prognostics

Prognostic has a very important role in order for an accurate and reliable decision making. Prognostics can be used effectively in utilization and maintenance of machinery systems. Prognostics can use different machines health related indices including temperature, oil-debris analysis, acoustics, and vibration. Among these, vibration based prognostics is quite common in condition based maintenance (CBM) of machines. In [70], researchers have used different health monitoring indices in gearboxes such as gear-wear, gear-chipped, gear-crack, gear-pitting, and shaft misalignment for projecting the gearbox health information in future. They have used neuro-fuzzy approaches for modeling and prediction of gearbox dynamics. A comprehensive review on prognostics for CBM is presented in [71], where decision making process based on diagnostics and prognostics is discussed. The three main steps involved in CBM are data acquisition, signal processing, and decision making. Data acquisition step includes data pre-processing, analogue to digital (A/D) conversion, and storage of data in computers. Signal processing step involves data processing for extraction of machine health related features. The representative features include time-domain, frequency-domain, and joint time and frequency domain signal processing. Prognostic is a part of decision making

process in the CBM of machines. Prognostic is performed by estimating the temporal evolution of the features over time [72 - 74]. Different methods are proposed for prognostics of machine health related indices in the past [75]. Among those, vibration based prognostics is more common. In vibration based prognostics, vibration signals emanating from sensitive components inside the gearbox are recorded and health related features are extracted. Different techniques are used to extract the health related features and time series prediction techniques are applied to the features trends for prognostics. In [76], researchers have used different time-domain and frequency-domain features for prognostics of gearbox's health. Statistical, evolutionary and soft computing approaches are used to estimate the predictors and the use of neural networks [77] and neuro-fuzy methods are very common [78]. In [79] authors have used general path model (GPM) for estimation of remaining useful life (RUL). The GPM uses statistical based techniques and degradation measures to estimate the time of failure (TOF) for a turbofan machine.

2.3 Conclusion

This chapter explores different vibration based fault diagnosis methods proposed in open research. Among these, adaptive filtering techniques are commonly used to filter vibration signals emanating from rotating machines. Objective functions like Shannon entropy, signal energy, kurtosis maximization, and singular value decomposition have been used in these techniques [80-83]. However, these techniques lack proper optimization framework in selecting suitable values of adaptive filter parameters. Some studies have used evolutionary algorithms for optimization of filter parameters but at the expense of computational time. On the other hand, different time-domain features extraction methods such as kurtosis, root-mean-square, peak-to-peak, and crest factor exhibit poor performance if used on raw vibration signals. Therefore, there is a need for an intelligent and integrated system that can do the followings: (1) harness the power of optimization algorithms and at the same time makes it suitable for real time applications such as wind turbine gearboxes, (2) extract the gearbox health related features for clustering and prognostics, and (3) devise a framework for intelligent decision support. In the next chapter, we will review such systems already implemented in the wind industry

with their analysis capabilities and shortcomings, and we will justify a need to enhance the performance of such systems.

CHAPTER 3

A Review on Condition Monitoring of Wind Turbines in Industry

3.1 Introduction

As discussed in chapter 1, the detection of incipient faults in wind turbines is of utmost importance while the wind turbines are in operation. Wind turbines installed offshore, exhibit more probability of faults development as compared to onshore installations. Also the offshore wind turbines are subjected to severe environmental conditions that restrict maintenance activities. Therefore, it is a technological challenge to monitor wind turbines non-destructively. There are many condition monitoring (CM) systems of wind turbines already developed in the industry [84]. This chapter discusses the overall architecture of wind turbine monitoring systems and lists some commercially available CM systems along with their analysis capabilities, shortcomings, and a need to enhance their performance in the wind industry.

3.2 Monitoring Structures of a Wind Turbine

The CM of wind turbine is carried out because of many reasons: (1) obtaining information commercially important like amount of energy generated and operation status of a turbine, (2) monitoring the health of the wind turbine, (3) detecting faults and failures through alarm systems, and (4) finding the locations of faulty components or causes of faults for scheduling maintenance activities. The monitoring of a wind turbine consists of four major sub-systems as shown in Figure 3.1. The monitoring techniques are different for different sub-systems and mainly depend on the properties and behaviours of the structures of sub-systems and the environmental condition they are subjected to.

3.2.1 Supervisory Control and Data Acquisition (SCADA)

SCADA is required on all modern wind turbine systems these days. SCADA monitors different important information from a wind turbine like, amount of energy generated, operational status of the wind turbine, turbine's controller and related parameters. Although, SCADA monitors at comparatively high frequency, the measurements are averaged over a long period of time (5-10 minutes) and fed back to the monitoring and control panel [85]. SCADA also incorporates algorithms like trend analysis to monitor potential deviations of recorded parameters above the pre-set thresholds. SCADA systems provide alarms for different malfunctions occurred during the operation of a wind turbine.

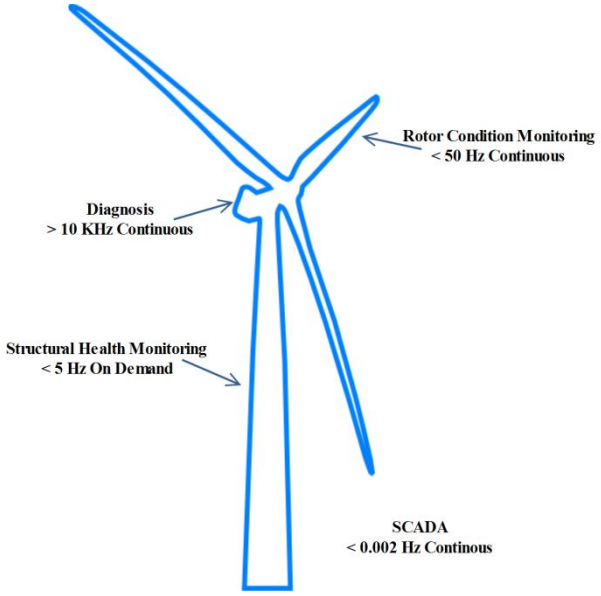


Figure 3.1: Structural Health Monitoring of a Wind Turbine

Most recent SCADA systems include the following different parameters [85]:

(a) Reactive power.
(b) Active power output.
(c) Wind speed.
(d) Gearbox lubrication oil temperature.
(e) Gearbox bearing temperature.
(f) Gearbox winding temperature.
(g) Power factor.

(h) Phase currents.

(j) Nacelle temperature.

(k) Drivetrain vibration readings.

3.2.2 Structural Health Monitoring

Structural health monitoring (SHM) systems determine the integrity of the wind turbine towers, structures, and foundations. In SHM, low frequency monitoring accelerometers (<5Hz) are used to detect faults in wind turbine's structures induced by blade-passing frequencies. Recent trends in monitoring technologies of SHM systems include fibre optics strain calculations and acoustic emission sensing techniques.

3.2.3 Rotor Condition Monitoring

Wind turbine's blades have increased in size more than 80m in recent years and they weigh more than 17 tones. Also, wind turbines' rotors must bear more than 10^8 stress cycles over an average span of 20 years of its life time. This makes the rotor's condition monitoring a crucial task. Different issues in rotor blade's monitoring include [86]:

(a) Rotor balance and aerodynamic symmetry.

(b) Moisture uptake.

(c) Icing.

(d) Surface roughness.

Manufacturing defects or impacts and damages during transportation may cause the rotor's imbalance and aerodynamic asymmetry. Also an accumulation of non-uniform ice and moisture results in an imbalanced rotor. The operation of the turbine must be stopped to mediate the effect of severe imbalance until either the ice melts or maintenance has been performed to rebalance the rotor. The variations in the rotor speed owing to asymmetry can be detected through variations in the power spectral density (PSD) of the electrical output recorded by the turbine's SCADA system. In PSD, a peak at the rotor frequency indicates either rotor imbalance or aerodynamic asymmetry owing to yaw error.

3.2.4 CM-FDS System for a Wind Turbine

CM-FDS system provides an overall health status of a wind turbine's mechanical and electrical drive trains. A reliable indication of incipient faults, determination of faults severity, and locations of faults are the tasks handled by CM-FDS system. Figure 3.2 shows different critical sections monitored by CM-FDS system in wind turbine's nacelle.

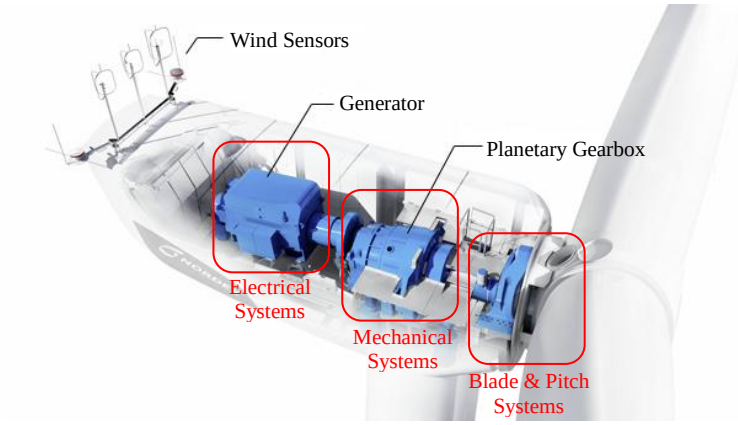


Figure 3.2: Condition Monitoring and Diagnosis in Nacelle
(*Source: http://www.newtobegreen.com/green-energy/wind-turbine/*)

These days, every modern wind turbine is partially or completely equipped with some form of CM-FDS system. The CM-FDS system includes system-level or subsystem-level fault detection and monitoring of a wind turbine. Different parameters like vibration of a wind turbine's drive train [87], bearing temperature, oil particulate content, and optical strain measurements [88] are monitored in CM-FDS. The condition monitoring and fault diagnosis system for system-level and subsystem-level of a wind turbine pose great technological challenges that have led to various modeling and solution approaches in the open research literature [89-91].

As stated above, an indication of a fault location is very important in CM-FDS system. An occurrence of a fault triggers further detailed investigations that include diagnosis and cause and consequence analysis. The identification of a faulty component in CM-FDS system bridges the gap between CM and FDS for wind turbines. Such a system is already proposed in Figure 1.7. A diagnosis process gives initial indications about the location of the faulty system and if possible the sub-system. The monitoring engineer could then start

further detailed investigations in order to pin point the exact location of the faulty component and estimate the physical nature of the fault. Keeping in view the above aims, the CM-FDS system should provide enough information to reliably monitor the wind turbine and its sub-systems and at the same time should not flood the monitoring engineer or data transmission bandwidth with excess and redundant information.

3.3 Commercially Available CM-FDS Systems of Wind Turbines

Some of the commercially available CM-FDS for wind turbines are shown in Figure 3.3 [92]. The CM-FDS systems shown in Figure 3.3 focus on the following major areas of the wind turbine:

(a) Blades.
(b) Main rotor bearing.
(c) Gearbox gears.
(d) Gearbox bearings.
(e) Generator bearings.

Some recent modern CM-FDS systems for wind turbines monitor generator windings, converters and pitch control systems. It is evident that 70% (14 out of 20) of commercially available systems described in Figure 3.3 use vibration analysis for condition monitoring of wind turbines. Three systems are specified for oil debris monitoring, one for vibration based monitoring of turbine blades and two for fibre optic based strain measurements.

It is clear from Figure 3.3 that industry is extensively using vibration based analysis methods for condition monitoring of wind turbines. All 14 vibration based condition monitoring systems also imply some form of diagnosis capability after a fault is detected. All vibration based monitoring systems use fast Fourier transform (FFT) for diagnosis. The FFT is a well-established technique in the industry and it is a very useful tool to detect fault related frequencies and harmonic patterns present in a vibration signal.

3.4 Justification for Condition Monitoring of Wind Turbines

Although, condition monitoring systems used in the industry today are quite matured and based over decades of research, they suffer from the following shortcomings:

(1) **Analysis of Non-stationary Signals.** The vibration signals emanating from wind turbines mechanical drive trains or for that matter gearboxes are non-stationary in nature. Hence, there is a need for methods to perform signal processing of non-stationary signals such as WA or STFT. The conventional FFT method used in existing CM-FDS systems loses the time information and is not suitable for analysing non-stationary signals.

(2) **Background Noise.** The problem of huge background noise in vibration signals restricts the performance of analysis and features extraction methods. An intelligent noise removal technique is needed to enhance the signal SNR. The existing CM-FDS systems incorporate many analysis and features extraction methods that use raw vibration signals. Therefore, they suffer from low detection quality and sometimes false alarms.

(3) **Effective Decision Support.** The existing CM-FDS systems offer some kind of fault diagnosis based on FFT analysis but lack with advanced and intelligent decision techniques.

Increasing the reliability of fault detection methods and automating the fault diagnosis process is a way forward for future research in wind turbine's condition monitoring systems. Also applying intelligent signal processing to infer meaningful information from noisy data and methods used for data dimensionality reduction pave their way in modern condition monitoring systems for wind turbines. A lot of research has already been done in academia in this context as discussed in Chapter 2 but lacks in devising an appropriate optimization framework for background noise removal and fault detection that performs in comparatively less time and preserves the quality of results.

3.5 Conclusion

New technologies are encroaching on traditional methods for condition monitoring systems of wind turbines and are developing over the time. The future direction for

research on condition monitoring systems should be to develop and refine the existing techniques to suit the new working environments. The new technologies should be incorporated with existing condition monitoring systems in order to bridge the technological gaps and enhance their capabilities. Efforts should be made to monitor the high risk components in a wind turbine such as gearboxes and electrical system. This thesis work develops an integrated and intelligent framework for condition monitoring and fault diagnosis in wind turbine gearboxes. In the following chapters, we will develop conventional and evolutionary methods to intelligently detect faults in vibration signals emanating from wind turbine gearboxes. We will present transient based features extraction techniques and apply soft computing theory to predict the fault behaviour over time. We will also present a framework for intelligent decision support system in fault diagnosis of wind turbine gearboxes.

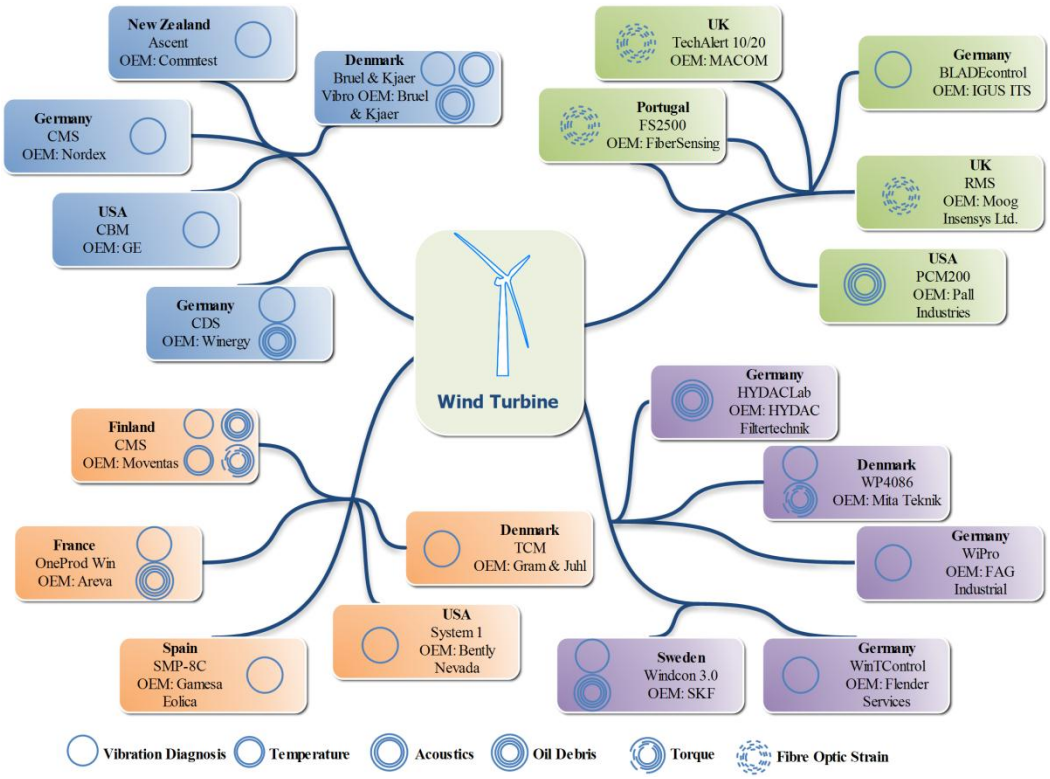


Figure 3.3: Commercially Available CM-FDS Systems [92]

CHAPTER 4

Conventional Optimization in Background Noise Removal

4.1 Motivation

One of the major problems in vibration based feature extraction is the presence of huge background noise. This makes fault detection a difficult process especially when the SNR of the faulty signal is very low and the fault is at an early stage of its development. Intelligent filtering techniques are required to remove the background noise and detect the valuable faulty information early in time. The early fault detection enables timely repair of the system to avoid further development of the faults into severe conditions and eventually catastrophic failures. This chapter develops conventional optimization methods to filter the huge background noise present in vibration signals emanating from wind turbine gearboxes. The methods are implemented in front end fault detection framework as shown in Figure 1.8.

4.2 Adaptive Filtering

Adaptive filtering is a process of iteratively calculating filter parameters to satisfy some objective functions. Two types of filtering methods are used in this chapter, band-pass filtering and wavelet analysis. In this chapter, we use conventional optimization framework to adaptively tune the filters and optimize an objective function. In conventional optimization techniques, one-dimensional golden section search and multi-dimensional Nelder-Mead methods are used. A flow chart of the conventional optimization framework is shown in Figure 4.1, where coarse tuning and fine tuning steps are included for optimization of filter parameters. In coarse tuning step, the filter parameters are optimized through one-dimensional search techniques and an objective function is maximized. The optimized values of these parameters are then passed to the fine tuning step. In the fine tuning step, a multi-dimensional optimization method is used

to fine tune all the filter parameters simultaneously. The main purpose behind using coarse tuning step is to benefit from the speed of one-dimensional search techniques and to give an advantageous start to multi-dimensional search methods. The process also helps in avoiding the multi-dimensional search methods to get stuck at local extremum.

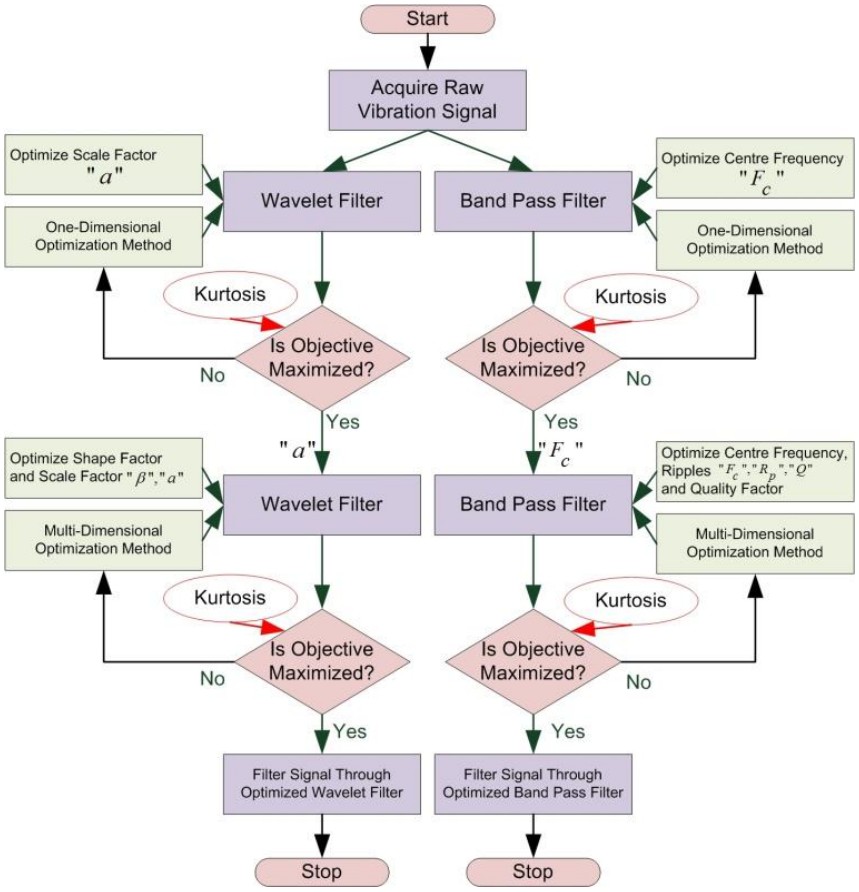


Figure 4.1: The Adaptive Filter Framework

In this chapter, a combination of band-pass filter with kurtogram is also explored. Kurtogram is a fourth order spectral analysis tool and is suitable to detect non-stationarities in a signal. Since vibration signals emanating from inside a wind turbine contain non-stationarity, kurtogram is a very useful tool to detect this phenomenon because of the sensitivity of kurtosis towards the faulty pulses present in the signal. The complexity of computing kurtogram is $n \log n$ similar to the FFT and is suitable for implementation in real time fault detection applications. A detailed description about the

kurtogram is presented in [22]. Kurtogram gives an estimation of the faulty bandwidth present in the vibration signal and gives an advantageous start for band-pass filtering.

4.2.1 The Objective Function

The objective function or fitness function used for this work is maximization of kurtosis. Kurtosis is the degree of peakedness and is defined as a normalized form of the fourth central moment μ_4 of a distribution. The kurtosis based objective function is calculated as [38]

$$\begin{aligned} & f(x) = \max Kurt = \frac{\frac{1}{n}\sum_{i=0}^{n-1}\left(x_t(i)-\mu\right)^4}{\sigma^4}. \\ & s.t. \\ & c1_{\min} \le c1 \le c1_{\max} \\ & c2_{\min} \le c2 \le c2_{\max} \end{aligned} \tag{4.1}$$

Here, n is total number of samples in the signal, $x_t(i)$ is the i^{th} sample, μ is the signal mean, σ is the standard deviation, $c1_{\min} \le c1 \le c1_{\max}$ and $c2_{\min} \le c2 \le c2_{\max}$ are the constraints on the search space. We use a penalty method and drop constraints of non-linear objective function by substituting new terms in the objective function penalizing infeasibility in the form [93]

$$\max or \min F(x) = f(x) \pm \eta \sum_i p_i(x). \tag{4.2}$$

Where, "+" is used for minimization problems and "–" is used for maximization, η is a positive penalty multiplier and p_i are functions satisfying

$$p_i(x)\begin{cases} = 0 & \text{if } x \text{ satisfies constraint } i \\ > 0 & \text{otherwise} \end{cases}. \tag{4.3}$$

The new unconstrained objective function becomes

$$\begin{aligned} \max Kurt = \frac{\frac{1}{n}\sum_{i=0}^{n-1}\left(x_t(i)-\mu\right)^4}{\sigma^4} - \eta\Big[& \max^2\{0, c1_{\min} - c1\} + \max^2\{0, c1 - c1_{\max}\} + \dots \\ & \max^2\{0, c2_{\min} - c2\} + \max^2\{0, c2 - c2_{\max}\}\Big]. \end{aligned} \tag{4.4}$$

As an example, if the constraints $c1_{\min} \le c1 \le c1_{\max}$ and $c2_{\min} \le c2 \le c2_{\max}$ are satisfied, the η part in Equation (4.4) becomes zero, and if the constraints are not satisfied, a squared penalty is subtracted from the objective function that restricts the objective function to be maximized. The inclusion of penalty functions in the objective function makes the algorithms converge faster. In our case, we start η with a very small value less than 1 and sequentially increase it with some multiplier δ so that $\eta_{t+1} = \delta\eta_1$.

4.2.2 Chebyshev Band-Pass Filter

Chebyshev band-pass filter is commonly used in many filtering application for fault detection. Figure 4.2 shows different design parameters for a typical Chebyshev band-pass filter. The Chebyshev band pass filter has an advantage over other filters because of its speed since it is carried out by recursion rather than convolution. Although, it cannot be compared with a windowed sinc filter in performance but it presents an advantage in terms of speed and is widely acceptable in many real time applications [94]. We make use of the Chebyshev band-pass filter as a trade-off between speed and performance. The Chebyshev filter applies a mathematical strategy to achieve a faster roll off by allowing ripples in frequency response. We design type1 Chebyshev filter as it allows ripples only in the pass band.

4.2.2.1 Conventional Optimization of Chebyshev Band-Pass Filter

We need to control three parameters: centre frequency F_c, quality factor Q and pass band ripples R_p. Filter order N is kept constant at value equal to 4. Initially, the kurtogram is used to detect the faulty bandwidth present in the vibration signal. The filter's centre frequency F_c is estimated by the kurtogram, where the search process is started in the range $F_c \pm \frac{F_c}{2}$. Setting the quality factor $Q = 0.707$, $R_p = 1$ and the order $N = 4$ then applying one dimensional accelerated search to optimize F_c, Q and R_p parameters of the filter one by one in the coarse tuning step. The optimized results from the coarse tuning are then passed to the Nelder-Mead algorithm for multi-parameters

optimization. Here, F_c, Q and R_p are optimized for fine tuning. Since, the transition band is related to the pass band ripple R_p in the Chebyshev filter design, we ignore the parameters Fs_1 and Fs_2 in Figure 4.2. We also ignore the stop band ripples Rs as type1 Chebyshev filters do not exhibit stop band ripples. The values for Fp_1, Fp_2 are calculated from the following equations [95]:

$$Fp_1 = F_c\left(\sqrt{1+\frac{1}{4Q^2}}-\frac{1}{2Q}\right)$$
$$Fp_2 = F_c\left(\sqrt{1+\frac{1}{4Q^2}}+\frac{1}{2Q}\right). \quad (4.5)$$

After data acquisition and conditioning, the kurtogram is computed for the raw vibration signal. As described earlier, the kurtogram gives an initial estimate about the centre frequency of the filter that maximizes the kurtosis.

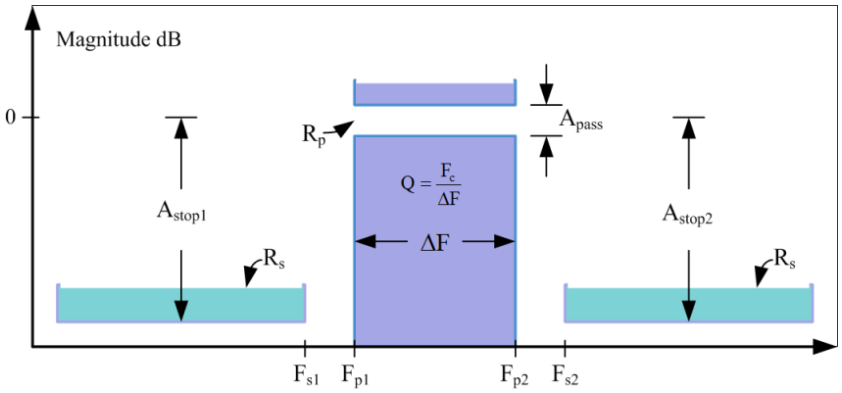


Figure 4.2: Band-Pass Filter Specifications

A Chebyshev bandwidth filter is initialized, and the filter parameters are set according to the initial estimates from the kurtogram. A direct search method with a derivative free search is used to tune the band-pass filter and the kurtosis is used as an optimization cost function as shown in Equation (4.4). A combination of genetic algorithm (GA) with the kurtogram for detection of rolling element bearings faults is proposed in [45]. Although, GA is given with initial parameters in order to search for the extremum, the process is still inherently slow if considered for real time implementation. In order to speed up the performance, a coarse tuning step is included, that performs a speedy one dimensional

search on Chebyshev band-pass filter one by one. The coarse tuning step makes use of the golden section search combined with parabolic interpolation for fast and guaranteed convergence [96]. Derivative free direct search method like Nelder-Mead is a good candidate for fast convergence and it outperforms evolutionary search algorithms in terms of computational complexity. The Nelder-Mead method fine tunes of the band-pass filter. It provides enough information and reasonable extraction of faulty features in a reasonable amount of time as compared to evolutionary algorithms. In each step of the optimization process, techniques with comparatively less computational complexity are implemented for real time realization of faults.

4.2.2.2 Accelerated One Dimensional Search

An accelerated one dimensional search is performed in the coarse tuning step as shown in Figure 4.1. As discussed earlier, we use golden section one dimensional search technique. The golden section search deals with a unimodal objective by rapidly narrowing an interval guaranteed to contain optimum [61]. Figure 4.3 illustrates the idea of the golden section search, where four carefully spaced points are iteratively considered. Leftmost $x^{(lo)}$ is a lower bound on the optimal x^* and $x^{(hi)}$ is an upper bound. The function optimum lies between the interval, $\left[x^{(lo)}, x^{(hi)}\right]$. Points $x^{(1)}$ and $x^{(2)}$ are the intermediate points. Each iteration determines whether the objective is better at $x^{(1)}$ or $x^{(2)}$, if $x^{(1)}$ proves better, the direction for the next iteration is left and $x^{(2)}$ becomes $x^{(hi)}$, and if $x^{(2)}$ proves better, the direction for the next iteration is right and $x^{(1)}$ becomes $x^{(lo)}$. The choice of interior points define the efficiency of the search, whether it is $\left[x^{(1)}, x^{(hi)}\right]$ or $\left[x^{(lo)}, x^{(2)}\right]$ interval and golden section search proceeds by keeping both these intervals equal in length. The two middle points of the golden section search are spaced according to the following equation:

$$\begin{aligned} x^{(1)} &= x^{(hi)} - \alpha(x^{(hi)} - x^{(lo)}) \\ x^{(2)} &= x^{(lo)} + \alpha(x^{(hi)} - x^{(lo)}) \end{aligned} \tag{4.6}$$

Where $\alpha = 0.618$ is the golden ratio and $\alpha = (-1+\sqrt{5})/2$. Algorithm 1 [93] gives main steps in performing golden section optimization search. Although, the golden section search algorithm is reliable, it is slow and narrowing of the optimum containing interval requires considerable computation before an optimum is defined with considerable efficiency. We combine quadratic fit search with golden section search for rapid convergence, taking full advantage of the three point pattern fit. We can fit a quadratic function through three points and have a unique minimum or maximum, whichever we are seeking for the given objective function $F(x)$. The unique optimum of the quadratic function agreeing with $F(x)$ at three point pattern $(x^{(lo)}, x^{(mid)}, x^{(hi)})$ occurs at [93]

$$x^{(qu)} = \frac{1}{2} \frac{F^{(lo)}\left[s^{(mid)} - s^{(hi)}\right] + F^{(mid)}\left[s^{(hi)} - s^{(lo)}\right] + F^{(hi)}\left[s^{(lo)} - s^{(mid)}\right]}{F^{(lo)}\left[x^{(mid)} - x^{(hi)}\right] + F^{(mid)}\left[x^{(hi)} - x^{(lo)}\right] + F^{(hi)}\left[x^{(lo)} - x^{(mid)}\right]}. \quad (4.7)$$

Where $F^{(lo)} = F(x^{(lo)})$, $F^{(mid)} = F(x^{(mid)})$, $F^{(hi)} = F(x^{(hi)})$, $s^{(lo)} = (x^{(lo)})^2$, $s^{(mid)} = (x^{(mid)})^2$ and $s^{(hi)} = (x^{(hi)})^2$. The algorithm starts with the golden section search and calculates four points $(x^{(lo)}, x^{(1)}, x^{(2)}, x^{(hi)})$. It then determines the search direction (right or left) and fits a quadratic function with either $(x^{(lo)}, x^{(1)}, x^{(2)})$ or $(x^{(1)}, x^{(2)}, x^{(hi)})$. It calculates the quadratic fit $x^{(qu)}$ from Equation (4.7) and again applies criteria similar to the golden section search to discard one point and so on. Whether to use the golden section search or parabolic interpolation in the next iteration of the algorithm depends on the value of the cost function. If the golden section search yields better value of the objective function than parabolic interpolation, the golden section search will be used for the next iteration, otherwise parabolic interpolation. The combination of parabolic interpolation and the golden section search can speed up the optimal search process by 35-40% as compared to the golden section search only [97]. Figure 4.4 shows the process of parabolic interpolation in combination with the golden section search.

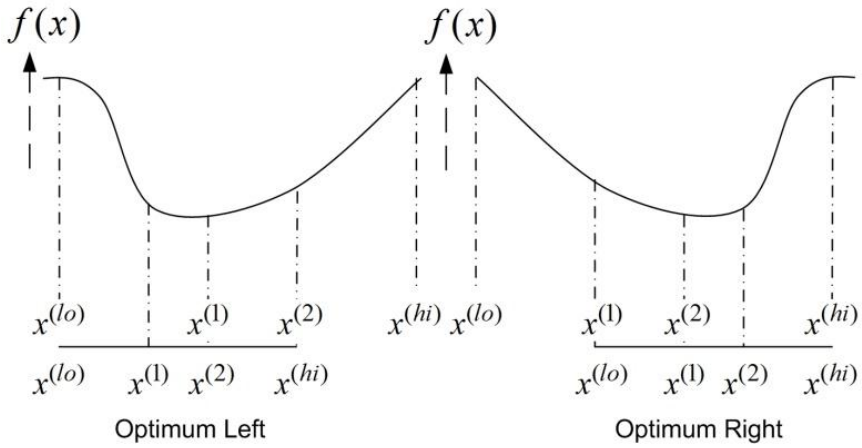


Figure 4.3: Golden Section Search

Step1 : Initialization: Lower Bound $\Rightarrow x^{(lo)}$ *on* $f(x)$
Upper Bound $\Rightarrow x^{(hi)}$ *on* $f(x)$
Stopping Tolerance $\Rightarrow \varepsilon > 0$

Step2 : Computation: Evaluate $f(x)$ on all four points $x^{(lo)}, x^{(hi)}, x^{(1)}, x^{(2)}$

$$x^{(1)} \leftarrow x^{(hi)} - \alpha(x^{(hi)} - x^{(lo)})$$
$$x^{(2)} \leftarrow x^{(lo)} + \alpha(x^{(hi)} - x^{(lo)})$$
$$\alpha \leftarrow 0.618$$
$$t \leftarrow 0$$

Step3 : Stopping: if $(x^{(hi)} - x^{(lo)}) \leq \varepsilon$ Then **Stop**
and

$$x^{*} \leftarrow \frac{1}{2}(x^{(lo)} + x^{(hi)}) \Rightarrow \text{Approx. Optimal Solution}$$

Step4 : Direction: if $f(x^{(1)}) < f(x^{(2)})$ Then **Left Search**

$$x^{(hi)} \leftarrow x^{(2)}$$
$$x^{(2)} \leftarrow x^{(1)}$$
$$x^{(1)} \leftarrow x^{(hi)} - \alpha(x^{(hi)} - x^{(lo)})$$
$$t \leftarrow t + 1$$

Go to Step3

Step5 : Direction: if $f(x^{(1)}) > f(x^{(2)})$ Then **Right Search**

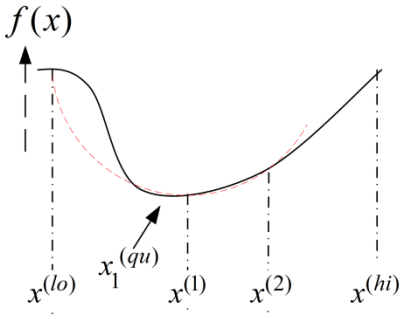
$$x^{(lo)} \leftarrow x^{(1)}$$
$$x^{(1)} \leftarrow x^{(2)}$$
$$x^{(2)} \leftarrow x^{(lo)} + \alpha(x^{(hi)} - x^{(lo)})$$
$$t \leftarrow t + 1$$

Go to Step3

Algorithm 1: Golden Section Search [93]

4.2.2.3 Nelder-Mead Derivative Free Search

The parameters found by the golden section search are passed to the direct search method for fine tuning. The Nelder-Mead is one of the most popular schemes for unconstrained search without calculating derivatives [93]. The accuracy and computation time of any search method greatly depends on the initial values. Since, we use the golden section search and parabolic interpolation to speed up the process, the Nelder-Mead method converges in less time and becomes computationally suitable to implement in real time applications. The Nelder-Mead strategy, in contrast to other greedy search techniques, maintains a set of $n+1$ distinct solutions for n decision variables. The distinct solutions $y^{(1)}, y^{(2)}, ...y^{(n+1)}$ are sorted in descending order having $y^{(1)}$ as the best objective function value and $y^{(2)}$ as the second best and so on. The Nedler-Mead method replaces the worst solution $y^{(n+1)}$ with a better one in each iteration. Algorithm 2 [93] gives main steps in performing the Nedler-Mead search for multi-decision variables. In the case of any derivative free optimization method, the decision for the search direction is critical, and it directly relates to the algorithm's convergence properties. The Nedler-Mead method chooses a search direction away from the worst solution $y^{(n+1)}$ in each iteration.



Since $f(x^{(1)}) < f(x^{(2)})$ parabolic fit occurs at $x^{(lo)}, x^{(1)}, x^{(2)}$ and $x_1^{(qu)}$ is calculated.
Since $f(x^{(lo)}) > f(x_1^{(qu)}) < f(x^{(1)})$
Parabolic fit occurs at $x^{(lo)}, x_1^{(qu)}, x^{(1)}$
and $x_2^{(qu)}$ is calculated and so on till stopping tolerance is reached.

Figure 4.4: Parabolic Interpolation and Golden Section Search

Step 0: Initialization: Choose $(n+1)$ distinct solutions as starting set, $\left[y^{(1)}, y^{(2)}, \ldots y^{(n+1)}\right]$ evaluate $f(y^{(1)}), f(y^{(2)}), \ldots f(y^{(n+1)})$ and initialize iteration index $t \leftarrow 0$.

Step1: Centroid: Rearrange the solutions $y^{(i)}$ in non-improving sequence and compute best n centroid $x^{(t)} = \frac{1}{n}\sum_{i=1}^{n} y^{(i)}$.

Step2: Stopping: If all solution values $f(y^{(1)}), f(y^{(2)}), \ldots f(y^{(n+1)})$ are sufficiently close to centroid value $f(x^{(t)})$ stop and report the best of $y^{(1)}$ and $x^{(t)}$.

Step3: Direction: $\Delta x^{(t+1)} = x^{(t)} - y^{(n+1)} \Rightarrow$ Direction away from worst.

Step4: Reflection: Try $\lambda = 1$ by computing $f(x^{(t)} + 1\Delta x^{(t+1)})$ if the new value is as good as current best $y^{(1)}$, go to Step 5 and expand. If it is no better than second worst value, go to Step 6 and contract. Otherwise accept $\lambda \leftarrow 1$ and proceed to Step 8.

Step5: Expansion: Try $\lambda = 2$ by computing $f(x^{(t)} + 2\Delta x^{(t+1)})$. If this value is no worse than $f(x^{(t)} + 1\Delta x^{(t+1)})$ fix $\lambda \leftarrow 2$ otherwise set $\lambda \leftarrow 1$ and proceed to Step 8.

Step6: Contraction: If reflection value $f(x^{(t)} + 1\Delta x^{(t+1)})$ is better than worst Current $f(y^{(n+1)})$, try $\lambda = \frac{1}{2}$ by computing $f(x^{(t)} + \frac{1}{2}\Delta x^{(t+1)})$. If not try $\lambda = -\frac{1}{2}$ by Computing $f(x^{(t)} - \frac{1}{2}\Delta x^{(t+1)})$. If the result improves from worst current $f(y^{(n+1)})$ fix $\lambda \leftarrow \pm\frac{1}{2}$ tried and go to Step 8. Otherwise go to Step 7 and shrink.

Step7: Shrink: Shrink the current solution set towards the best $y^{(1)}$ by $y^{(i)} = \frac{1}{2}(y^{(1)} + y^{(i)}) \ \forall \ i = 2, \ldots n+1$. Then compute new $f(y^{(2)}), f(y^{(3)}), \ldots f(y^{(n+1)})$, advance $t \leftarrow t+1$ and return to Step 1.

Step8: Replacement: Replace worst $y^{(n+1)}$ in the solution set by $x^{(t)} + \lambda\Delta x^{(t+1)}$ then advance $t \leftarrow t+1$ and return to Step 1.

Algorithm 2: Nelder-Mead Multi-Parameter Optimization Method [93]

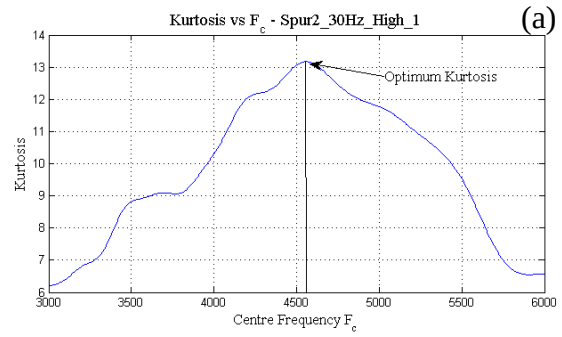
4.2.2.4 Simulations and Discussions

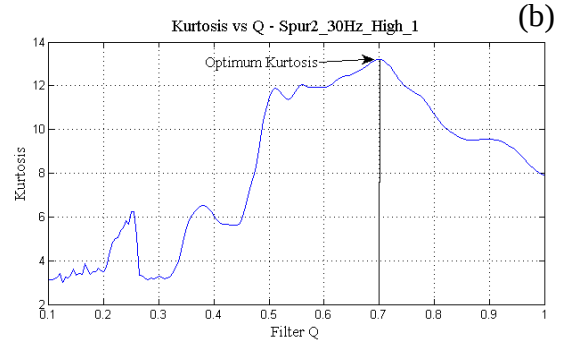
The vibration data used in this section are taken from "Spur2" case listed in Table 1.2 in Section 1.10.2. In this case, one tooth on the 32 teeth gear is chipped. Vibration signals coming from input accelerometer sensor are used in this section. Simulation results in Table 4.1 and Figures 4.5 and 4.6 show the results. A combination of different algorithms is used in this section in order to get comparisons. It is described in section 4.2.2.1 that, at the initial stage, we change F_c, Q and R_p parameters for the band-pass filter and apply

accelerated one dimensional search for rapid convergence and coarse tuning. Kurtogram gives an initial estimate for the centre frequency F_c. Initially, we set $Q = 0.707$ and $R_p = 1\text{dB}$. Order N is always set to 4. We run golden section search combined with parabolic interpolation to search for a value F_c of the band-pass filter which maximizes kurtosis of the filtered signal. Let us take the case in Table 4.1, the value of the centre frequency returned by the kurtogram for a particular signal is $F_c = 5403.59\text{Hz}$. Figure 4.5(a) shows the change in the kurtosis value plotted against the centre frequency of the band-pass filter F_c, in the range $\left[F_c - \frac{F_c}{2}, F_c + \frac{F_c}{2}\right]$. Running an accelerated one dimensional search on F_c, the optimum is found just in nine iterations of the algorithm because of the reduction in search space by the kurtogram and the accelerated nature of the golden section search and parabolic interpolation combination strategy. The termination criteria we use, is to stop the execution when further improvement is less than 0.001. Figure 4.5(a) plots kurtosis behaviour against the centre frequency F_c. Table 4-5-1 in Figure 4.5 shows the results of this run. The initial run calculates the kurtosis value equal to 13.1735 for centre frequency $F_c = 4557.83\text{Hz}$. Freezing the value of the centre frequency at $F_c = 4557.83\text{Hz}$ and applying one dimensional search to the filter quality parameter Q in the range $[0.1 - 1]$ and looking for any improvement in the kurtosis value. Figure 4.5(b) plots kurtosis behaviour against filter quality parameter Q. Table 4-5-2 in Figure 4.5 shows the results of this run. The algorithm converges in nine iterations, and there is an improvement in kurtosis value from 13.17 to 13.21 for $Q = 0.70017$. Further, freezing the values of $F_c = 4557.83\text{Hz}$ and $Q = 0.70017$ and applying one dimensional search to the filter pass band ripples R_p in the range $[1 - 10]\text{dB}$. Figure 4.5(c) plots kurtosis behaviour against filter pass band ripples R_p. Table 4-5-3 in Figure 4.5 shows the results of this run. The algorithm converges in six iterations improving kurtosis value from 13.217 to 13.219 for $R_p = 1.103\text{dB}$. So, there is no considerable improvement in kurtosis value for R_p search.

We have optimized three parameters $F_c = 4557.83\text{Hz}$, $Q = 0.70017$ and $R_p = 1.103\text{dB}$ individually. The final kurtosis value we have obtained is 13.219. We use the Nelder-Mead method for further fine tuning of the filter parameters. The Nelder-Mead method takes all the three parameters and applies multi-objective optimization as discussed in section 4.2.2.3. The Nelder-Mead method converges in 11 iterations with 22 function evaluations and produces the final values of $F_c = 3839.81\text{Hz}$, $Q = 0.600$ and $R_p = 4.1\text{dB}$ as shown in Table 4-5-4 in Figure 4.5. The Nelder-Mead method can often handle discontinuity, particularly if it does not occur near the solution. To avoid the Nelder-Mead method to be stuck at local extremum, we use one dimensional accelerated search. The strategy gives good starting point for the Nelder-Mead method and ensures convergence at global extremum. Experiments show that if we use the Nelder-Mead method just after the kurtogram estimations and use $F_c = 5403.59\text{Hz}$, $Q = 0.707$ and $R_p = 1\text{dB}$ as a starting point, the method sometimes converges at local extremum. The kurtogram combined with GA as proposed in [45] gives a reasonable detection of faulty pulses but is computationally expensive and not suitable for real time implementation. Instead, after using the fast Chebyshev filter design, the computational complexity of GA is still not reasonable for real time implementation. On the other hand, the combination of kurtogram with golden section search and Nelder-Mead method gives comparable results in a reasonable amount of time for online real time implementation. In Figures 4.6(a) and 4.6(b), a sample of a raw vibration signal from the case “Spur2” is shown with respective kurtogram. After initial parameter estimations through kurtogram, a band-pass filter is designed, and filtered signal is obtained. The faulty pulses in the signal are prominent as shown in Figure 4.6(d) but kurtosis value is not maximized (kurtosis = 10.20) because the band-pass filter is not yet tuned as shown in Figure 4.6(c). The tuned band-pass filter is then obtained through kurtogram-golden-Nelder-Mead combination as shown in Figure 4.6(c), and the raw signal is filtered again. Faulty pulses are now more prominent as SNR is improved with kurtosis value maximized to 13.82 as shown in Figure 4.6(e). The above mentioned simulations are repeated for another case “Helical2” listed in Table 1.2 of

Section 1.10.2. In this case, 24T gear is chipped. Figure 4.7 and Table 4.2 show the optimization results. The kurtosis is maximized from 19.37 to 24.45.





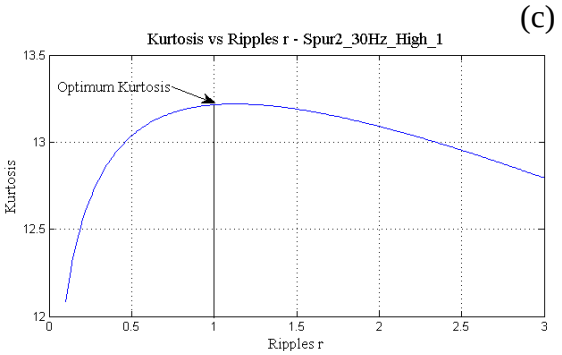


Table 4-5-1: One-Dimensional Search F_c

Count	F_c (Hz)	Kurtosis	Method
1	4145.9	11.6888	Initial
2	4854.1	12.0664	Golden
3	5291.8	10.7178	Golden
4	4584.52	13.1509	Parabolic
5	4525.67	13.139	Parabolic
6	4562.96	13.1726	Parabolic
7	4558.21	13.1735	Parabolic
8	4557.73	13.1735	Parabolic
9	4557.83	13.1735	Parabolic

Table 4-5-2: One-Dimensional Search Q

Count	Q	Kurtosis	Method
1	0.44376	5.65486	Initial
2	0.65623	12.5827	Golden
3	0.78753	11.1468	Golden
4	0.67871	12.9517	Parabolic
5	0.70012	13.2146	Parabolic
6	0.73351	12.3684	Golden
7	0.71287	13.0666	Golden
8	0.69820	13.2111	Parabolic
9	0.70017	13.2147	Parabolic

Table 4-5-3: One-Dimensional Search R_P

Count	R_P (dB)	Kurtosis	Method
1	1.2077	13.2175	Initial
2	1.8923	13.1153	Golden
3	0.7845	13.1791	Golden
4	1.2054	13.2176	Parabolic
5	1.1421	13.2191	Parabolic
6	1.1032	13.2192	Parabolic

Table 4-5-4: Multi-Dimensional Search (Nelder-Mead)

Count	Function Eval.	Kurtosis	Procedure
1	4	13.2192	Initial simplex
2	6	13.2252	Contract inside
3	8	13.2297	Contract inside
4	10	13.2859	Contract inside
5	12	13.3150	Contract inside
6	13	13.5710	Reflect
7	15	13.5734	Contract inside
8	17	13.6158	Contract inside
9	19	13.6192	Contract inside
10	21	13.8201	Expand
11	22	13.8211	Reflect

$F_c = 3839.81\text{Hz}$, $Q = 0.600$ and $R_p = 4.12\text{dB}$

Figure 4.5: Accelerated One-Dimensional and Multi-Dimensional Search

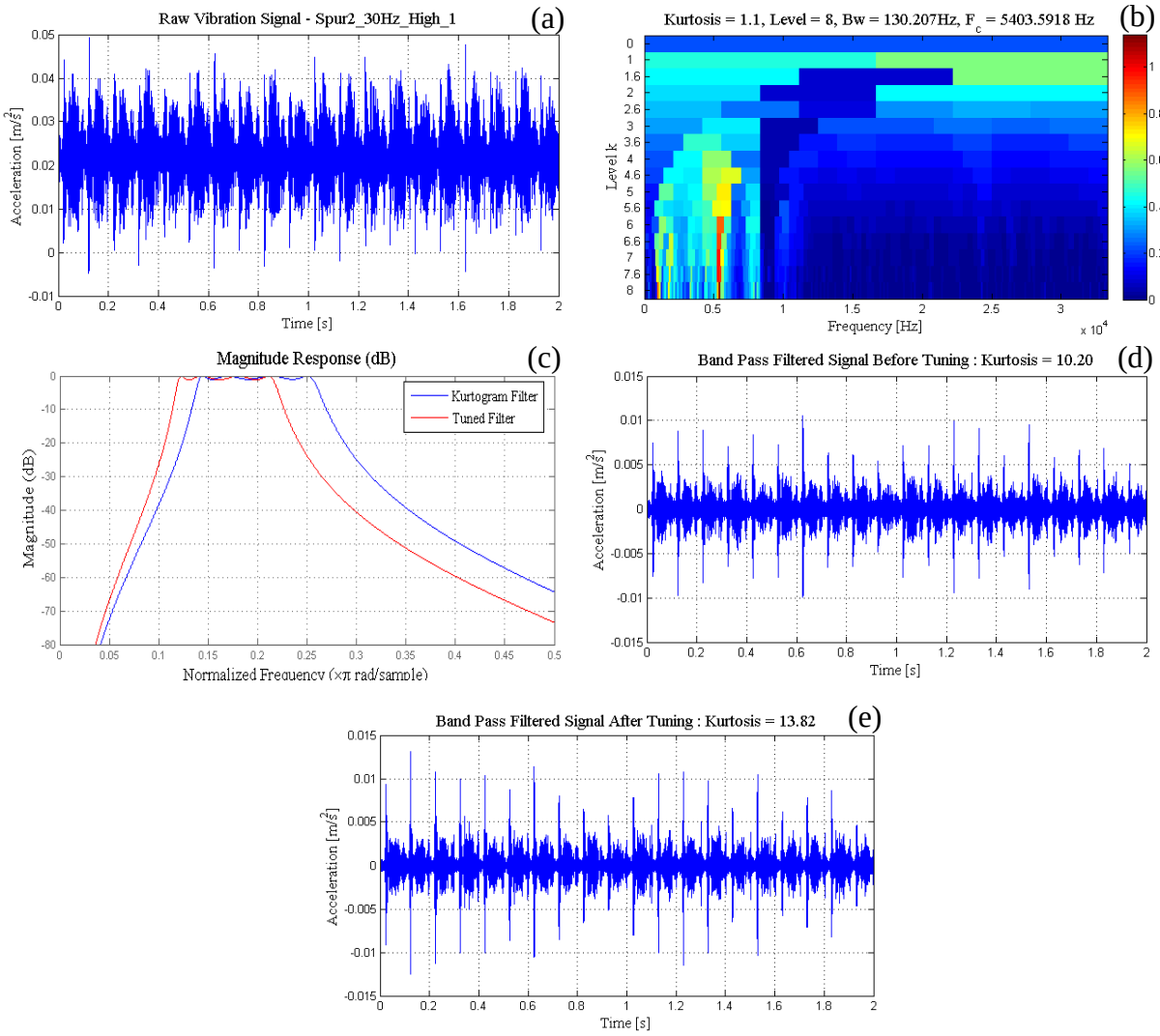


Figure 4.6: Band-Pass Filter Tuning (Case - Table 4.1)

Table 4.1: Filter Tuning Algorithms – Different Combinations: 1800 RPM, Load: High (Accelerometer Location: Gearbox Input, Type of Gears: Spur, Case: Spur2) (Platform: Intel Core Duo 2.27 GHz, 4GB RAM)

Params	Kurtogram	Filter Initialization *(Chebyshev)*	Filter Tuning Algorithmic Combinations			
			Kurtogram GA/ Kurtogram Nelder-Mead	Kurtogram GA Nelder-Mead	Kurtogram Golden GA	Kurtogram Golden Nelder-Mead
F_c	5403.59	5403.59	4535.18/4025.72	4521.51	4546.18	3839.83
Q	-	0.7071	0.707/0.562	0.6831	0.7521	0.600
N	-	4	4/4	4	4	4
R_p	-	1 (dB)	1.217/1.434 (dB)	1.452 (dB)	1.016 (dB)	4.12 (dB)
Time(s)	1.976	0.069	84.344/13.3259	92.359	63.156	10.178
Kurt	1.1	10.20	13.20/13.05	13.011	11.913	13.821

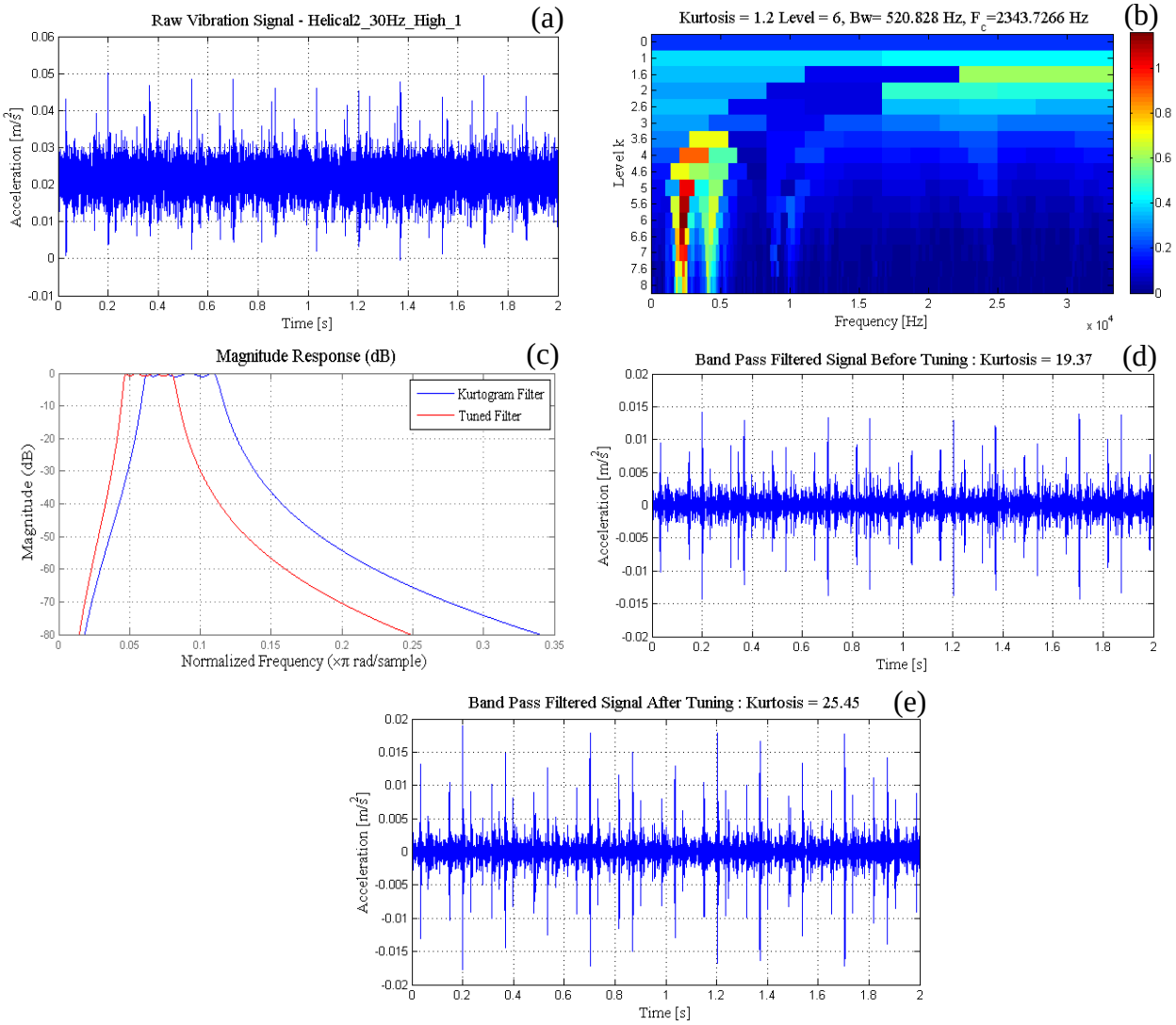


Figure 4.7: Band-Pass Filter Tuning (Case - Table 4.2)

Table 4.2: Filter Tuning Algorithms – Different Combinations: 1800 RPM, Load: High (Accelerometer Location: Gearbox Input, Type of Gears: Helical, Case: Helical2) (Platform: Intel Core Duo 2.27 GHz, 4GB RAM)

Params	**Kurtogram**	**Filter Initialization *(Chebyshev)***	**Filter Tuning Algorithmic Combinations**			
			Kurtogram GA/ Kurtogram Nelder-Mead	**Kurtogram GA Nelder-Mead**	**Kurtogram Golden GA**	**Kurtogram Golden Nelder-Mead**
F_c	2343.7266	2343.7266	1650.41/1445.30	1701.23	1709.15	1705.49
Q	-	0.7071	0.7215/0.3451	0.6541	0.7035	0.6993
N	-	4	4/4	4	4	4
R_p	-	1 (dB)	1.3785/0.5752	1.841	1.053	0.9932
Time(s)	1.89	0.065	89.344/13.349	98.854	65.126	13.178
Kurt	1.20	19.37	23.12/13.70	24.15	25.94	24.45

In Table 4.1, the computation time improvement is from 92.35 sec (kurtogram-GA-Nelder-Mead combination) to 10.17 sec (kurtogram-golden-Nelder-Mead combination) or 89%. As shown in Tables 4-5-1 to 4-5-4 of Figure 4.5, the total function evaluations are 46 for kurtogram-golden-Nelder-Mead combination. For kurtogram-GA-Nelder-Mead combination, the function evaluations are more than 300. GA implementation is not discussed in this chapter as we have used MATLAB® GA toolbox for comparison purposes. In Table 4.2, the computation time improvement is from 98.85 sec (kurtogram-GA-Nelder-Mead combination) to 13.17 sec (kurtogram-golden-Nelder-Mead combination) or 86.6%. Based on the above two cases, we can say, the proposed algorithm gives approximately 87% improvement in speed with almost no degradation in kurtosis values.

The faulty pulses emanating from a particular fault in the gearbox is the result from a certain vibration mode of gears. These pulses should be present in the same bandwidth throughout the presence of the fault. The adaptive concept we have used here is to detect the faulty bandwidth and tune a band-pass filter in order to maximize the kurtosis. Tuning time is a critical factor for mission critical applications as this plays an important role for early and timely detection of faults. After the filter is tuned, it can be used to extract the faulty information from the same bandwidth and from the same vibration mode.

4.2.3 Morlet Wavelet Filter

Wavelet transform is successfully used in de-noising vibration signals with huge background noise and increasing SNR. Wavelet transform provides a joint time-frequency analysis paradigm for signal filtering. Wavelet transform is also suitable for the detection of transient faults present in the vibration signals emanating from a faulty machine. We use an adaptive wavelet filter based on Morlet wavelet for signal de-noising. Morelet wavelet is a sinusoidal function with an exponential decay on both sides and very much like a faulty pulse in shape. When the Morlet wavelet with a suitable shape and centre frequency is convolved with the vibration signal, the filtered signal is obtained. The filtered signal reveals the faulty pulses that match with the shape of the wavelet used [81]. To get good performance and optimized results, the design parameters

of the Morlet wavelet must be selected carefully. As discussed in Chapter 2, previous studies on this problem have used different selection methods based on different objective functions like Shannon entropy, kurtosis maximization and singular value decomposition. However, these studies lack proper optimization techniques in selecting suitable values of Morlet wavelet design parameters. In [82], the authors used entropy as an objective function to be minimized in searching optimized wavelet parameters. The authors selected a range of the wavelet parameters that encapsulates the minimum of the entropy, and evaluated the objective function by varying the parameter values in pre-determined steps. This approach is quite cumbersome and time consuming, especially in real time fault detection applications. Similarly in [83], the authors used energy to entropy ratio as an objective function to maximize and follow the same step-wise approach described above to optimize the objective functions. The authors have estimated the computational complexity of their proposed method as $O\left(n^3 \log_2(n)\right)$. Where n being the total number of data points in the signal. However, they have reduced the computational complexity to the order of $O\left(n^2 \log_2(n)\right)$ by choosing a fixed bandwidth for practical applications. The computational complexity of the filtering process can be further reduced by choosing a suitable optimization framework that converges faster in the solution search space. In this section, we use optimization methods like golden section search and Nelder-Mead in order to select optimized values of Morlet wavelet design parameters.

Wavelet transform is a convolution or inner product between the signal and that of the wavelet family. The wavelet family contains one mother wavelet and several daughter wavelets. The daughter wavelets are derived from mother wavelet by translation and dilation. Let $\psi(t)$ be the mother wavelet, the daughter wavelets are obtained by $\psi_{a,b}(t) = \psi(t-b)/a$, where a is the scale parameter and b is time translation. The Morlet mother wavelet is defined as

$$\psi(t) = \exp(-2\pi^2(\nu - \nu_0)^2), \tag{4.8}$$

which is a complex wavelet and can be decomposed into real and imaginary parts as

$$\psi_r(t) = \frac{1}{\sqrt{2\pi}} \exp(-\beta^2 t^2 / 2) \cos(2\pi \nu_0 t)$$
$$\psi_i(t) = \frac{1}{\sqrt{2\pi}} \exp(-\beta^2 t^2 / 2) \sin(2\pi \nu_0 t), \tag{4.9}$$

Where ν_0 is a constant and β is a shape parameter. Generally, only the real part of the Morlet wavelet is used. The real part of the Morlet wavelet is a cosine signal that decays exponentially on both the left and the right sides, and its function shape is very much similar to an impulse. This similarity makes the Morlet wavelet very attractive and widely applied in mechanical fault diagnostic applications. A daughter Morlet wavelet is obtained by time translation and scale dilation of the mother Morlet wavelet as [81]

$$\psi_{a,b}(t) = \exp(-\beta^2 (t-b)^2 / 2a^2) \cos(\frac{\pi(t-b)}{a}). \tag{4.10}$$

Where, a is scale parameter for dilation and b is time translation. By carefully choosing the Morlet wavelet design parameters, a daughter wavelet that matches in shape with the faulty pulses present in the signal can be constructed. When the signal is convolved with Morlet wavelet, the filtered signal is obtained, and the faulty information present in the signal is revealed. Figure 4.8 shows the Morlet wavelet with parameters $\beta = 0.05$, $a = 3$ and $b = 0$.

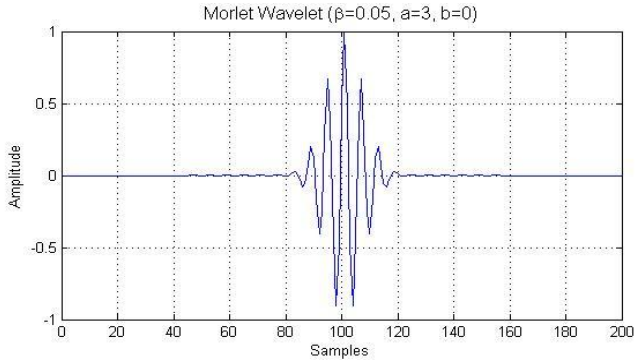


Figure 4.8: The Morlet Wavelet

One of the important properties of fast Fourier transform (FFT) is that convolution in one domain is a multiplication in another domain. Filtering operation is done by converting the time domain signals into the frequency domain and performing multiplication operation as

$$W(a,b) = F^{-1}(X(f)\psi^{*}(f)). \tag{4.11}$$

Where, $X(f)$ and $\psi(f)$ are the Fourier transforms of time signals $x(t)$ and $\psi(t)$, respectively, and $F^{-1}(.)$ is the inverse Fourier transform. Figure 4.9 shows the wavelet transform operation.

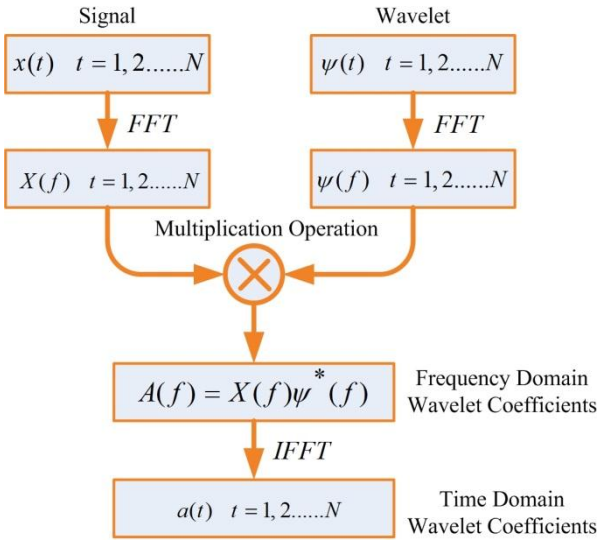


Figure 4.9: The Wavelet Analysis

4.2.3.1 Simulations and Discussions

The vibration data are taken from NREL presented in Section 1.10.1. In this section, we use signals from test case "CM_2a" listed in Table 1.1. The data are emanating from the sensor AN3, which is located on input side of the planetary gearbox inside the wind turbine's mechanical drive train assembly. One dimensional golden section search with parabolic interpolation is first used to optimize the shape factor β and the scale parameter a one by one and keeping the time translation $b = 0$. The values of β and a are passed to Nelder-Mead derivative free search for fine tuning and maximizing the kurtosis as the objective function.

Figure 4.10 shows the flow chart of the adaptive Morlet wavelet filtering. The simulation results are shown in Figure 4.11. In Figure 4.11(a), a raw vibration signal emanating from the gearbox is shown. Figure 4.11(b) plots the variation in the kurtosis value against scale

parameter a and shape parameter β. In Figure 4.11(b) the optimum value of kurtosis is estimated to lie at $\beta = 0.0005$ and $a = 34$. This is the global optimum that will be searched by the optimization process. The optimum values of a and β for the kurtosis maximization are searched by the one-dimensional golden section search one by one.

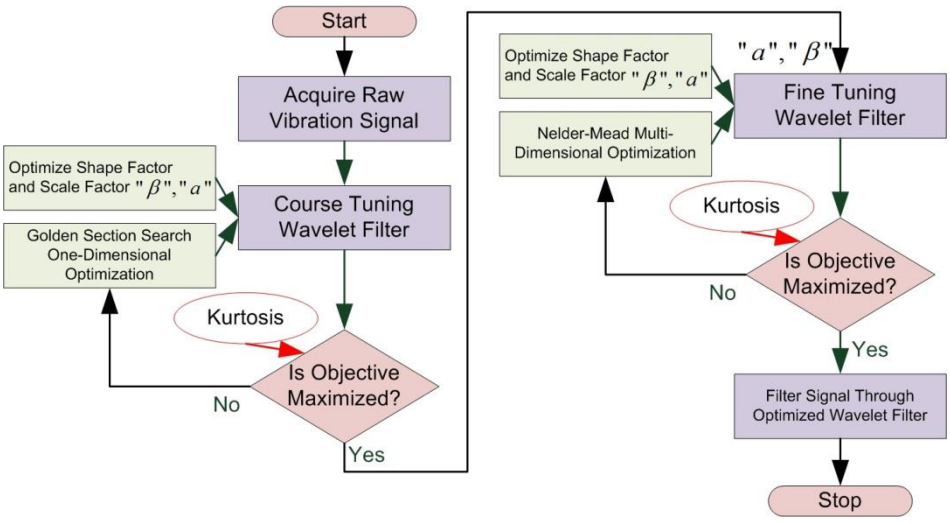


Figure 4.10: Flow Chart of the Adaptive Morlet Wavelet Filtering

These values are then passed as a starting point to the Nelder-Mead method for fine tuning the results and finding maxima in Figure 4.11(b). Figure 4.11(c) shows the Morlet wavelet designed with final values searched by the Nelder-Mead method. This is the optimized wavelet to reveal the faulty pulses present in this particular vibration signal. The filtered signal is shown in Figure 4.11(d). The faulty pulses are clearly revealed in the filtered signal and the kurtosis value is maximized to 4.66. Table 4-11-1 in Figure 4.11 shows the results of the one-dimensional run for the scale parameter a, keeping the shape parameter $\beta = 0.0005$. The algorithm stops in 7 iterations, and the kurtosis value of 3.59 is found for the scale parameter $a = 32.89$. In table 4-11-2 in Figure 4.11, the value of the shape parameter β is varied while keeping $a = 32.89$. The algorithm stops in 7 iterations and maximizes the kurtosis to 4.64 for $\beta = 0.00046$. In table 4-11-3 in Figure 4.11, the Nelder-Mead multi-parameter optimization method is used where $a = 32.89$ and $\beta = 0.00046$ are passed as initial values. The Nelder-Mead algorithm converges in 10 iterations and finds the kurtosis value of 4.667 for $a = 33.48$ and $\beta = 0.0005$. The whole

optimization process takes 2.3 seconds on an Intel Core Duo, 2.27 GHz processor with 4GB of RAM.

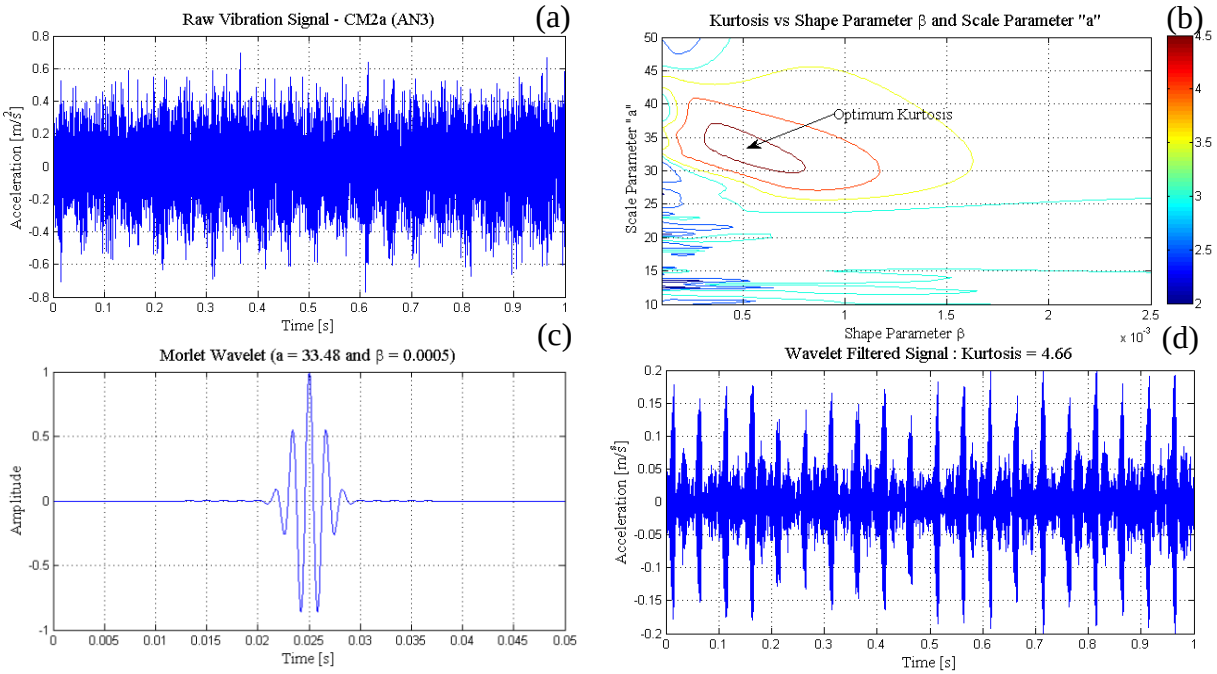


Table 4-11-1: One-Dimensional Search a

Count	a	Kurt	Method
1	25.27	2.326	Initial
2	34.72	3.009	Golden
3	33.15	3.491	Golden
4	32.67	3.503	Parabolic
5	32.90	3.592	Parabolic
6	32.89	3.593	Parabolic
7	32.89	3.593	Parabolic

Table 4-11-2: One-Dimensional Search β

Count	β	Kurt	Method
1	0.00082	4.3372	Initial
2	0.00127	3.8108	Golden
3	0.00054	4.6581	Golden
4	0.00037	4.5311	Golden
5	0.00058	4.6405	Parabolic
6	0.00051	4.6630	Parabolic
7	0.00046	4.6412	Golden

Table 4-11-3: Nelder-Mead

Count	Eval.	Kurt	Procedure
1	1	4.6392	Initial simplex
2	3	4.6541	Contract outside
3	5	4.6596	Contract outside
4	7	4.6635	Contract outside
5	9	4.6668	Reflect
6	10	4.6668	Contract inside
7	12	4.6668	Contract outside
8	14	4.6668	Reflect
9	16	4.6673	Contract inside
10	18	4.6673	Contract inside

$a = 33.48$, $\beta = 0.0005$

Figure 4.11: Adaptive Morlet Wavelet Analysis - Optimization

4.3 Conclusion

In this chapter, a conventional optimization method for real time gearbox fault detection is proposed. The method uses two types of adaptive filtering approaches, the Chebyshev band-pass filter and the Morelet wavelet filter. The Chebyshev band-pass filter method uses a combination of kurtogram, one dimensional and multi-dimensional optimization techniques. The Morelet wavelet approach uses only one dimensional and multi-dimensional optimization techniques. Both techniques are proven to find faulty information in a reasonable amount of time, and are suitable for real time applications. As compared to the Chebyshev adaptive band-pass filtering, the wavelet adaptive filtering is computationally less expensive and approximately ten times faster. The proposed method demonstrates a reasonable computational complexity and improves response time, which proves its applicability for real time fault detection in mission critical applications. One of the main contributions of this chapter is a combination of different conventional techniques to speed up the optimization process. We use kurtogram for estimating the faulty band width in the signal and use golden section search to reduce the solution search space for the Nelder-Mead method. This gives an advantageous start to the Nelder-Mead method and it converges faster.

In the following chapter, we will harness the optimization powers of evolutionary algorithms to tackle the background noise problem in vibration signals.

CHAPTER 5

Evolutionary Optimization in Background Noise Removal

5.1 Motivation

Biologically inspired or nature inspired optimization algorithms are extensively used in vibration based fault detection. Because of their inherently slow nature, they are not suitable for real time vibration based fault detection. However, we can harness their extraordinary capabilities by carefully reducing the solution search space. Researchers have proposed different flavours of evolutionary algorithms for speedy convergence. They combined the evolutionary algorithms with other algorithms to reduce the solution search space and speed up the convergence. Chapter 2 gives an overview of similar work in this field. In this chapter, we explore two types of evolutionary algorithms as part of the research methodology: real coded genetic algorithm (RCGA) and particle swarm optimization (PSO) as shown in Figure 1.8. RCGA is a flavour of GA that uses real values of design variables as compared to the binary coded GA that converts the design variable into binary strings for evolution process. As described in Chapter 2, PSO is less computationally expensive as compared to GA. We combine RCGA and PSO with conventional one-dimensional search techniques to speed up the convergence process and reduce the solution search space. The objective function we use is kurtosis maximization as described in Section 4.2.1. The RCGA tunes the Chebyshev band-pass filter presented in Chapter 4, and we apply PSO on both the Chebyshev band-pass and Morlet wavelet filters.

5.2 Fault Detection Using Real Coded Genetic Algorithm (RCGA)

In fault detection using RCGA, we follow the same flow as shown in Figure 4.1 but use RCGA instead of the multi-dimensional Nelder-Mead method for fine tuning. The kurtogram and one-dimensional optimization methods reduce the solution search space for RCGA. The RCGA itself provides a framework to accelerate the search process as it works on real coded values instead of binary coded bit strings.

5.2.1 Real Coded Genetic Algorithm (RCGA)

Genetic algorithm (GA) is a general purpose derivative free global search algorithm. GA uses principles of natural evolution or survival of the fittest. Holland first described the basic principles of GA in 1975 [98]. In GA, the idea is to maintain and evolve a population of knowledge structures. The evolution takes place through a process of competition and controlled variations called crossover and mutation operators. Each structure in the population represents a candidate solution to a particular problem at hand and has an associated fitness value to determine which structure can be used to form a new competitive population. In the evolution process, a subset of relatively good structures is selected for reproduction and replaces a relatively bad solution in the new population. Crossover operator makes use of the information contained into parents and combines them in a new population in order to increase the average quality of the population. The mutation operator randomly changes the new individuals to help avoid local extremum.

Figure 5.1 shows a flow chart of the basic genetic operations. Here, an initial random population of candidate solutions is generated in the form of chromosomes. Each chromosome is ranked according to its fitness value. The fitness value is calculated based on pre-defined objective functions to be optimized. Individuals based on their fitness values are selected and evolved by performing crossover and mutation operators. The process continues until required criteria are met. Genetic algorithms can be designed to evolve multiple parameters at the same time. The tuning of band-pass filters for vibration based features extraction presents two main restrictions.

- The evaluation is based on multiple parameters and sometimes on multiple objectives (kurtosis maximization, pass band ripple reduction, etc.). This fact adds complexity to the search space because we must obtain the best trade-off among different criteria.
- The evolved band-pass filter accuracy is assessed by means of simulations which usually take a long time. This causes the run time of the algorithms to be extremely long.

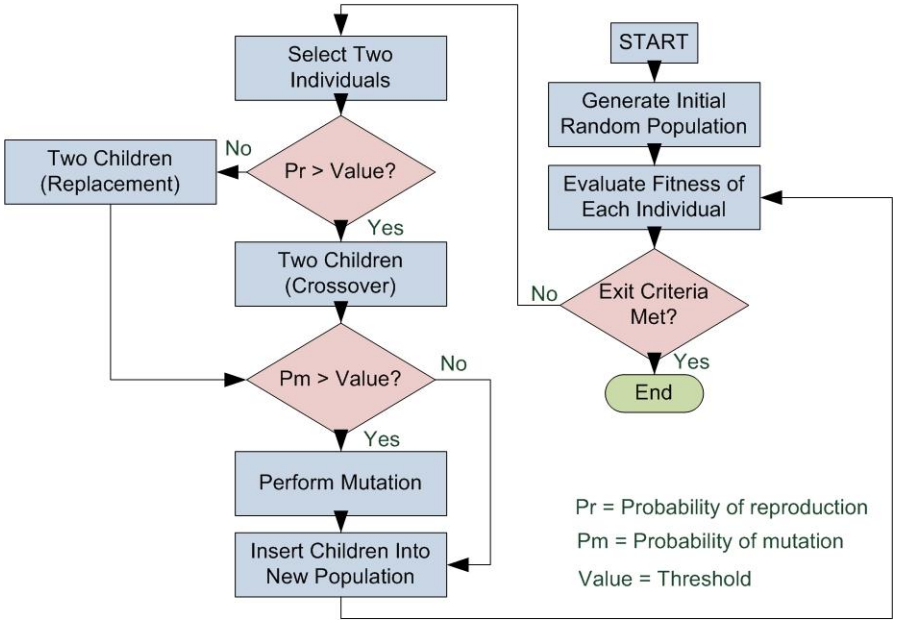


Figure 5.1: Genetic Algorithm Flow Chart

The first restriction will be solved by using a multi-parametric and multi-criteria GA optimization. The GA optimization allows working with fitness functions comprised of different competitive objectives. In this case, we could obtain not only an optimal solution, but also a set of different possible solutions. In order to solve the second restriction, we use two approaches. First, a real coded or integer coded genetic algorithm (RCGA) to reduce the computational complexity of encoding/decoding to/from binary strings. Second, penalty multipliers are included in the objective function to make unfit chromosomes die soon. Section 4.2.1 explains the penalty function phenomenon in detail. This will considerably reduce the run time for the genetic algorithm search. We also use a steady-state RCGA, which involves selecting two of the best individuals in the

population and combines them to obtain the two off-springs. This approach improves the convergence performance and simultaneously decreases the number of evaluations. The following genetic operators are used.

5.2.1.1 Encoding

The genetic algorithm used in this work is real coded to decrease the computational complexity and increase the convergence speed. Here, real coded numbers represent the parameters of the band pass filter in the evolution process.

Band-pass filter parameters chromosome: [F_c Q R_p N *]*

5.2.1.2 Selection

We use a very simple selection approach called a roulette-wheel selection. The roulette-wheel selection is also named as a stochastic sampling with a replacement. In the roulette-wheel selection, a continuous segment of a line maps the individuals in such a way that each individual's segment is equal in size to its fitness. An individual whose segment spans a generated random number is selected for next generation. The process is repeated until the desired number of individuals is obtained (a mating population).

5.2.1.3 Arithmetic Crossover

An arithmetic crossover is used for reproduction and creation of new offspring from the mating of two selected parents. We use an arithmetic crossover operator that defines a linear combination of two chromosomes [43]. Two chromosomes, selected randomly for a crossover, C_i^{Gen} and C_j^{Gen} may produce two offspring, C_i^{Gen+1} and C_j^{Gen+1}, which is a linear combination of their parents as

$$\begin{aligned} C_i^{Gen+1} &= \gamma.C_i^{Gen} + (1-\gamma).C_j^{Gen} \\ C_j^{Gen+1} &= \gamma.C_j^{Gen} + (1-\gamma).C_i^{Gen}. \end{aligned} \tag{5.1}$$

Where, C^{Gen} and C^{Gen+1} are the individuals from old and new generations, respectively and γ is a weighting factor which governs the dominant individual in the reproduction process. The value of γ ranges from 0 to 1.

5.2.1.4 Arithmetic Mutation

The mutation operator is used to inject new genetic material into the population, and it is applied to each new structure individually. A given mutation involves the random alteration of each gene with a small probability. We generate a random real value which makes a random change in the k^{th} element selected randomly from the chromosome. In arithmetic mutation, we involve two chromosomes but mutate only one according to some mutation probability. The two chromosomes C_i^{Gen} and C_j^{Gen}, selected for the crossover in the previous section, can undergo mutation and they have equal probability of selection for the mutation. Let us say, two offsprings, C_i^{Gen+1} and C_j^{Gen+1} are produced in crossover operation, and the k^{th} element of C_i^{Gen+1} is selected for mutation. The mutation will occur as follows

$$C_i^{Gen+1} = \left\langle C_{i1}, C_{i2}, ...C_{i(k-1)}, \gamma.C_{jk}^{Gen+1} + (1-\gamma).C_{ik}^{Gen+1}, ...C_{i(n-1)}, C_{in} \right\rangle. \quad (5.2)$$

5.2.1.5 Elitism

A process to select better individuals is called elitism. In other words, elitism selects individuals with a bias towards the better ones and allows the solutions to get better over time.

5.2.1.6 Simulations and Discussions

The vibration data are taken from NREL presented in Section 1.10.1. In this section, we use signals from test case "CM_2a" as listed in Table 1.1. The data are emanating from the sensor AN3, which is located on input side of the planetary gearbox inside the wind turbine's mechanical drive train assembly.

We change F_c, Q and R_p parameters for the band-pass filter and apply accelerated one dimensional search for the rapid convergence and coarse tuning. Fast kurtogram gives us an initial estimate for the centre frequency F_c of the band-pass filter. Initially, we set $Q = 0.707$, $R_p = 1\text{dB}$ and order $N = 4$. Then run a golden section search combined with

parabolic interpolation to search for a value F_c of the band-pass filter which maximizes the kurtosis of the filtered signal. Consider the case in Table 5.1, the value of centre frequency returned by kurtogram for the particular signal is $F_c = 546.87\text{Hz}$ in 1.076 seconds. Filtering the signal with these filter parameters ($F_c = 546.87\text{Hz}$, $Q = 0.707$, $R_p = 1\text{dB}$ and $N = 4$) gives kurtosis value of 4.55. Running the accelerated one dimensional search with centre frequency range $F_c = [100 \;\; 800]\text{Hz}$, the optimum is found at $F_c = 527.37\text{Hz}$ and kurtosis value jumps to 4.85. The search only takes 0.24 seconds because of reduction in search space by the kurtogram and accelerated nature of the golden section search and parabolic interpolation combination. The termination criteria used to stop the execution is when further improvement is less than 0.001. Let us freeze the value of the centre frequency at $F_c = 527.37\text{Hz}$ and apply the one dimensional search to the filter quality parameter Q in the range $[0.1, 2]$ and look for any improvement in the final kurtosis value. There is no considerable improvement in the kurtosis value. Therefore, we keep the value of $Q = 0.707$. Further, we apply one dimensional search to the filter pass band ripples R_p in the range $[0.1, 10]\text{dB}$. There is still no considerable improvement in the kurtosis value. Therefore, we keep the value of $R_p = 1$ and use the RCGA for further fine tuning of the filter parameters. The RCGA takes all the four parameters including filter order N, and applies multi-parameter optimization. We also restrict the pass band ripples below a maximum value of $\xi = 2.5\text{dB}$ and formulate the objective function for RCGA as

$$\max F_{obj} = \frac{\frac{1}{n}\sum_{i=0}^{n-1}\left(x_t(i) - \mu\right)^4}{\sigma^4} - \eta\left[\max^2\left\{0, R_p - \xi\right\}\right]. \tag{5.3}$$

Where, $\dfrac{\frac{1}{n}\sum_{i=0}^{n-1}\left(x_t(i) - \mu\right)^4}{\sigma^4}$ is the kurtosis and $\eta\left[\max^2\left\{0, R_p - \xi\right\}\right]$ is a penalty function with η as a penalty multiplier, R_p is the evolved ripple value and ξ is the maximum allowable ripples in the pass band. If $R_p > \xi$, the term $\eta\left[\max^2\left\{0, R_p - \xi\right\}\right]$ will subtract

a squared penalty value from the objective function and restrict the objective function to be maximized. This step also makes unfit chromosomes in the population worse so that they can be replaced or die soon and the population converge soon. The RCGA method with Population Size = 20 and No. of Generations = 100 converges in average of 17 seconds and produces the final values of the filter parameters as shown in Table 5.1. In Table 5.1, different runs of the RCGA are shown. RCGA1-RCGA4 runs are for reproduction plan of steady state, replace worst and elitism is ON. RCGA5-RCGA8 runs are for reproduction plan of steady state, replace random and elitism is OFF. RCGA9-RCGA12 runs are for reproduction plan of steady state, replace random and elitism is ON. The kurtosis value increases from 4.85 to 5.58 in the case "RCGA4" in Table 5.2. To avoid RCGA be stuck at local optimum, we make use of the one dimensional accelerated search. The strategy gives a good starting point for the RCGA and ensures convergence at a global optimum. The experiment shows that if we use RCGA just after kurtogram estimation and use $F_c = 546.87\text{Hz}$, $Q = 0.7071$, $R_p = 1\text{dB}$ and $N = 4$ as a starting point, the method sometimes converges at local optimum. As shown in Table 4.2, the kurtogram, the golden section search and the conventional GA takes 65.12 seconds on Intel Core Duo 2.27 GHz platform with 4GB RAM. A combination of the kurtogram, the golden section search and the RCGA takes an average of 17 seconds, as shown in Table 5.1, to find an optimum kurtosis value. Hence, there is an improvement of 74% in speed as compared to classical GA and kurtogram combination.

A raw vibration signal and its kurtogram is shown in Figure 5.2(a) and Figure 5.2(b), respectively. After initial parameters estimation from kurtogram, a band-pass filter is initialized in Figure 5.2(c) and a filtered signal is obtained in Figure 5.2(d). The faulty pulses in the signals are quite prominent but the band-pass filter is not yet tuned. The kurtosis value reaches at 4.55. The tuned band-pass filter is then obtained through golden-RCGA tuning combination as shown in Figure 5.2(e) and the raw signal is filtered again in Figure 5.2(f). Faulty pulses are now more prominent and the kurtosis value calculated is 5.58. Figure 5.3 shows the evolution in the cost function for all cases of the RCGA runs. In Figure 5.3, we can easily see the effect of elitism in the RCGA runs. If elitism is OFF and the reproduction plan is random replacement, the cost function trend can even

go towards opposite of the optimum as shown in Figure 5.3(b). It is also evident from Figure 5.3(a) that "replace worst, elitism ON" strategy converges the RCGA faster than others (Figure 5.3(b) and Figure 5.3(c)). In Figure 5.3(a), the RCGA4 run case converges faster in 26 generations with maximum value of the fitness function as compared to any other RCGA run case in Figure 5.3.

Physical inspection of the gearbox confirms that there is a severe scuffing on the HSS pinion gear revolving at the speed of 1200 RPM as shown in Figure 5.4. The faulty pulses emanating from a particular fault in the gearbox are the result from a certain vibration mode of gears and these pulses should be present in the same bandwidth throughout the presence of the fault. The adaptive concept used is to detect the faulty bandwidth and tune a band pass filter in order to maximize the kurtosis. Tuning time is a critical factor here and for mission critical applications this plays an important role for early and timely detection of faults.

Many researchers have proposed different techniques to avoid local minima and stagnation in the search process of evolutionary algorithms. Although, it helps to achieve a better convergence but at the same time it slows the convergence speed. Therefore, it is not suitable for time critical applications, where time is more important parameter as compared to the quality of the results. In order to increase the convergence speed and establish a trade-off between time complexity and result quality, we proposed to reduce the solution search space for the RCGA through using the kurtogram and the golden section search methods. This gives the RCGA an advantageous start and the convergence takes less time.

Table 5.1: Filter Tuning Algorithm – Kurtogram, Golden Search and RCGA: 1200 RPM, (Accelerometer Location: AN3 Gearbox High Speed Shaft (HSS) Pinion) (Platform: Intel Core Duo 2.27 GHz, 4GB RAM)

	F_c **(Hz)**	Q	N	R_p **(dB)**	**Kurtosis**	**Time (sec)**
Kurtogram	546.87	0.707	4	1	4.55	1.076
Golden Section	527.37	0.707	4	1	**4.85**	0.240
Pop. Size: 20, Generations: 100, Cross-Over Prob.: 0.8, Mutation Rate: 0.01, Reproduction: SS Replace Worst, Elitism: ON						
RCGA1	416.40	0.536	3	2.29	5.149	15.00
RCGA2	371.93	0.546	7	1.78	5.200	17.00

RCGA3	475.15	0.723	7	2.53	5.454	17.00
RCGA4	404.61	0.605	8	2.41	**5.581**	14.00
Pop. Size: 20, Generations: 100, Cross-Over Prob.: 0.8, Mutation Rate: 0.01, Reproduction: SS Replace Random, Elitism: OFF						
RCGA5	453.58	0.696	9	1.58	5.087	19.00
RCGA6	451.42	0.693	9	1.92	5.100	19.00
RCGA7	318.39	0.378	2	0.50	4.811	15.00
RCGA8	414.82	0.536	3	1.35	5.019	15.00
Pop. Size: 20, Generations: 100, Cross-Over Prob.: 0.8, Mutation Rate: 0.01, Reproduction: SS Replace Random, Elitism: ON						
RCGA9	390.47	0.501	3	2.44	5.178	15.00
RCGA10	435.98	0.577	3	1.80	5.079	15.00
RCGA11	308.95	0.378	2	2.67	5.164	15.00
RCGA12	462.87	0.711	9	2.38	5.250	17.00

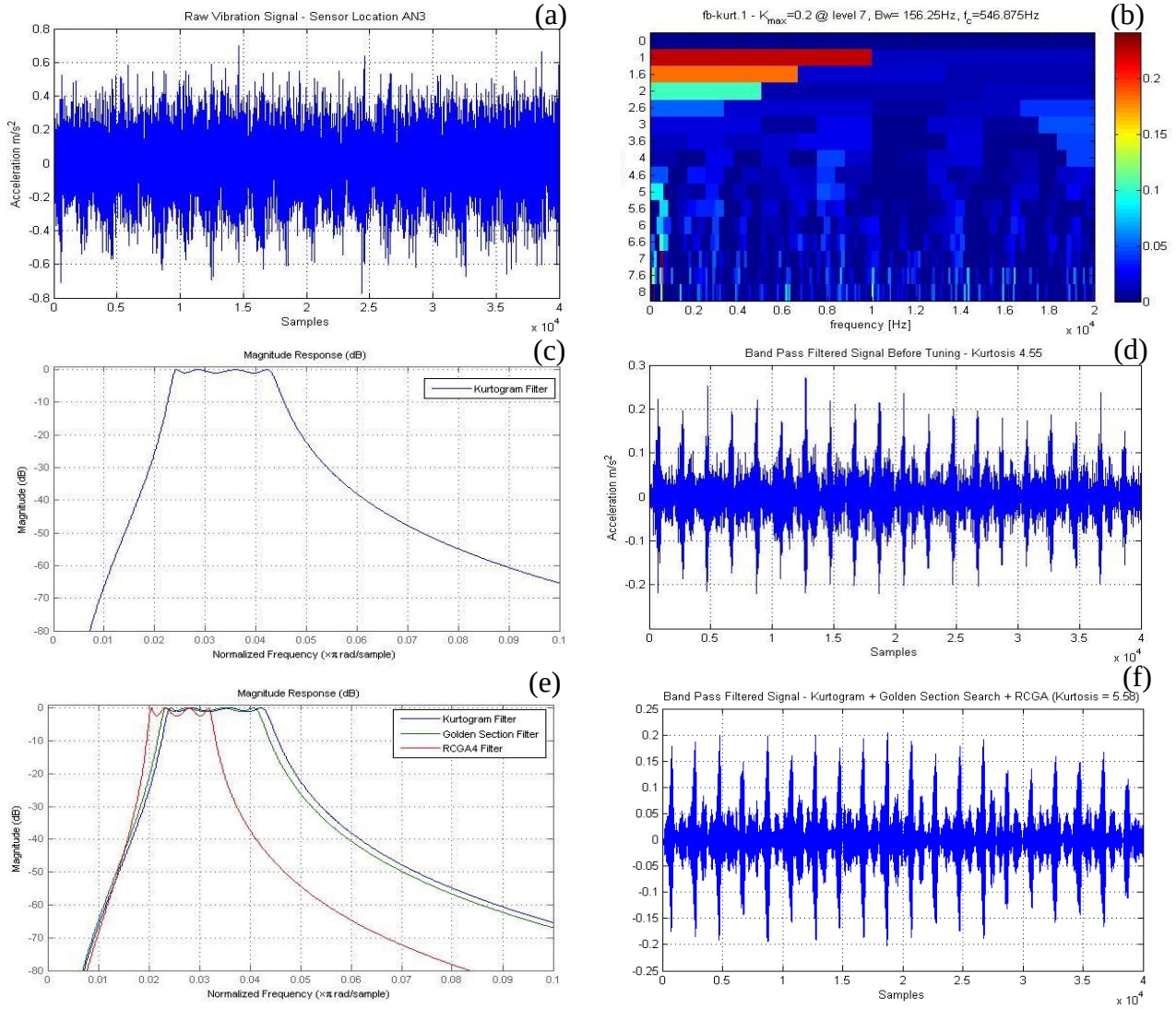


Figure 5.2: Band-Pass Filter Tuning (Case: Table 5.1 - RCGA4)

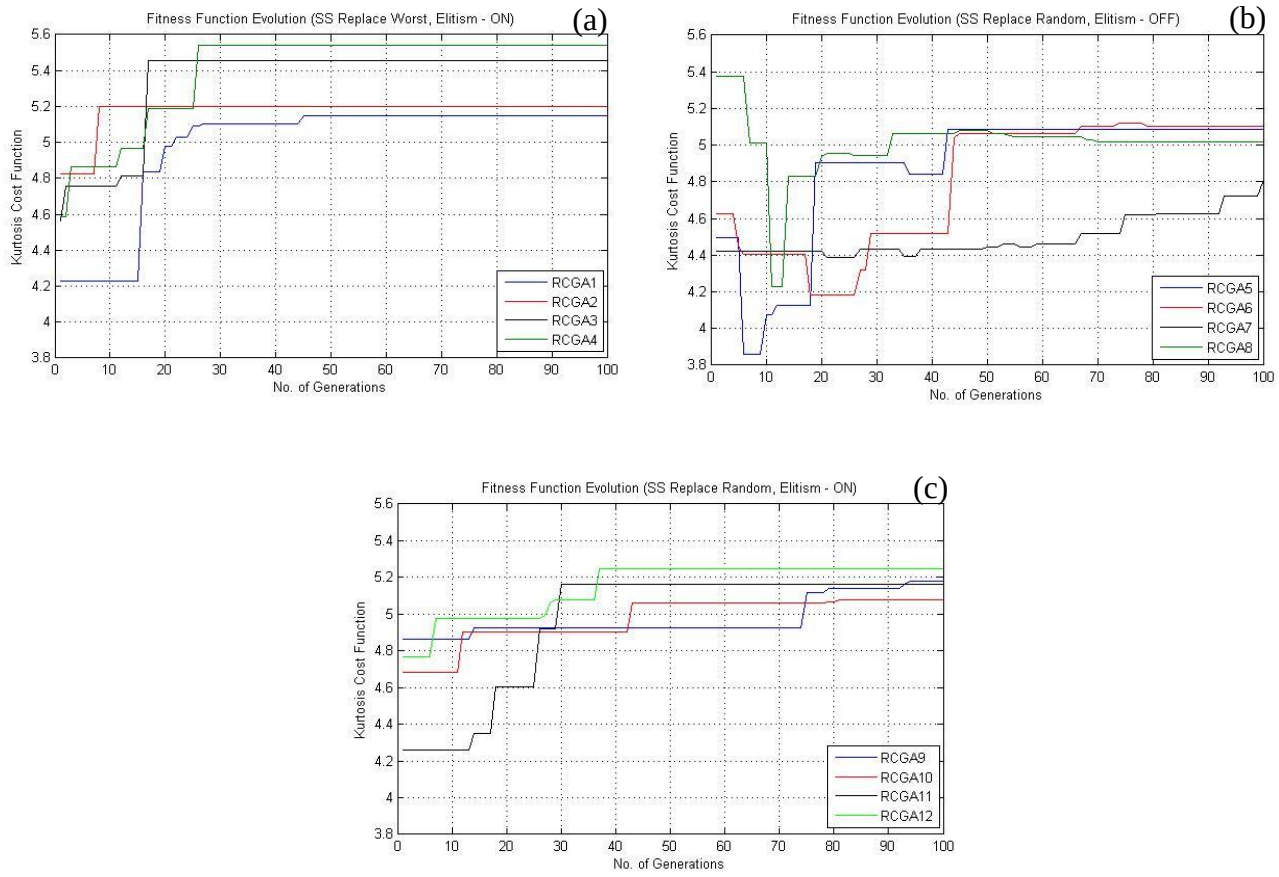


Figure 5.3: Fitness Function Evolution Trend (Case: Table 1 - RCGA1-RCGA12)

Figure 5.4: Severe Scuffing on the High Speed Pinion [15]

5.3 Fault Detection Using Particle Swarm Optimization (PSO)

In fault detection using PSO, we follow the same flow framework as shown in Figure 4.1. For PSO, the kurtogram and one-dimensional optimization methods reduce the solution

search space. PSO itself provides a framework to accelerate the search process as compared to GA.

5.3.1 Particle Swarm Optimization (PSO)

PSO is a population based stochastic optimization technique. PSO was first proposed by Kennedy and Eberhart in 1995 [50]. PSO mimics the social behaviour of birds and fish schooling. Similar to other population based stochastic optimization techniques, like GA, PSO starts with a population of random solutions and eventually converges to find an optimal solution. The population in PSO consists of particles that fly in n-dimensional solution search space and follow the position of the best particle. Each particle in the solution search space is described by its position vector $\mathbf{x}_i$ and velocity vector $\mathbf{v}_i$. The velocity and position update equations of the PSO are

$$\begin{aligned} \mathbf{v}_{k+1}^{i} &= w\mathbf{v}_{k}^{i} + d_1 r_1 \frac{\left(\mathbf{p}^{i} - \mathbf{x}_{k}^{i}\right)}{\Delta t} + d_2 r_2 \frac{\left(\mathbf{p}_{k}^{g} - \mathbf{x}_{k}^{i}\right)}{\Delta t}. \\ \mathbf{x}_{k+1}^{i} &= \mathbf{x}_{k}^{i} + \mathbf{v}_{k+1}^{i}.\Delta t \end{aligned} \tag{5.4}$$

Where $\mathbf{p}^{i}$ is the best position of each particle, $\mathbf{p}_{k}^{g}$ is the global best position of a particle in the swarm. The constants d_1 and d_2 are self-confidence factor and swarm-confidence factor, respectively. The value of d_1 and d_2 is normally taken to be in the range $[1,\ 2]$ as described in [56]. The parameters r_1 and r_2 are randomly generated and uniformly distributed in the range $[0,\ 1]$. This avoids any entrapment in a local optimum and provides a good convergence of the solution search space. Δt is a time step and can be taken as 1. In Equation (5.4), w is inertia weighting and is usually a linear descending function given by

$$w = w_{\max} - \left(w_{\max} - w_{\min}\right)\frac{T}{T_{\max}}. \tag{5.5}$$

Where T and $T_{\max}$ are current and maximum iteration, respectively. Each particle's best position is evaluated through a fitness function that maximizes the kurtosis. Figure 5.5 shows a description of position and velocity updates of the particles in PSO graphically.

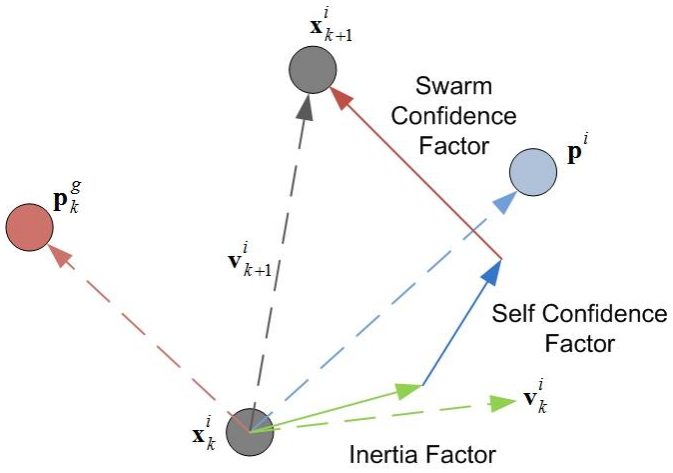


Figure 5.5: Velocity and Position Updates in PSO

5.3.1.1 Chebyshev Band-Pass Filter Optimized by PSO

The vibration data are taken from NREL presented in Section 1.10.1. In this section, we use signals from the test case "CM_2a" as listed in Table 1.1. The date are emanating from the sensor AN3, which is located on input side of the planetary gearbox inside the wind turbine's mechanical drive train assembly. The one-dimensional search from section 4.2.2.2 is used to search for the filter's centre frequency F_c that maximizes the objective function. Other filter parameters like the quality factor Q, the filter order N and the pass band ripples R_p are kept constant at values $Q = 0.707$, $N = 4$ and $R_p = 1$, respectively. Figure 5.6(b) gives an idea where the kurtosis maximizes at 4.85 for $F_c = 525\,\text{Hz}$ and found by one-dimensional accelerated search method. Figure 5.6(c) plots kurtosis against F_c and Q, where it is evident that the kurtosis maximization occurs in the range of $300 < F_c < 500\,\text{Hz}$ and $0.3 < Q < 0.7$. We then initialize the PSO search for parameter ranges: $500 < F_C < 1500\,\text{Hz}$, $0.3 < Q < 2$, $2 < N < 8$ and $1 < R_p < 6$. We see that the PSO maximizes the kurtosis to 5.18 with $F_c = 420.20\,\text{Hz}$, $Q = 0.54$, $N = 3$ and $R_p = 2.99$. Figure 5.6(d) plots the final filtered vibration signal with kurtosis value equal to 5.18. Figure 5.7 plots each particle's best position versus iterations. It is shown in Figure 5.7 that particle-1 reaches the optimum point first in 50 iterations and other

particles follow it. With particle-4, converged at iteration 80, the whole population converges at iteration 80. The PSO parameters used in this section are as follows.

No. of Iterations = 100
No. of Swarm Particles = 5
Inertia Weight w = Varies from 0.9 to 0.4
Self Confidence d_1 =1.496
Swarm Confidence d_2 = 1.4961
Parameters r_1 and r_2 = rnd[0-1]
Threshold for Success = 0.0001

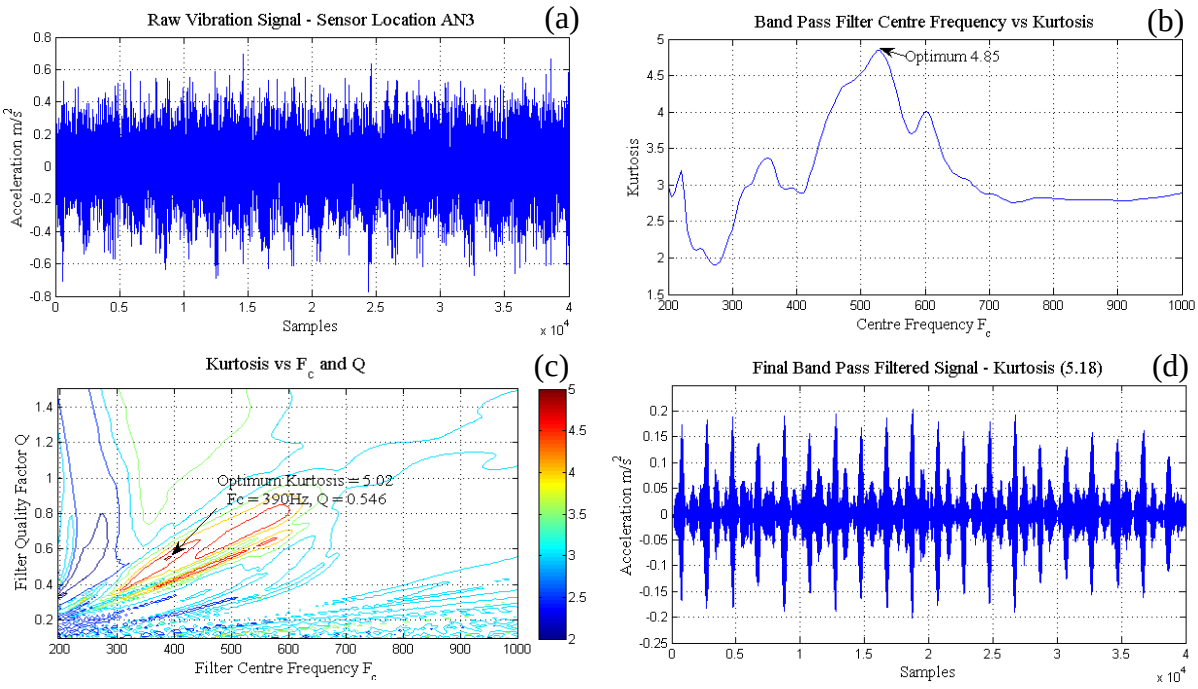


Figure 5.6: The Chebyshev Band-Pass Filter Optimized by PSO

5.3.1.2 Morlet Wavelet Filter Optimized by PSO

The main parameters we will tune for the Morlet wavelet are the shape parameter β and the scale parameter a. A suitably starting value for the scale parameter a is determined through the one dimensional accelerated search method. As seen in Figure 5.8(a), the kurtosis is optimized for the scale parameter $a = 35$ found by the one-dimensional search. Figure 5.8(b) plots the kurtosis against a and β. In Figure 5.8(b), it is evident that the kurtosis maximization occurs in the range of $10 < a < 50$ and $0.0001 < \beta < 0.02$. We

then initialize the PSO search for parameter ranges, $10 < a < 50$ and $0.0001 < \beta < 0.02$. The PSO maximizes the kurtosis to 4.66 with $a = 34.12$ and $\beta = 0.00047$. Figure 5.8(c) plots the final filtered vibration signal with kurtosis value equal to 4.66. The PSO parameters used in this section are as same as described in the previous section.

5.3.2 Comparison of Band-Pass and Wavelet Filters Optimized by PSO

The performance of the band-pass filter and the wavelet filter is compared in this section. As discussed before, we design the Chebyshev band-pass filter because of its speed. The band-pass filter has four parameters to be optimized by the PSO as compared to the Morlet wavelet that has two parameters to be optimized. Table 5.2 gives a comparison between the two.

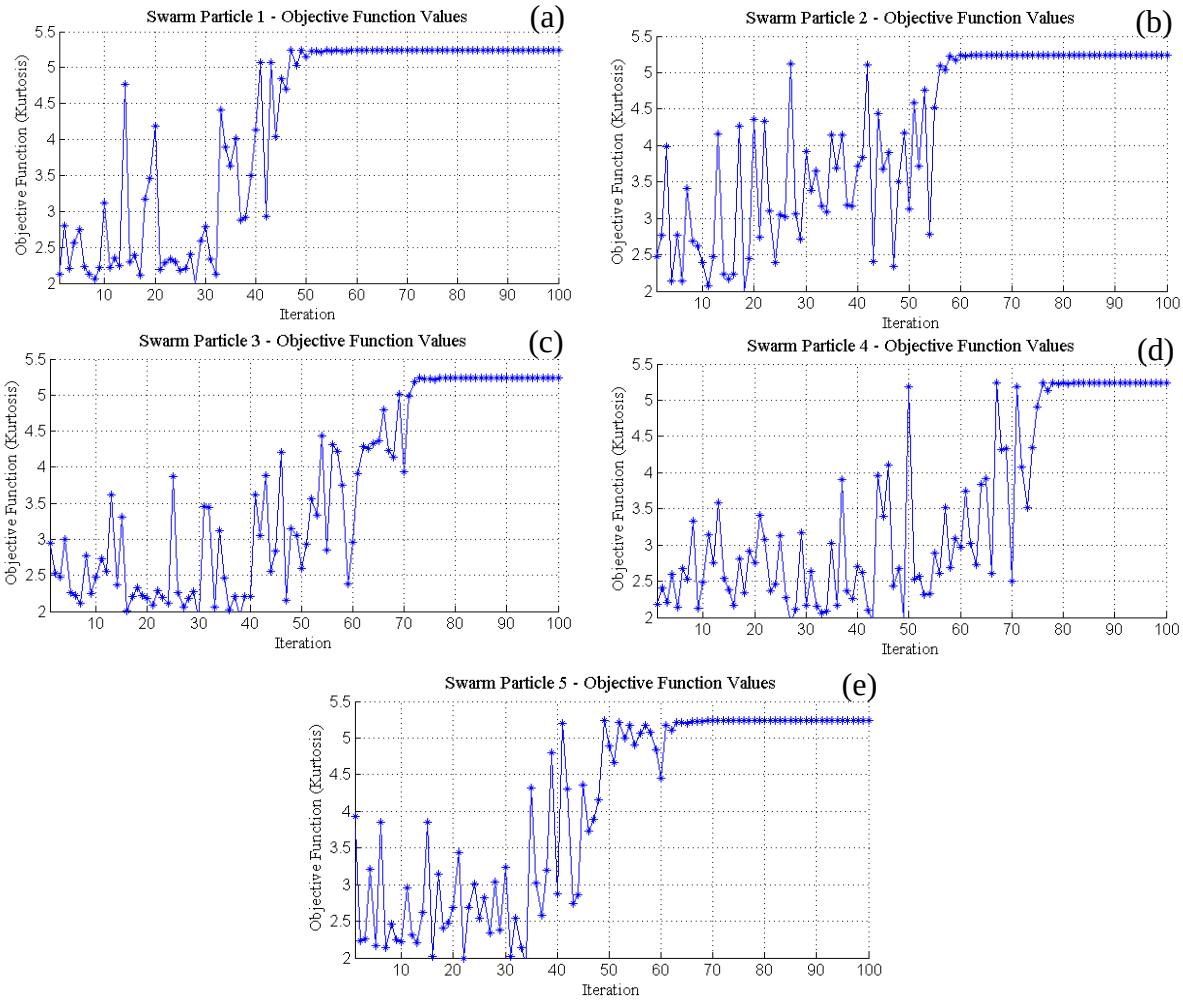


Figure 5.7: Swarm Particles Local Best Evaluations

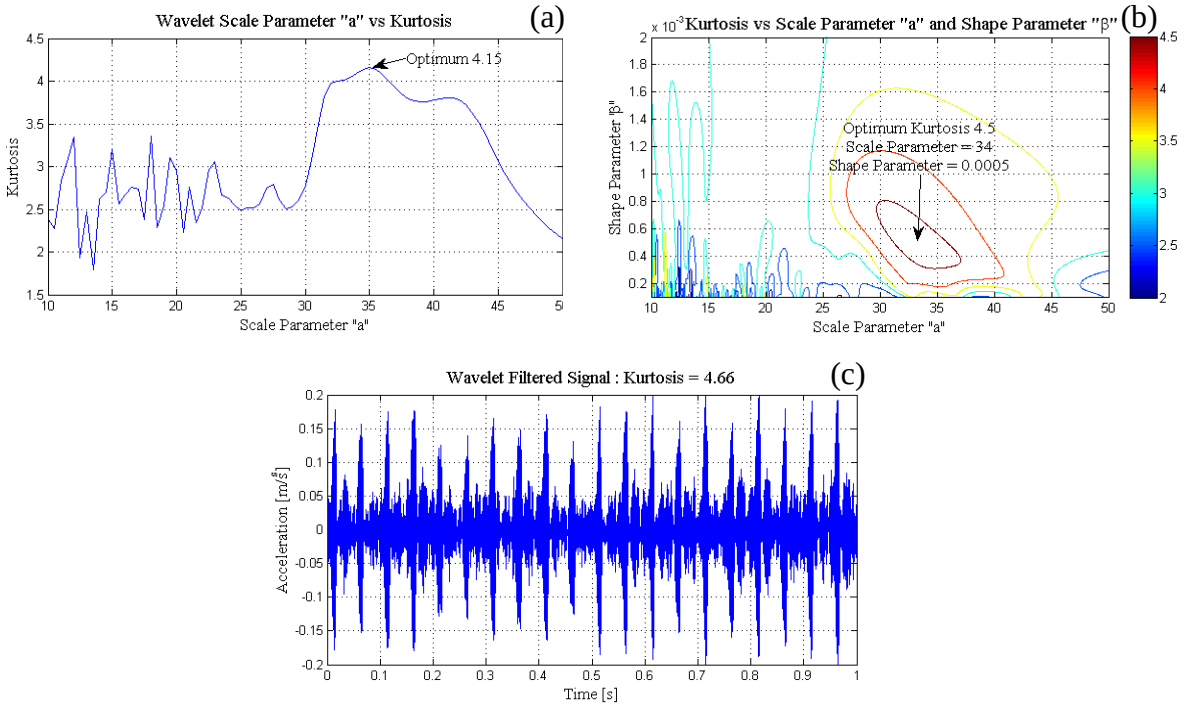


Figure 5.8: The Morlet Wavelet Filter Optimized by PSO

In Table 5.2, PSO takes 8.756 seconds to find an optimum value of the kurtosis in case of the Morlet wavelet filter and 14.654 seconds in case of the band-pass filter. On the other hand, the quality of the results degrades in the case of the Morlet wavelet filter. The Morlet filter finds a kurtosis value of 4.66, whereas, the band-pass filter finds a kurtosis value of 5.18. Also, the one-dimensional accelerated search takes less number of iterations to converge for the Morlet wavelet filter. Based on the application requirements, either the band pass filtering or the Morlet wavelet filtering can be chosen.

Table 5.2: Comparison of Band-Pass vs Morlet Wavelet Filters
Platform: Intel Core i7 CPU, 2Ghz and 8GB RAM

Parameters	Band Pass Filter	Morlet Wavelet Filter
Accelerated Search Dimensions	1 (F_c)	1 (a)
PSO Dimensions	4 (F_c, Q, R_p, N)	2 (a, β)
Filtering Time (s)	0.11307 s	0.0459 s
Accelerated Search Iterations	8	5
PSO Iterations to Converge	80	65
PSO Time (s)	14.654 s	8.756 s
Kurtosis Value	5.18	4.66

5.4 Conclusion

In this chapter, the evolutionary optimization methods such as RCGA and PSO are proposed for gearbox fault detection. The optimization process is based on the kurtogram, the one-dimensional accelerated search methods, and the evolutionary search schemes. Although, the evolutionary optimization algorithms exhibit strong ability to optimize multi-parametric function, they are inherently slow and are not suitable for real time applications. The kurtogram and the one-dimensional optimization techniques are used to reduce the solution search space and speed up the convergence process for the evolutionary algorithms. It is demonstrated that the developed hybrid optimization scheme has successfully achieved more than 70% improvement in convergence speed as compared to classical evolutionary algorithms. This is because of the three main reasons: (1) inclusion of conventional kurtogram technique to reduce the solution search space for the evolutionary algorithm, (2) use of real coded-genetic algorithm to avoid encoding and decoding from binary to real numbers and vice versa and (3) inclusion of a squared penalty in the objective function that makes unfit or bad chromosomes or swarms worse and avoids them to be used in next generations or to be followed by others.

After the fault detection process is complete and the faulty information is revealed. We extract the representative features that describe the health status of the gearbox. The following chapter develops such methods.

CHAPTER 6

Vibration Based Features Extraction for Gearbox Fault Detection

6.1 Motivation

Extracting representative features that best represent the health of a system is of utmost importance in any CM-FDS system. The extracted features should be sensitive to the machine condition and should reflect the changes in the machine behaviour or possible faults development over time. Features extraction is a first step before fault clustering, classification and diagnosis. Features extraction also helps in reducing the data dimensionality and prevents the over flooding of the information. In this chapter, we present vibration based features extraction methods for gears and bearings as shown in Figure 1.8. We develop three methods for time-based features extraction to represent the health of the gearboxes and present some existing frequency-based features for gears and bearings fault detection.

6.2 Gear's Features Extraction

Gears are an inseparable part of most of rotating machines in the industry today. Gears transfer torque and power from one shaft to another. They are essential components of any rotating machinery where the availability of the whole system depends on the smooth function of gears and gearboxes. Reliable monitoring techniques are required to detect incipient faults in gearboxes and to avoid catastrophic breakdowns.

6.2.1 Time Domain Analysis

Time-domain analysis is a powerful technique to detect incipient faults in gearboxes that cannot be detected by conventional frequency-domain analysis techniques. In this section, we develop transient based features extraction methods for gearboxes.

6.2.1.1 Feature-extraction Based on Shock Response Spectrum

Shock response spectrum (SRS) was first introduced by Maurice Biot in 1932 [99]. Since then, SRS was under development and use by the US military and R&D facilities as mechanical shock analysis. With the advent of today's modern high speed computers and signal processing techniques, it is now possible to simulate SRS of a structure based upon well-defined transient forces. SRS is the maximum response of a series of single degree of freedom (SDOF) systems of the same damping to a given transient signal. SRS actually gives the severity of transient information present in the signal. Therefore, it is useful in extracting transient features present in the vibration signatures of faulty machines. A combination of wavelet analysis and SRS is already proposed in [100], where authors have demonstrated the use of SRS to investigate transient nature of process variables and extract complex interaction patterns among them. Another technique with SRS and adaptive wavelet filtering is used for gearbox fault detection in [101]. The process of SRS features extraction is shown in Figure 6.1. The faulty pulses or transient information, revealed after adaptive filtering, is divided into blocks and is fed into a sequence of SDOF oscillator systems with different natural frequencies. Block division is done in such a way that each block contains at least one faulty pulse inside. We use wavelet peak detection to detect the faulty peaks and then make block division. A detailed description about wavelet peak detection can be found in [102]. The SRS applies the transient information as a base excitation to an array of SDOF systems, as shown in Figure 6.2. Here, the SDOF modeling for each block is performed individually.

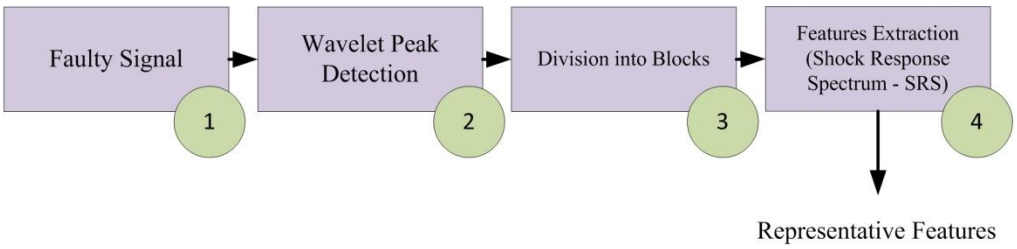


Figure 6.1: SRS Transient Features Extraction

In SDOF modeling, each system is assumed to have no mass-loading effect on the base input and consists of a mass with value M , spring with stiffness k and a damper with a

damping coefficient c. The resonance frequency f_N, the critical damping ratio ζ and the damped vibration frequency ω_d of each SDOF is calculated from the following equations:

$$\begin{aligned} f_N &= \frac{1}{2\pi}\omega_n = \frac{1}{2\pi}\sqrt{\frac{k}{M}} \\ \zeta &= c\frac{1}{2\sqrt{kM}} \\ \omega_d &= \sqrt{1-\zeta^2}\,\omega_n \end{aligned} \tag{6.1}$$

Where ω_n is natural frequency in radians. According to Smallwood [103], the above mentioned SDOF system can be simulated with the help of a digital recursive filter. The output of the filter using a sampled input is assumed to be a measure of the response of the SDOF system. The filter transfer function is presented as

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 - 2Cz^{-1} + E^2 z^{-2}}. \tag{6.2}$$

Where $E = e^{-\zeta\omega_n t}$, $C = E\cos K$ and $K = T\omega_d$. The interval T is the sampling time of the signal and the coefficients b_0, b_1 and b_2 can be calculated as

$$b_0 = \frac{1}{T\omega_n}\left[2\zeta(C-1) + \frac{(2\zeta^2-1)S}{\sqrt{1-\zeta^2}} + T\omega_n\right], \tag{6.3}$$

$$b_1 = \frac{1}{T\omega_n}\left[-2CT\omega_n + 2\zeta(1-E^2) - \frac{2(2\zeta^2-1)S}{\sqrt{1-\zeta^2}}\right], \tag{6.4}$$

and

$$b_2 = \frac{1}{T\omega_n}\left[E^2(T\omega_n + 2\zeta) - 2\zeta C + \frac{2(2\zeta^2-1)S}{\sqrt{1-\zeta^2}}\right]. \tag{6.5}$$

Where $S = E\sin K$. The damping ratio of each system is typically taken as $\zeta = 0.05$ and the natural frequency f_N is an independent variable. The calculation is performed for a number of independent SDOF systems, each with a unique natural frequency. Any arbitrary set of unique natural frequencies can be used for the shock response spectrum

calculation. A typical scheme, however, is based on a proportional bandwidth, such as 1/6 octave. This means that each successive natural frequency is $2^{1/6}$ times the previous natural frequency. Thus, a sample set of q calculation frequencies in units of Hertz would be 10, 11.2, 12.6, 14.1……,$[10\times 2^{1/6(q-1)}]$. Note that the bandwidth increases with frequency. Each of the natural frequencies in the above example represents the center frequency of an octave band. Continuing with the 1/6 octave format, the following relationships hold for the lower frequency f_l and the upper frequency f_u

$$\begin{aligned} f_l &= \frac{1}{2^{1/12}} f_c \quad Hz, \\ f_u &= 2^{1/12} f_c \quad Hz. \end{aligned} \tag{6.6}$$

Figure 6.2 presents the idea behind the SDOF modelling of transient information present in the filtered signal. As discussed above, the transient information from the filtered vibration signal is divided into blocks and fed to the SDOF system. The shock response spectrum of the SDOF is constructed according to Figure 6.2. The SRS for each block is calculated and maximum value $G_{\max}$ is taken. The following equation is used to extract the SRS features.

$$SRS = 10.\log_{10}\left(\frac{\frac{1}{N_B}\sum_{i=1}^{N_B}\left(G_{\max}\right)_i^2}{G_{ref}^2}\right). \tag{6.7}$$

Where, N_B is a number of blocks the signal is divided into, $G_{\max}$ is a maximum shock value in each block and G_{ref} is a reference shock value. If an engineering unit g or m/s^2, we select $G_{ref} = 1$and if the engineering unit is Pascal (pa), we select $G_{ref} = 20\times 10^{-6}$. A squared term for $G_{\max}$ is included to favour for large shock values. The SRS in Equation (6.7) is actually a logarithmic average of the maximum shock values in each block.

The vibration data are taken from NREL presented in Section 1.10.1. In this section, we use signals from the test case "CM_2a" as listed in Table 1.1. The date are emanating from the sensor AN3, which is located on input side of the planetary gearbox inside the

wind turbine's mechanical drive train assembly. The signals are adaptively filtered and divided into blocks. Signal in each block is fed to the base of SDOF systems for calculation of SRS and SRS index according to equation (6.7) and Figure 6.2.

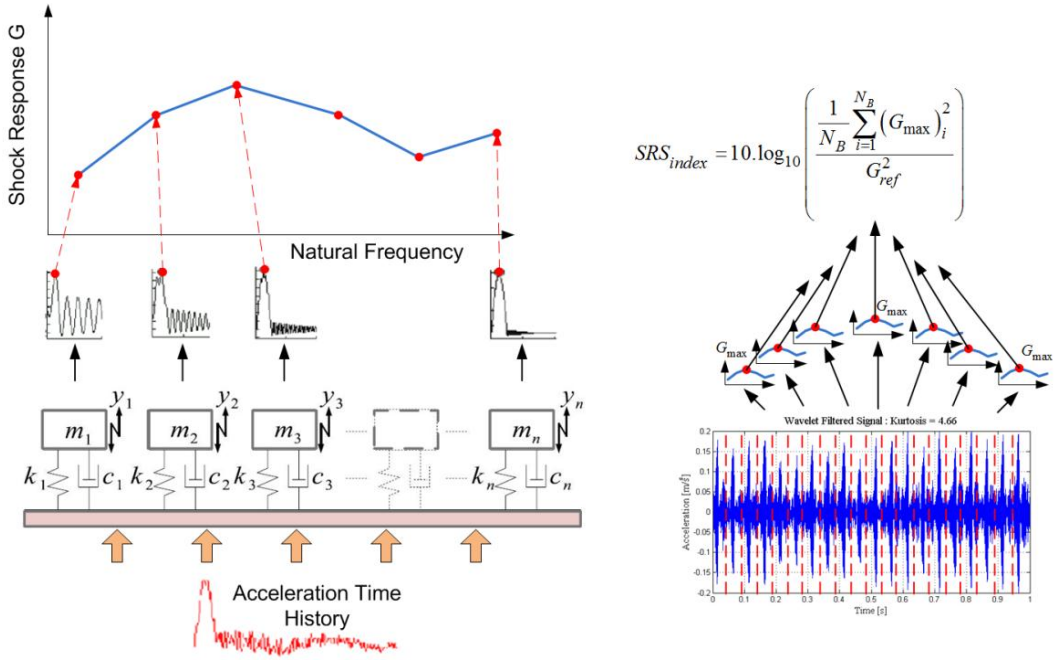


Figure 6.2: SRS SDOF Oscillation System

The value of SRS index can go below zero or negative if the averaged $G_{\max}$ shock energy is less than 1.0 (very less amount of shock), since $\log_{10}$ of a number less than 1.0 is a negative value. In Figure 6.3(b), normalized SRS index's trend is shown. The SRS index is increasing as the fault is getting worse with respect to time.

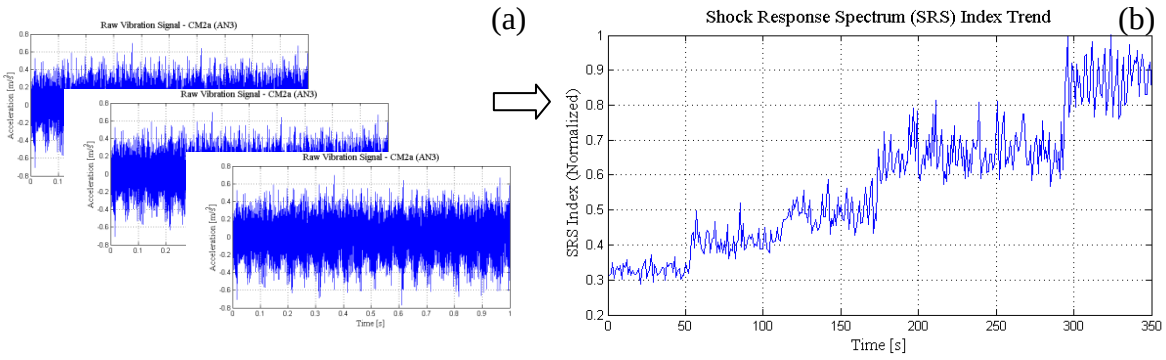


Figure 6.3: SRS Features Extraction

6.2.1.2 Feature-extraction Based on Pulse Shape Analysis

In pulse shape analysis, we analyse rise and fall transition durations, slopes and amplitudes of the faulty pulses present in the signal. Normally, in control theory, rise time is defined as the time for a waveform to go from 10% to 90% of its final value. We use the same reference levels here for the measurement of the rise and fall times of the faulty pulses. Hence, the rise and fall slopes S_r and S_f are calculated as

$$\begin{aligned} S_r &= \frac{High \quad Reference \quad Level - Low \quad Reference \quad Level}{Rise \quad Time} \\ S_f &= \frac{High \quad Reference \quad Level - Low \quad Reference \quad Level}{Fall \quad Time}. \end{aligned} \tag{6.8}$$

The varying fault conditions, severity and development can be analysed from the shape and width of the faulty pulses. The following features are extracted

$$\begin{aligned} F_{SR} &= \frac{1}{n}\sum_{p=1}^{n} S_{rp} \quad where \quad p = 1, 2, ...n \quad No. \ of \quad pulses \\ F_{SF} &= \frac{1}{n}\sum_{p=1}^{n} S_{fp} \quad where \quad p = 1, 2, ...n \quad No. \ of \quad pulses \\ F_{A} &= \frac{1}{n}\sum_{p=1}^{n} A_{p} \quad where \quad p = 1, 2, ...n \quad No. \ of \quad pulses. \end{aligned} \tag{6.9}$$

Where F_{SR}, F_{SF} and F_A are the slope rise, the slope fall and the amplitude features of the faulty pulses present in the signal. Figure 4.6(e) is reproduced here as Figure 6.4(a). After Hilbert transform envelope detection in Figure 6.4(b), pulse shape analysis is performed in Figure 6.4(c).

The Hilbert transform detects the envelope of the band passed signal. Hilbert transform can be expressed as follows

$$\hat{x}(t) = x(t) * \frac{1}{\pi t} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau. \tag{6.10}$$

Hilbert transform creates an artificially complex signal $u(t)$ from $x(t)$. The real part $x(t)$ of the $u(t)$ is the original signal and the imaginary part $\hat{x}(t)$ is the Hilbert transform of the real part. Thus, $u(t)$ is defined as $u(t) = x(t) + j\hat{x}(t)$. The magnitude and phase of $u(t)$ is computed as $A(t) = \sqrt{x(t)^2 + \hat{x}(t)^2}$ and $\theta(t) = \arctan \hat{x}(t) / x(t)$. The magnitude $A(t)$ is the envelop of the signal and is always a positive function. Figure 6.4(d) plots slope versus amplitude features of the faulty pulses for normal and faulty conditions of the gearbox. In Figure 6.4(d), five distinct clusters are seen. The cluster with good gearbox "Spur1-Good" has less amplitude and slope values as compared to other clusters "Spur2-Chipped", "Spur4-Broken" and "Spur5-Chipped&Broken". For all clusters in Figure 6.4(d), the amplitude and slope values are linearly related. When the severity of a fault increases, the transients' amplitudes and slopes also increase. In the "Spur5-Chipped&Broken" case there are two distinct sub-clusters: one belongs to chipped gear and the other belongs to broken gear in the gearbox. Surely, the sub-cluster with high amplitude and slope values belong to the broken gear in the gearbox.

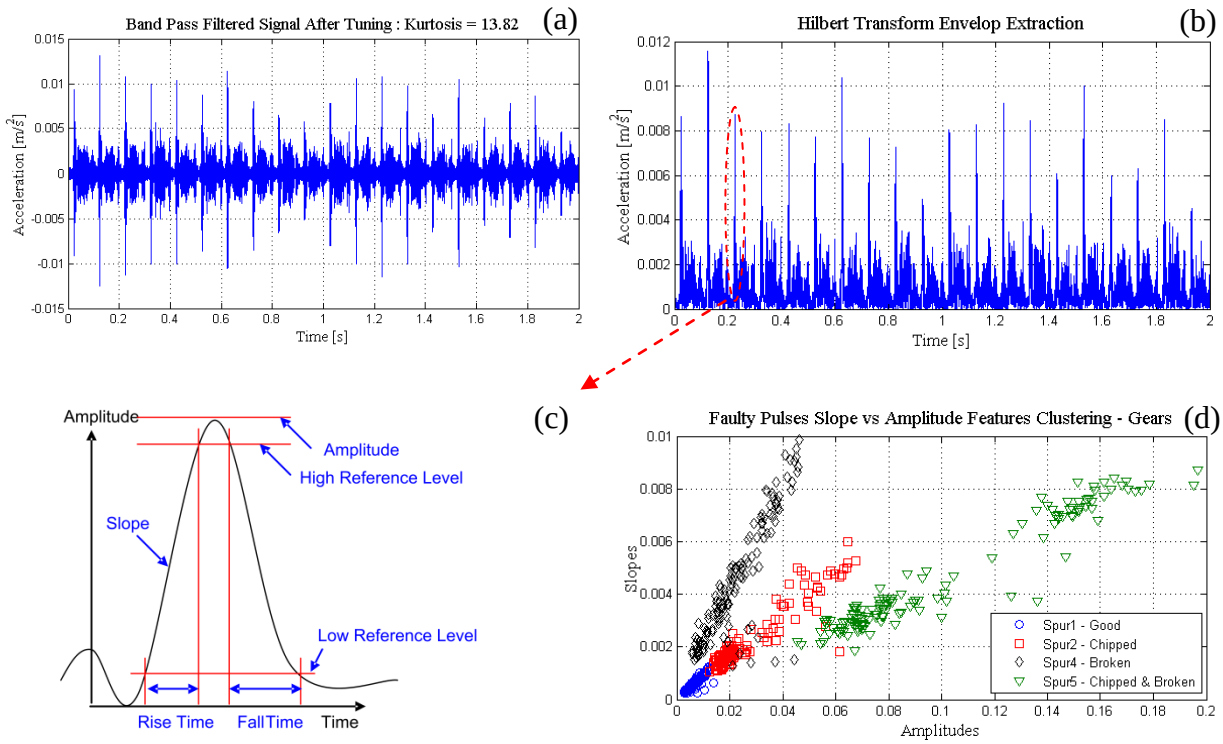


Figure 6.4: Pulse Shape Analysis

6.2.1.3 Feature-extraction Based on Second Order Transient Analysis

Meshing of gears can be modeled by a second order, critically damped, spring mass system [104] as shown in Figure 6.5(a). The spring is deflected by a force $f(t)$ acting on it, normally an impulse. A second order differential equation for the system can be derived from Newton's law as

$$\frac{1}{\omega_n^2}\frac{d^2x(t)}{dt^2}+\frac{2\zeta}{\omega_n}\frac{dx(t)}{dt}+x(t)=K_s f(t). \tag{6.10}$$

Where $\omega_n=\sqrt{k/M}$ is the natural frequency, K_s is the dc gain and $\zeta=b\omega_n/2k=b/2\sqrt{1/kM}$ is the damping ratio. The choice of ω_n, K_s and ζ represents very important characteristics of the response of the second-order system.

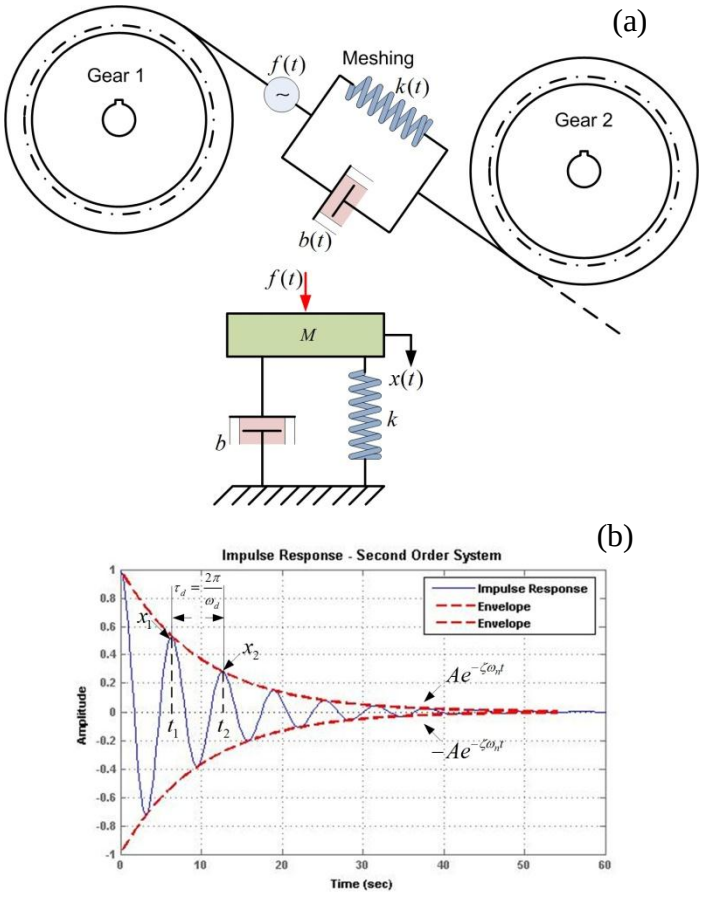


Figure 6.5: Second-Order Spring Mass System and Impulse Response

In Figure 6.5(b), an impulse response of a second-order system is plotted for values of $\zeta = 0.1,\ \omega_n = 1$ and $K_s = 1$. The system response shown in Figure 6.5(b) is for the underdamped case where $\zeta < 1$. When a fault occurs in gears, especially tooth crack, severe forces act on the gears and each time the cracked teeth mesh between the gears, transient forces are generated. These transient forces are responsible for faulty pulses in the vibration signals emanating from the gearboxes. The second-order feature-extraction method estimates the values of natural frequencies ω_n and damping ratios ζ of mass spring systems virtually present between the gears. The driving force $f(t)$, the spring constant $k(t)$ and the damping ratio $\zeta(t)$ change with time and are different for different faults. The envelope of the response is an exponential function of damping ratio ζ and natural frequency ω_n, denoted by $e^{-\zeta\omega_n t}$ or $e^{-t/\tau}$ as shown in Figure 6.5(b). The time constant of the complex envelope $\tau = 1/\zeta\omega_n$ is the time elapsed after the envelope falls down to 37% of its maximum value. The damping ratio ζ is experimentally calculated as from Figure 6.5(b).

$$\frac{x_1}{x_2} = \frac{Ae^{-\zeta\omega_n t_1}}{Ae^{-\zeta\omega_n (t_1+\tau_d)}} = e^{\zeta\omega_n \tau_d}. \tag{6.11}$$

The logarithmic decrement δ and the damping ratio ζ is

$$\begin{aligned} \delta &= \ln\left(\frac{x_1}{x_2}\right) = \zeta\omega_n\tau_d = \zeta\omega_n \frac{2\pi}{\omega_n\sqrt{1-\zeta^2}} = \frac{2\pi\zeta}{\sqrt{1-\zeta^2}}, \\ \zeta &= \frac{\delta}{\sqrt{2\pi^2+\delta^2}}. \end{aligned} \tag{6.12}$$

These values can be further used for fault classification and clustering. In Figure 6.5(b), ω_d is a damped natural frequency given by $\omega_d = \omega_n\sqrt{1-\zeta^2}$.

As described earlier, meshing of gears can be modeled by a second-order, critically damped spring mass systems. The properties of these systems differ for different faults. This phenomenon is used to calculate a set of representative faulty features. We estimate

the natural frequency ω_n and damping ratio ζ of such a system from the faulty pulse itself. The method is described in Figure 6.5 and Equation (6.12). Figure 5.6(d) is reproduced here as Figure 6.6(a). One faulty pulse from Figure 6.6(a) is zoomed in Figure 6.6(b). The amplitudes x_1 and x_2 are calculated as 0.1824 and 0.15011, respectively.

Using Equation (6.12), we calculate $\delta = 0.1952$ and $\zeta = 0.04$. The calculation of τ_d is performed by calculating the difference in two adjacent peaks in number of samples and as shown in Figure 6.6(b). In this particular case, this difference is $2.6814\times10^4 - 2.6740\times10^4 = 74$ samples. Since, the sampling frequency used in this case is $F_s = 40000$ Hz, τ_d becomes $\tau_d = \frac{1}{F_S}\times 74 = \frac{1}{40000}\times 74 = 1.85\times10^{-3}$ sec or 1.85 msec. The damped frequency ω_d is then calculated to be $\omega_d = \frac{2\pi}{\tau_d} = 3396.3 \;\; rad/s$ and natural frequency to be $\omega_n = \frac{\omega d}{\sqrt{1-\zeta^2}} = 3399.6 \;\; rad/s$.

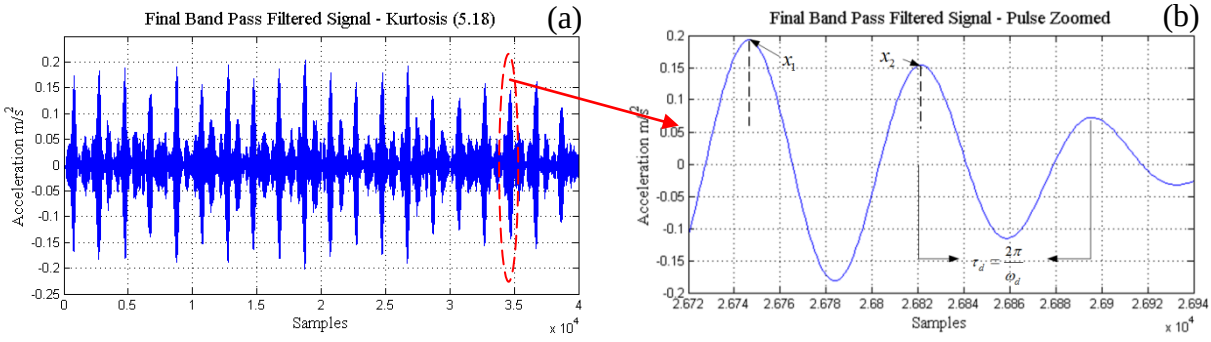


Figure 6.6: Second-Order Transient Analysis Feature-extraction

Figure 6.7 plots natural frequency ω_n versus damping ratio ζ . In Figure 6.7, four distinct clusters are visible, that belong to four levels of fault severity. When the fault's severity increases, the natural frequency and damping ratio also increases. Each cluster in Figure 6.7 represents a different second-order spring-mass system.

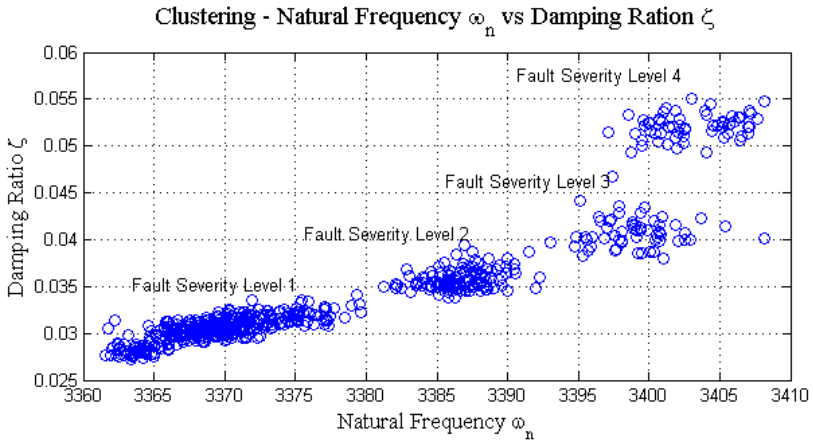


Figure 6.7: Extracted Second Order Transient Features

6.2.1.4 Pulse Position Analysis

Pulse position analysis is performed to locate the faulty pulse positions on the horizontal axis or time axis. Once the positions are located, a pulse repetition frequency can be easily calculated and converted to rotational speed (Revolution per Minute - RPM) of the faulty component. The following equation is used to calculate the average pulse repetition [105]

$$T_P = \frac{1}{F_S}\left[\frac{1}{\left(P-1\right)}\sum_{p=2}^{P}\left(S_p - S_{p-1}\right)\right]. \tag{6.13}$$

Where, S is the peak position (pulse position) on the time axis in terms of samples, p is an index representing peak number and P is the total number of peaks found in the signal. The sampling frequency is F_S. The RPM of the faulty component is then calculated as $RPM = \left(1/T_p\right)\times 60$. After the raw vibrational signal is filtered and the faulty pulses are revealed, Hilbert transform envelop detection is performed and peaks are detected through wavelet peak detection. Let us take the final band-pass filtered signal from Figure 5.6(a) and plot it again in Figure 6.8(a). In Figure 6.8(b), the envelop detected through Hilbert transform and peaks detected through wavelet analysis are shown.

Equation 6.13 is applied to calculate the RPM of the faulty gear as shown in Figure 6.8(c). In this particular case the RPM of the faulty gear calculated through equation 6.13 is 1203.83. This indicates the high speed pinion gear on high speed shaft (HSS) to be the faulty component in the gearbox. We know from the kinematics of gearbox that the RPM of the HSS is 1200 and it confirms the high speed pinion gear as the faulty component. Physical inspection of the gearbox further proves the hypothesis. Figure 5.4 shows a severe scuffing on the HSS pinion gear revolving at the speed of 1200 RPM.

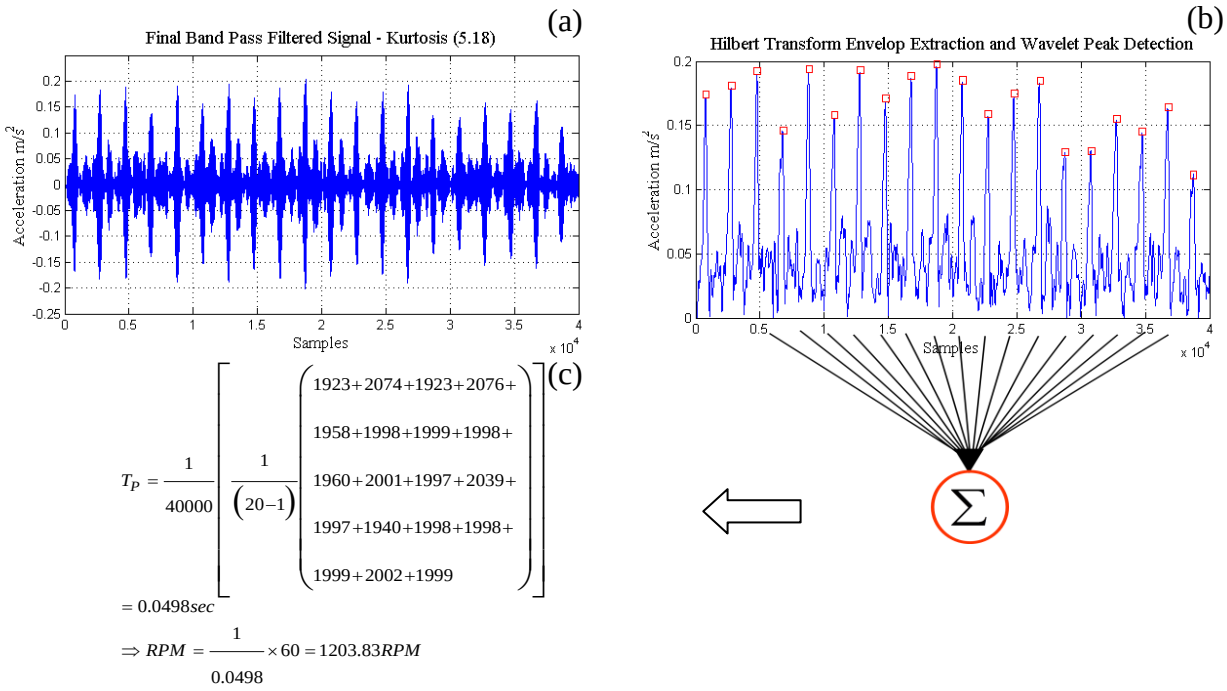


Figure 6.8: Pulse Position Analysis

6.2.2 Frequency Domain Analysis

In this section, some of the problems with gears listed in Table 1.2 of section 1.10.2 are discussed with corresponding fault related peaks in frequency domain.

6.2.2.1 Gears Tooth Wear (Chipped and Broken)

Gear wear causes high amplitude peaks in FFT at gear mesh frequency (GMF) and its side bands. In the "Spur2" case listed in Table 1.2, the 32T gear revolving at 1800 RPM (30 Hz) is chipped. It should cause high amplitude at GMF of $32x30 \cong 960$ Hz. Figure 6.9 shows the FFT analysis with GMF and sidebands. In Figure 6.9, the FFT of the

"Spur1" case with good gearbox and "Spur2" case with 32T chipped gear is plotted side by side for comparison of FFT peak amplitudes. In the "Spur5" case listed in Table 1.2, the 32T gear revolving at 1800 RPM (30 Hz) is chipped and 80T gear revolving at 360 RPM (6 Hz) is broken. It should cause high amplitudes at GMF of $32x30 \cong 960$ Hz and $80x6 \cong 480$ Hz. Figure 6.10 shows the FFT analysis with GMFs and sidebands. In Figure 6.10(a) the FFT of the "Spur1" case with good gearbox and "Spur5" case with 32T chipped gear and 80T broken gear is plotted side by side for comparison of FFT GMF peaks around 480 Hz that belong to 80T broken gear. In Figure 6.10(a) the main GMF peak for the "Spur5" case is less in amplitude as compared to its side bands. This confirms the broken tooth. In Figure 6.10(b) the GMF peaks and side bands around 960 Hz confirm the 32T gear chipped.

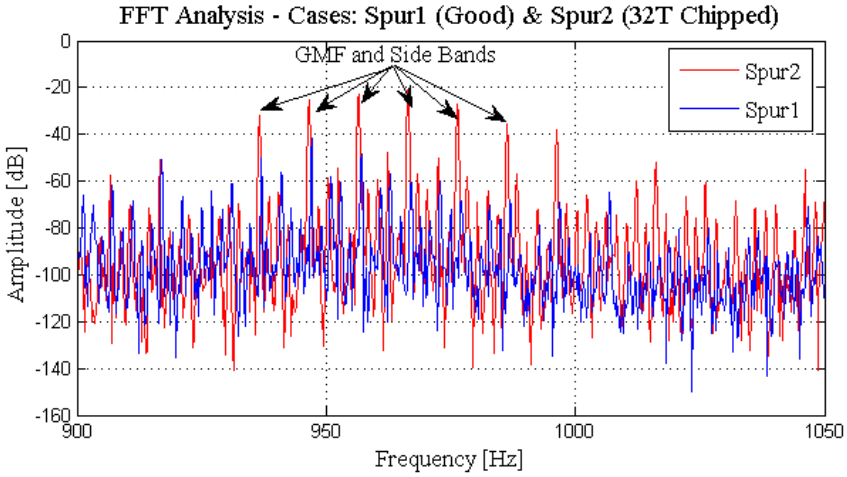


Figure 6.9: FFT Analysis - Spur1 and Spur2 Cases

6.2.2.2 Gear's Eccentricity

The gear's eccentricity is because of bent shaft. The eccentricity can be detected by higher amplitudes at GMFs and its sidebands. Figure 6.11 shows a case of gear's eccentricity. In Figure 6.11, the "Spur1" and "Spur3" cases from Table 1.2 are compared. In the "Spur3" case, 48T gear revolving at 600 RPM (10 Hz) is eccentric. This shows the peaks around GMF to be of $48x10 \cong 480$ Hz.

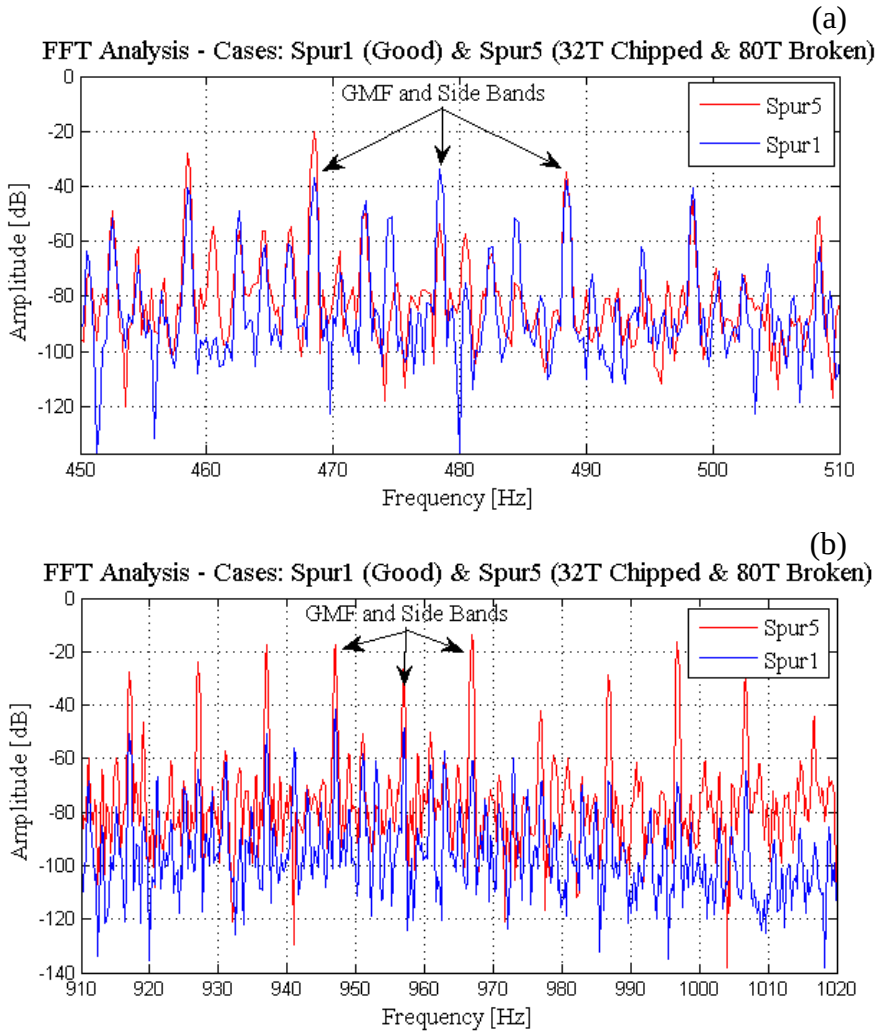


Figure 6.10: FFT Analysis - Spur1 and Spur5 Cases

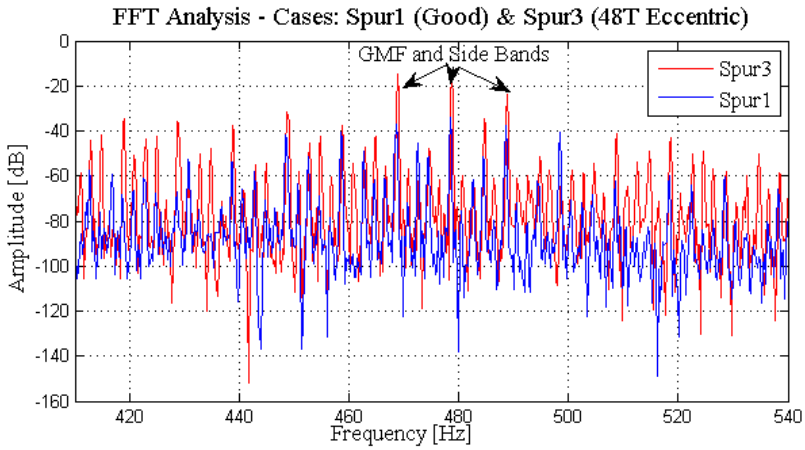


Figure 6.11: FFT Analysis - Spur1 and Spur3 Cases

6.3 Bearing's Features Extraction

Rolling element bearings are most widely used elements in industrial machines. Rolling element bearings provide an interface between moving and stationary parts of a machine. It is very important to detect incipient faults occurring in rolling element bearings in order to avoid catastrophic breakdown and unnecessary maintenance and shutdown costs. In this section, we discuss some conventional concepts about bearing's faults detection and present some features extraction methods representative of bearing's health conditions.

6.3.1 Signal Band-Pass Filtering

A very famous method for bearing fault detection is known as demodulated resonance analysis or envelope power spectral density analysis and is presented in [106]. The method filters a frequency range around a machine structural resonance, which is excited by the periodic impulsive excitation forces generated on bearings with localized defects. Demodulating this frequency range generates a ''signal envelope" which presents more clearly the impulsive nature of the fault than the signal itself. The technique is able to identify and characterize the nature of defects.

6.3.2 Characteristic Defect Frequencies

Defects in bearings exhibit periodic impulses in vibration signals recorded from faulty bearings. The shape and period of these pulses change according to the RPM, the fault type, fault location and bearing geometry. These characteristic defect frequencies are calculated as [107],

$$f_c = \frac{F_s}{2}\left(1-\frac{d}{D}\cos(\alpha)\right). \tag{6.14}$$

$$f_{bd} = \frac{D}{d} f_s \left(1-\frac{d^2}{D^2}\cos^2(\alpha)\right). \tag{6.15}$$

$$f_{id} = n_b\left(f_s - f_c\right) = \frac{n_b f_s}{2}\left(1-\frac{d}{D}\cos(\alpha)\right). \tag{6.16}$$

$$f_{od} = n_b f_c = \frac{n_b f_s}{2d}\left(1 - \frac{d}{D}\cos(\alpha)\right). \tag{6.17}$$

Where f_c is the fundamental cage frequency, f_{bd} is the ball defect frequency, f_{id} is the inner race defect frequency and f_{od} is the outer race defect frequency of a bearing. The shaft rotation frequency is $F_s = \frac{RPM}{60}$ Hz, n_b is the number of roller elements in the bearing, d is the roller diameter, D is the pitch diameter of the bearing and α is the contact angle as shown in Figure 6.12.

We use frequency domain analysis to extract the representative features for detection of bearing's faults. The frequency domain features extracted here are amplitudes of spectrum for bearing's characteristics frequencies calculated in Equations 6.14 – 6.17. Table 6.1 lists the bearing's characteristics frequencies for the vibration data taken from Prognostics and Health Management (PHM) society as listed in Table 1.2 in Section 1.10.2. The bearing's dimensions are as follows:

No. of elements = $n_b = 8$

Roller diameter = $d = 0.3125$ in

Pitch diameter = $D = 1.319$ in

Contact angle = $\alpha = 0$

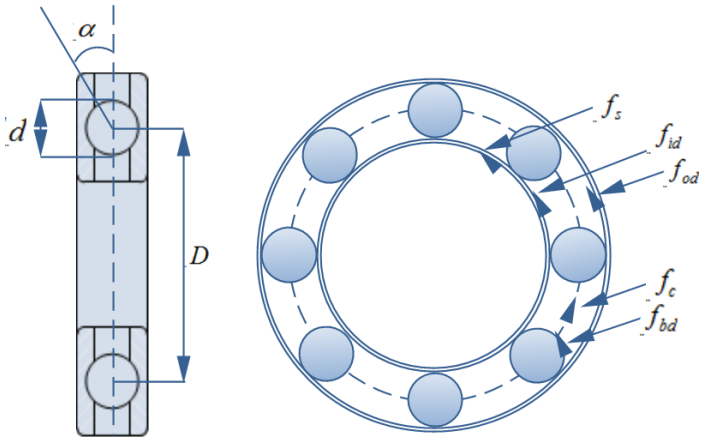


Figure 6.12: Bearing Defect Frequencies and Dimensions

Table 6.1: Characteristic Defect Frequencies – Bearings

Fault Locations	Characteristic Defect Frequencies (Hz)
Outer race	$f_{od} = 45.78$
Inner race	$f_{id} = 91.57$
Cage	$f_{c} = 11.44$
Balls	$f_{bd} = 119.51$

To diagnose the type of the defect, frequency harmonic analyses are usually used. Therefore, a plan for computing the distributed energy in the vicinity of characteristic defect harmonics has been used to extract frequency features for further classification. The signals are filtered through the band-pass adaptive filtering method presented in Chapter 4 and the power spectrum of the filtered signal's squared envelop is calculated. Figure 6.13 shows one case, "Spur4" with bearing's ball defect.

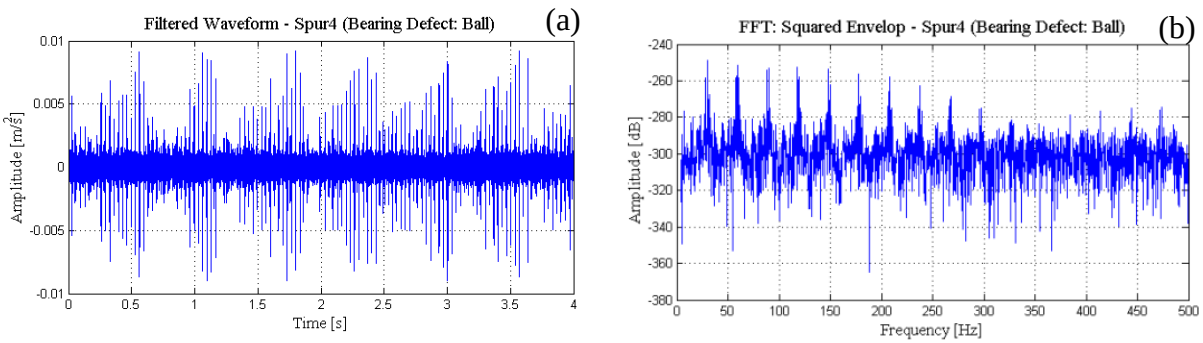


Figure 6.13: Bearing Ball Defect (a) Filtered Time waveform (b) Squared Envelop Spectrum

6.4 Conclusion

In this chapter, time domain transient features extraction algorithms for gears are presented: SRS, pulse shape analysis, and second-order transient analysis. SRS extracts features suitable for trend analysis or prognostics and the other two techniques extract features applicable for clustering. All three features extraction algorithms rely on transient information present in the signals. Vibration signals emanating from faulty machines contain transients of short duration. The shapes of these transients are sensitive

to types of faults present in the machines. Hence, features related with transients are sensitive to the faults present in the machines. This chapter also presents pulse position analysis for faulty gear identification. The pulse position analysis calculates the faulty pulse positions at time axis and estimates the rotation frequency of the faulty gear inside the gearbox. However, the technique can pose difficulties in estimating the faulty component's rotation frequency if there are two or more than two faulty gears inside the gearbox. In later case, we have to rely on FFT or clustering techniques. In this chapter, we also explore frequency domain analysis for detecting different types of faults for gears and bearings. The features extracted in this chapter would be used for prediction, classification and clustering in the next chapter along with some new features extraction techniques.

CHAPTER 7

Prognostic and Clustering of Machines Health Conditions

7.1 Motivation

Prognostic is a very important technique in machine health monitoring and calculation of remaining useful life (RUL) of machines. Prognostics become a difficult task, especially, in the presence of system noise, modeling inconsistencies and degraded sensor fidelity. In prognostics, current and past observations of certain health index are used to predict the future health states of the machine.

Clustering is another technique to group similar objects in one group. Clustering analysis is very often used in machines fault diagnosis. It is required to extract useful health related features from vibration signals emanating from machines before clustering. Chapter 6 is dedicated to different time-domain and frequency-domain feature extraction for gearbox health monitoring. As shown in Figure 1.8, we develop a new technique based on psychoacoustic filtering to extract gearbox health related vibration features. We then use neural networks and neuro-fuzzy approaches for prognostics. In this chapter, we also present fuzzy c-means and linear regression analysis clustering techniques. We use sunspot data-set for measuring the performance of the designed predictors.

7.2 Machines Prognostics

In order to perform prognostics we can either use the SRS features extracted in Figure 6.3(b) or develop a new technique for features extraction. As shown in Figure 1.8, we use psychoacoustic filtering to extract gearbox health related vibration features. Figure 7.1(a)

shows the process of features extraction and prognostics. Below, we discuss the each step one by one.

7.2.1 Health Monitoring Index

The vibration features extraction algorithm presented in this section works on the principle of transient analysis. Transients are very short and abrupt changes in sound waves due to non-linearity. Non-linearity could be mechanically introduced disturbances in electromechanical systems or unwanted clicks in transmission lines. The transient analysis algorithm presented here calculates a real time estimate of transients caused by non-linearity as perceived by the human ear. The Analyzer uses the knowledge about human ear filtering nature as presented in [108] ensuring that the transients are detected in a way that matches the nature of the cochlea and thereby as perceived by the human ear. Transient analysis gives a much better correlation to the perceived quality of sound than traditional measurements based on frequency analysis. The vibration features extraction algorithm presented in this chapter uses principles of auditory models developed by auditory physiologists. They have gathered considerable data which describes the response of mammalian auditory-nerve fibers and it is clear from the investigations that some of the frequency analysis is performed by human ear [109]. Figure 7.1(a) shows the block diagram of the transient analysis method. Pulses with short rise time or fall time will contain a broad spectrum of frequencies. Therefore it is possible to detect the instantaneous energy in frequency bands in the transient range of the ear. A common method for doing this is to use a filter bank containing a group of band-pass filters covering the frequency interval of interest. The purpose of the band pass filters is to detect the pulses in the frequency band where the pulses have most energy as perceived by the human ear. It will be the filter where the shape of the impulse response matches best the shape of the pulses. The optimal match will be an impulse response with shape equal to the pulse but reverse in time. A gamma tone filter bank is used in this study. Figure 7.1(b) shows a gamma tone filter bank. First designed by Patterson and Holdworth [110], gamma tone filter bank is an array of band pass filters which simulates the response of human ear cochlea. At each point along the cochlea, a psychoacoustic

measure of the width of the auditory filter is represented by an equivalent rectangular bandwidth (ERB).

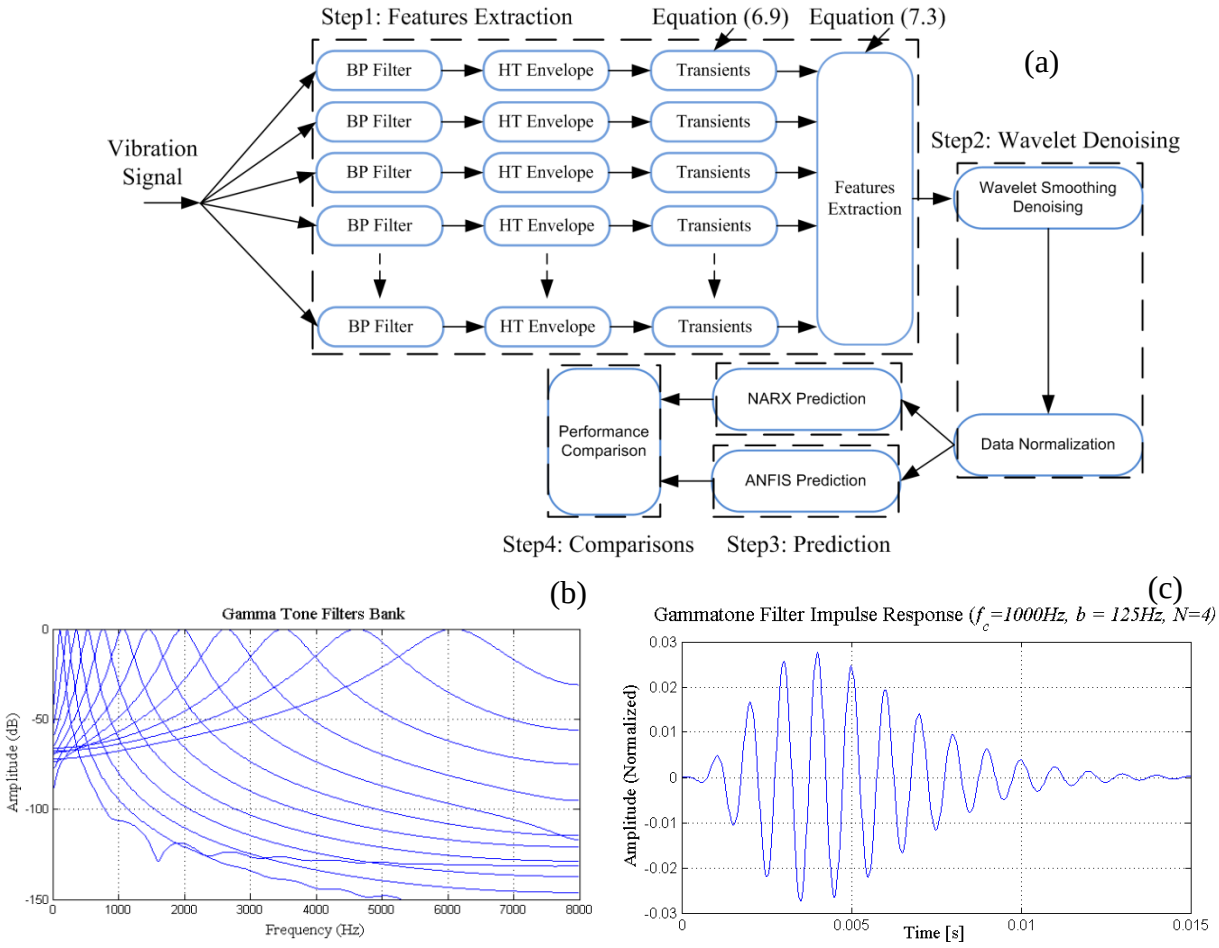


Figure 7.1: (a) Vibration Based Features Extraction (b) Gammatone Filters Bank (c) Gammatone Filters Impulse Response

The bandwidth of the filters is set by a critical band function and so filters bandwidth increases with increase in its center frequency. The relationship between ERB and centre frequency F_c Hz is given by the following equation

$$ERB = 24.7 + 0.108F_c. \tag{7.1}$$

The impulse response of a band pass filter is defined by the following relation [110]

$$h(t) = Rt^{N-1}e^{-2\pi mt}\cos(2\pi F_c t + \phi). \tag{7.2}$$

Where R is an arbitrary factor that is typically used to normalize the peak magnitude transfer to unity, N is the filter order, m is a parameter that determines the duration of the impulse response and thus the filter's bandwidth, F_c is the filter's center frequency, and ϕ is the phase of the tone. Figure 7.1(c) shows an impulse response with $F_c = 1000Hz$, $m = 125Hz$ and $N = 4$. To detect the energy in the channels the output signal from the band-pass filter is Hilbert transformed. The Hilbert transform detects the envelope of the band-pass filtered signal and extracts the instantaneous energy of the faulty pulses.

After the pulses are extracted through envelope detection, feature extraction block follows. In feature extraction block, amplitude and steepness of the pulses are calculated as per Equation (6.9). Amplitude (maximum magnitude) is a linear detector for pulses in the full frequency area. Steepness (maximum slope) is very sensitive for catching nonlinear sounds in the high frequency area. Both metrics can be summarized to a logarithmic index expressing the amount and size of the pulses. The index may be based on either the magnitude or the steepness of the pulses. It is expected that pulses with a short rise time will be more annoying to the ear, than pulses with longer rise times. Thus, by considering the pulse envelope, it seems reasonable to focus the measure on the steepness in leading edge. Every transient detected is characterized by the 80% amplitude and the maximum steepness in the leading edge. In order to obtain an equal number of detected transients in each frequency bands, the output from the transient analyzer is split into blocks of a pre-specified duration. The duration is found as a trade-off between preserving the complexity of the measurement and having an equal number of detected transients in each channel. In each block, only the transient with the maximum steepness is gathered. The maximum steepness is found by differentiating the signal and finding maximum amplitude of the differentiated signal. The logarithmic index is calculated as follows

$$Vb_S = 10.\log_{10}\left(\frac{\frac{1}{N_B N_C}\sum_{i,channels}\sum_{j=1}^{N_B}\left(S_{max}\right)_{ij}^2}{S_{ref}^2}\right). \qquad (7.3)$$

Where Vb_S is vibration steepness index, i is index for band pass filter and N_C is total number of band pass filters used. Index of sub-blocks in a band pass filtered signal is j and N_B is the total number of sub-blocks. The maximum steepness in band pass filter i and sub-block j is $\left(S_{max}\right)_{ij}$. The argument for squaring the maximum steepness is simply to put the large steepness values in favour. $S_{ref}=1$ if the measured amplitude is acceleration. For amplitude index Vb_A, $\left(S_{max}\right)_{ij}$ in Equation (7.3) is replaced by $\left(A_{max}\right)_{ij}$ where A_{max} is the maximum amplitude of the non-differentiated envelope signal. The vibration features extracted are de-noised through wavelet de-noising techniques and normalized before prediction.

7.2.2 Wavelet Denoising

Wavelets are limited duration, undulatory mathematical functions. The time integral of wavelet functions equals to zero. Figure 7.2 plots some common wavelets. Similar to Fourier transform, where we use sine and cosine as basis functions, wavelet transform uses wavelets as basis functions. Wavelets are used in many different fields including compression, signal processing, and de-noising [111]. In Fourier analysis, we approximate a function $f\left(x\right)$ by sines and cosines functions with different frequencies and amplitudes. Thus, the approximation equation becomes

$$\hat{f}\left(x\right)=a_o+\sum_{k=1}^{\infty}\left(a_i\sin\left(kx\right)+b_i\cos\left(kx\right)\right). \tag{7.4}$$

Where, a_o, a_i and b_i are calculated from Fourier transform as

$$F\left(\omega\right)=\int_{-\infty}^{\infty}f\left(t\right)e^{-j\omega t}dt. \tag{7.5}$$

Wavelet analysis is performed in the similar way as Fourier transform but with scaled and translated versions of mother wavelet $\psi\left(x\right)$ as basis functions. Mother wavelet can be any one from Figure 7.2 above and the scaling and translation is defined as child wavelets and can be calculated as

$$\psi_{j,k}(x) = \kappa\psi\left(2^j x - k\right). \tag{7.6}$$

Where κ is a constant, k is wavelet translation and 2^j is scale translation. We can estimate $f(x)$ from the following equation in wavelet analysis

$$\hat{f}(x) = \sum_{\forall j,\forall k} c_{j,k}\psi_{j,k}(x). \tag{7.7}$$

Where, $c_{j,k}$ are the wavelet coefficients and are obtained through the wavelet transform as

$$c_{j,k} = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)\,dx. \tag{7.8}$$

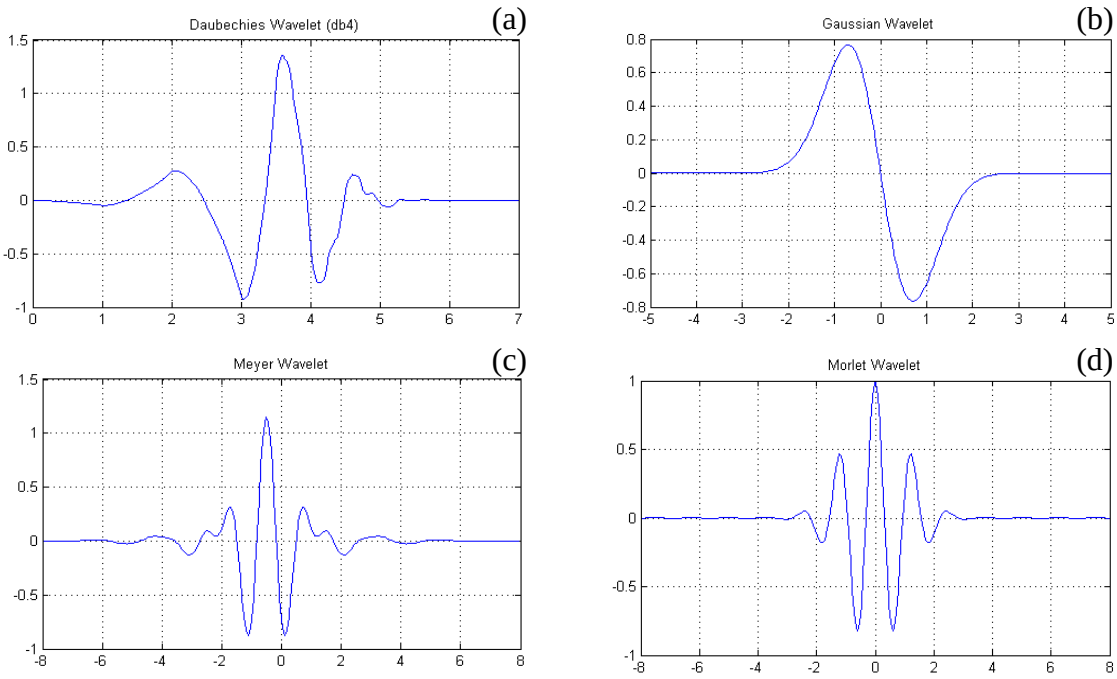


Figure 7.2: Commonly used Wavelets: (a) Daubechies, (b) Gaussian, (c) Meyer and (d) Morlet.

Wavelet analysis basis functions are finite and limited to one size and this makes wavelet analysis useful technique for detecting local features like discontinuities and spikes in a signal. On the other hand, Fourier analysis basis functions are infinite in nature and an approximation of a specific part of the signal affects the entire signal. Wavelet analysis is joint time-frequency analysis technique as contrast to the Fourier transform that is purely

frequency analysis. This feature also makes wavelet analysis to tell about when in time a particular event took place. Wavelet smoothing or wavelet trend analysis is used to remove high frequency components from the monitoring index which can be assumed as noise. The process is also called wavelet de-noising. De-noising the signal is one of the most effective applications of wavelets in signal processing. The wavelet transform-based de-noising methods can produce much higher de-noising quality than conventional methods. Furthermore, the wavelet transform-based methods retain the details of a signal after de-noising [111, 112]. We approximate the original signal from Equation (7.7). Each coefficient $c_{j,k}$ obtained in Equation (7.8) by wavelet transform is a contribution of the wavelet $\psi_{j,k}(x)$ in the whole approximation for the original signal. If the value of this coefficient $c_{j,k}$ is very small and its contribution to the approximation is considered negligible, we can omit the corresponding child wavelet $\psi_{j,k}(x)$ from the approximation. This procedure is called *thresholding* and it forms the basis for the wavelet de-noising. After de-noising the prediction trends by wavelet analysis, techniques for time series prediction comes in.

7.2.3 Time Series Prediction

The prediction of time series $x(t)$ at r time steps ahead, x_{t+r}, is obtained based on its values at present and past time steps $[x_{t-mr}, x_{t-(m-1)r}, x_{t-(m-2)r}, ..., x_{t-2r}, x_{t-r}, x_t]$ as $x_{t+r} = \chi(x_{t-mr}, x_{t-(m-1)r}, x_{t-(m-2)r}, ..., x_{t-2r}, x_{t-r}, x_t)$. Where χ is a predictor functions and can be approximated through various conventional, statistical and artificially intelligent techniques like Bayesian, support vector regression, Adaptive Neuro-Fuzzy Inference System **(**ANFIS) and neural networks (NN). This paper uses a dynamic neural network, the nonlinear autoregressive model with exogenous inputs (NARX) and ANFIS techniques to approximate the predictor function χ.

7.2.3.1 The NARX

The NARX is a dynamic neural network, used for modeling nonlinear dynamical systems. The NARX can be represented mathematically as

$$y(n+1) = f\left[y(n),...,y(n-d_y+1);u(n),u(n-1),...,u(n-d_u+1)\right]. \qquad (7.9)$$

Where, $\mathbf{u}(n)$ and $\mathbf{y}(n)$ are the input and output of the system at time step n, while $d_u \geq 1$ and $d_y \geq 1$, $d_u \leq d_y$, are the input-memory and output-memory orders. Equation (7.9) can also be written in compact form as $y(n+1) = f\left[\mathbf{y}(n);\mathbf{u}(n)\right]$, where $\mathbf{u}(n)$ and $\mathbf{y}(n)$ are the input and output regressor vectors respectively. The nonlinear mapping function $f(.)$ is approximated through a multi-layer perception (MLP) algorithm trained with plain back propagation algorithm. This research deals with nonlinear univariate time series prediction and for this we set $d_y = 0$. This reduces the NARX network to time delay neural network (TDNN) architecture and Equation (7.9) reduces to [113]

$$y(n+1) = f\left[u(n),u(n-1),...,u(n-d_u+1)\right]. \qquad (7.10)$$

Figure 7.3 shows the way the NARX is trained and tested. During the training phase, the feedback loops (dotted lines in Figure 7.3) are not used.

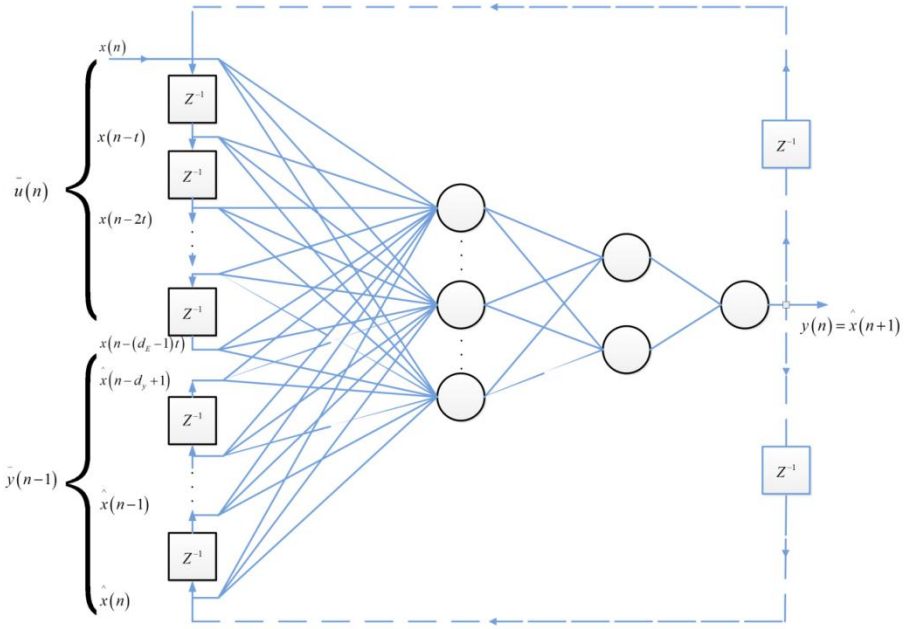


Figure 7.3: NARX for Training and Testing
(Feedback loops are required only during testing)

During the testing or prediction phase, if multistep-ahead predictions are required, the output values are fed back to both the input regressor $\mathbf{u}(n)$ and the output regressor $\mathbf{y}(n)$ at the same time. Thus, the resulting predictive model contains two feedback loops, one for the input regressor and another for the output regressor. Several types of activation functions can be used in the NARX neurons as shown in Figure 7.4 [126].

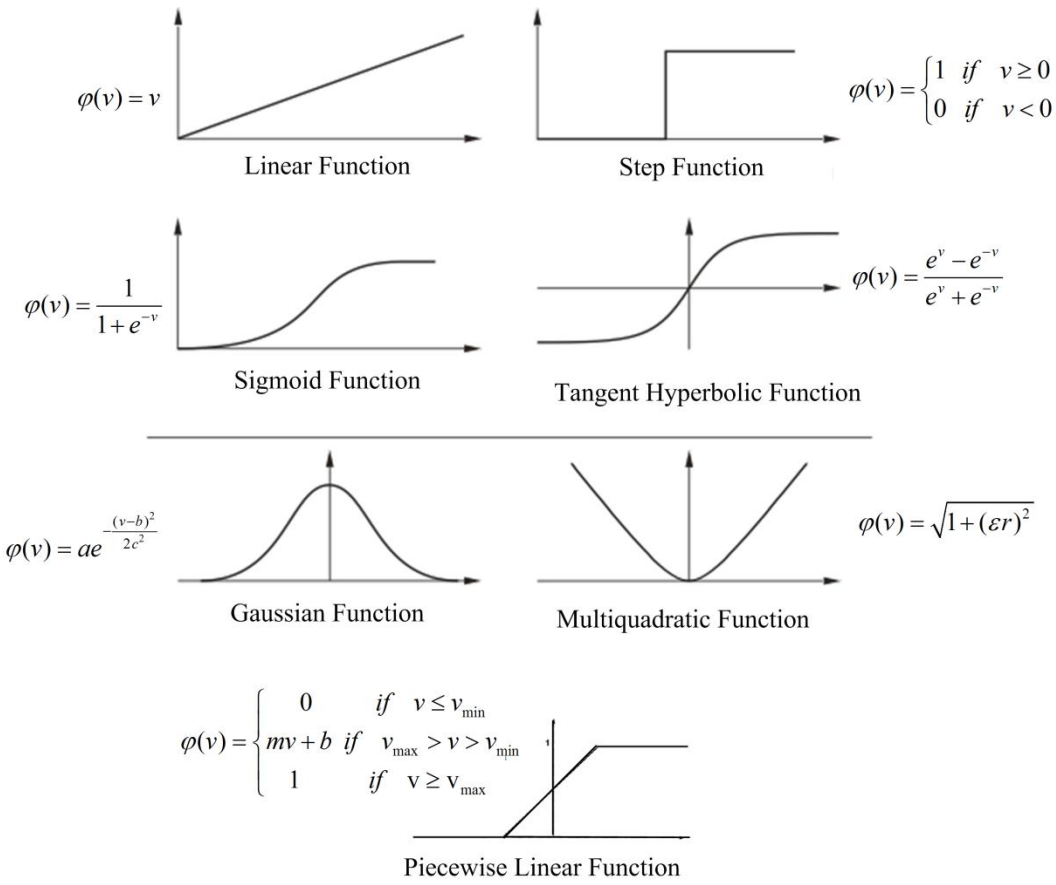


Figure 7.4: Different Types of Activation Functions used in NARX

We use the following settings for the NARX network:

```
Input delays: 2
Feedback delays: 2
Hidden layers: 1
Hidden layer size: 10
Training data: 70%
Validation data: 15%
Testing Data: 15%
Performance criteria: Mean squared error (MSE)
```

```
Training algorithm: Levenberg-Marquardt
Hidden layer activation function: Sigmoid
Output layer activation function: Linear
```

7.2.3.2 The Adaptive Neuro-Fuzzy Inference System (ANFIS)

The basic structure of the ANFIS is shown in Figure 7.5. The ANFIS has m inputs $(x_1, x_2 ..., x_m)$, each with n membership functions (MFs), R rules and one output y. When the ANFIS predicts a time series, the inputs are $\left(x_{t-mr}, x_{t-(m-1)r}, x_{t-(m-2)r}, ..., x_{t-2r}, x_{t-r}, x_t\right)$ and the output of the ANFIS is $y = x_{t+r}$. In the above mentioned case, the ANFIS predicts the time series r time steps ahead based on current and the previous m values. We use a Sugeno-fuzzy type inference system with five layers and $m = 4$ inputs. Number of nodes N in layer 1 is the product of number of inputs m and the input MFs n for each input, i.e., $N = m.n$. Number of nodes in layers 2-4 is equal to the number of rules R in the fuzzy rule base. Layer 1 is a fuzzufication layer and it transforms the crisp inputs x_i to linguistic labels A_{ij}. The examples of the linguistic labels are small, medium, large etc., and the transformation occurs with some degree of the MFs as $O_{ij}^1 = \mu_{ij}(x)$. Where, $i = 1, ..., m$, $j = 1, ..., n$ and μ_{ij} represents the jth membership function for the input x_i. Different types of MFs are used like triangular, trapeziodal, Gaussian etc. Layer 2 of the ANFIS is a product layer, where for each node k, the output represents weighting factor or firing strength of the rule R associated with k. The output w_k of this layer is $O_k^2 = \prod \mu_{ik}(x_i)$, and it is the product of all its inputs scaled according to the MFs μ_{ik}. Where $i = 1, ..., m$ and $k = 1, ..., R$. Layer 3 is a normalization layer and the output of each node k in this layer represents the normalized weighting factor $\bar{w}_k$ of the kth rule as $O_k^3 = \frac{w_k}{\sum_k w_k}$. Where $k = 1, ..., R$. Layer 4 is a de-fuzzification layer and the output of each node in this layer is a weighted output of the first order Sugeno-type fuzzy *if-then* rule as $O_k^4 = \bar{w}_k f_k$. Where, $f_k = \sum_j p_{kj} x_j + r_k$, $j = 1, ..., n$, $k = 1, ..., R$, f_k is the output of the kth rule, and the

parameters p_{kj} and r_k are called consequent parameters. Layer 5 is the final output layer and it contains only one node inside. The output of the layer 5 is an overall output y of the network as $O^5 = \sum_k \bar{w}_k f_k$. It is also a sum of all the weighted outputs of the rules.

We need a training dataset of desired input/output pairs $(x_1, x_2, ..., x_m, y)$ to train the ANFIS or model the target system. In training phase, the ANFIS adaptively maps the input features space $(x_1, x_2, ..., x_m)$ to the corresponding output y. The mapping in the ANFIS system is done through the membership functions (MFs), the rule base and the related parameters that emulate the training dataset. The training phase of the ANFIS uses hybrid learning method. It uses the gradient descent approach for fine tuning the parameters that define the MFs and applies the least squares method to identify the consequent parameters that define the coefficient of each output equation in Sugeno-type fuzzy rule base. The training process continues till the desired stopping criteria is reached, i.e., number of epochs or error tolerance.

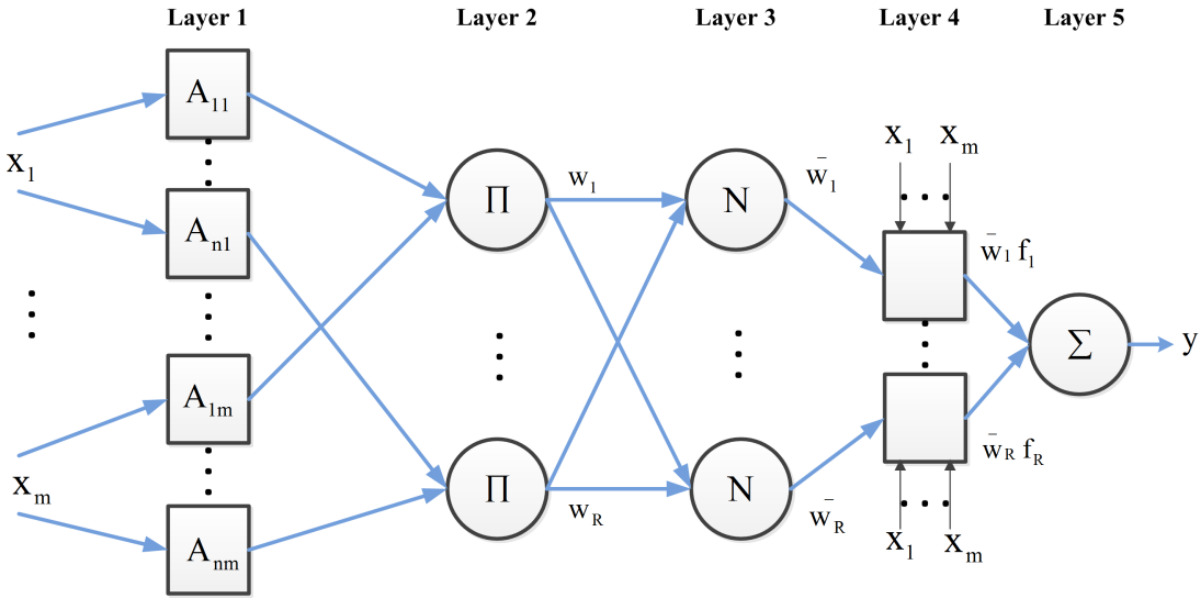


Figure 7.5: Basic Structure of ANFIS

7.2.3.3 Simulations and Discussions

In this section, the prediction accuracy of both the time series predictors, NARX and ANFIS is compared using standard dataset of sunspot activity for years 1749-2012 [114]. The sunspot activity data has non-linear, non-Gaussian and non-stationary characteristics and is suitable to test the performance of the predictors. Figure 7.6(a) shows the sunspot activity data along with wavelet smoothing. NARX and ANFIS prediction results are shown in Figure 7.6(b) and 7.6(c) along with errors box-plots in Figure 7.6(d). In Figure 7.6(d), the medians of the box-plots are centred at zero for both NARX and ANFIS predictors and the $\pm 2.698\sigma$ lines are at ± 0.01 showing about 99.3% of the error observations within the range of ± 0.01. Where σ is the error standard deviation. There are some outliers for both NARX and ANFIS cases but NARX depicts better performance in this case. The mean absolute error (MAE) and mean square error (MSE) values are less in the case of NARX (sunspot data-set) in Table 7.1.

In the framework of machine conditions prognostics, we calculate the machine's health index proposed in section 7.2.1. The experimental vibration data we use in this section emanate from a planetary gearbox inside a windmill. The data are provided by the National Renewable Energy Laboratory (NREL), through a consortium called the Gearbox Reliability Collaborative (GRC). The details about the data collected are shown in section 1.10.1 in Chapter 1. Figure 7.7(a) shows an example of a raw vibration signal, 1 second (40000 samples) in length, collected from the gearbox inside the windmill. A total of 350 such vibration signals are analysed in this section.

Vibration indices are calculated as proposed in section 7.2.1. Figure 7.7(b) shows a vibration index trend with wavelet smoothing. Figure 7.7(c) shows one step prediction for NARX and Figure 7.7(d) shows one step prediction for ANFIS. NARX seems to exhibit more promising results as compared to ANFIS in this case. Also, the MAE and MSE error values for NARX are less as shown in Table I (vibration index data-set). We can also use one step recursive prediction for NARX and ANFIS to predict as many future values as we want as shown in Figures 7.7(e) and 7.7(f). For multi-step recursive

prediction, we have to loop to output values back as discussed in section 7.4.1 in NARX case (dotted lines in Figure 7.3). Similar strategy is adopted in the ANFIS case.

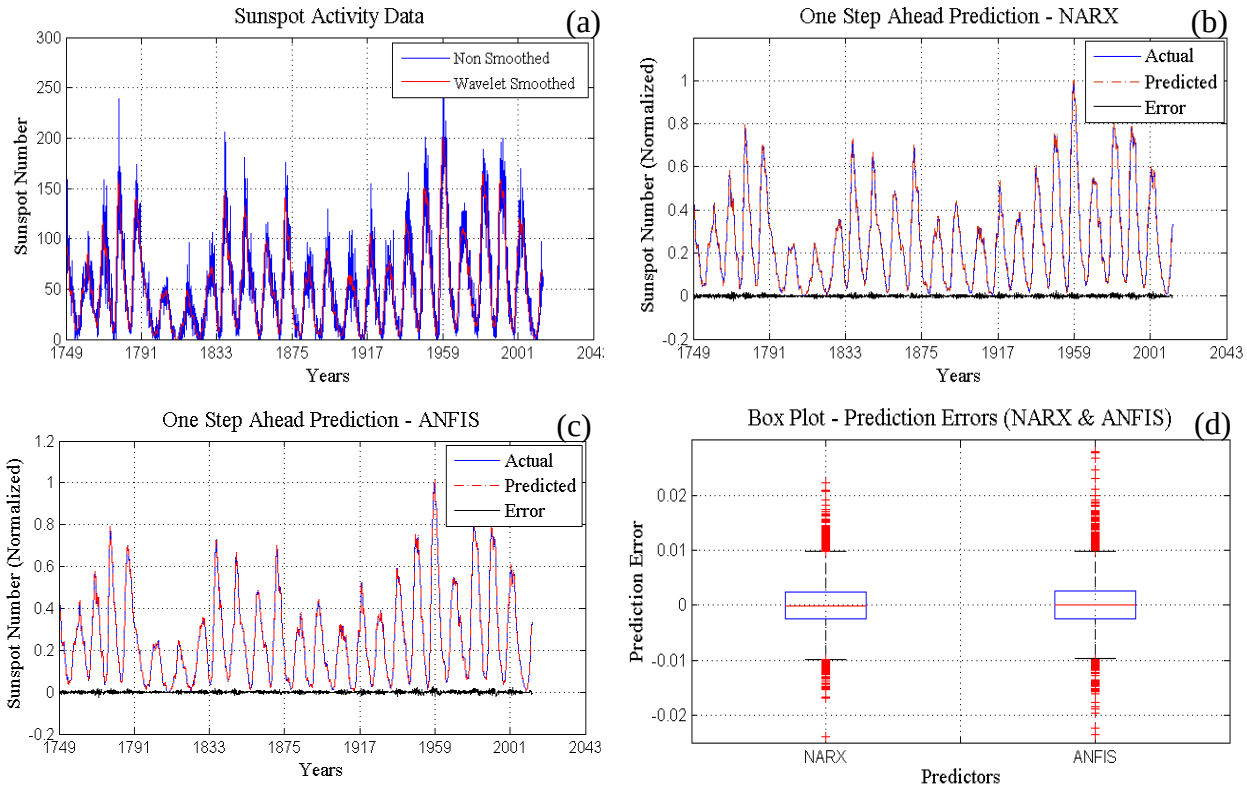


Figure 7.6: Sunspot Activity Data with Wavelet Smoothing and Prediction

7.3 Clustering

7.3.1 Fuzzy Logic and Fuzzy C-Means Clustering

Fuzzy c-means clustering was first presented in 1981 [116]. It is an extension of hard k-means clustering with the advantage of the fuzzy set theory. In fuzzy-set theory, the crisp or binary membership of a set is changed with gradual membership on the real and continuous interval $[0\ \ 1]$. Therefore, in fuzzy clustering a point belongs to a set by a partial membership function. Figure 7.8 shows the basic difference between crisp sets and fuzzy sets.

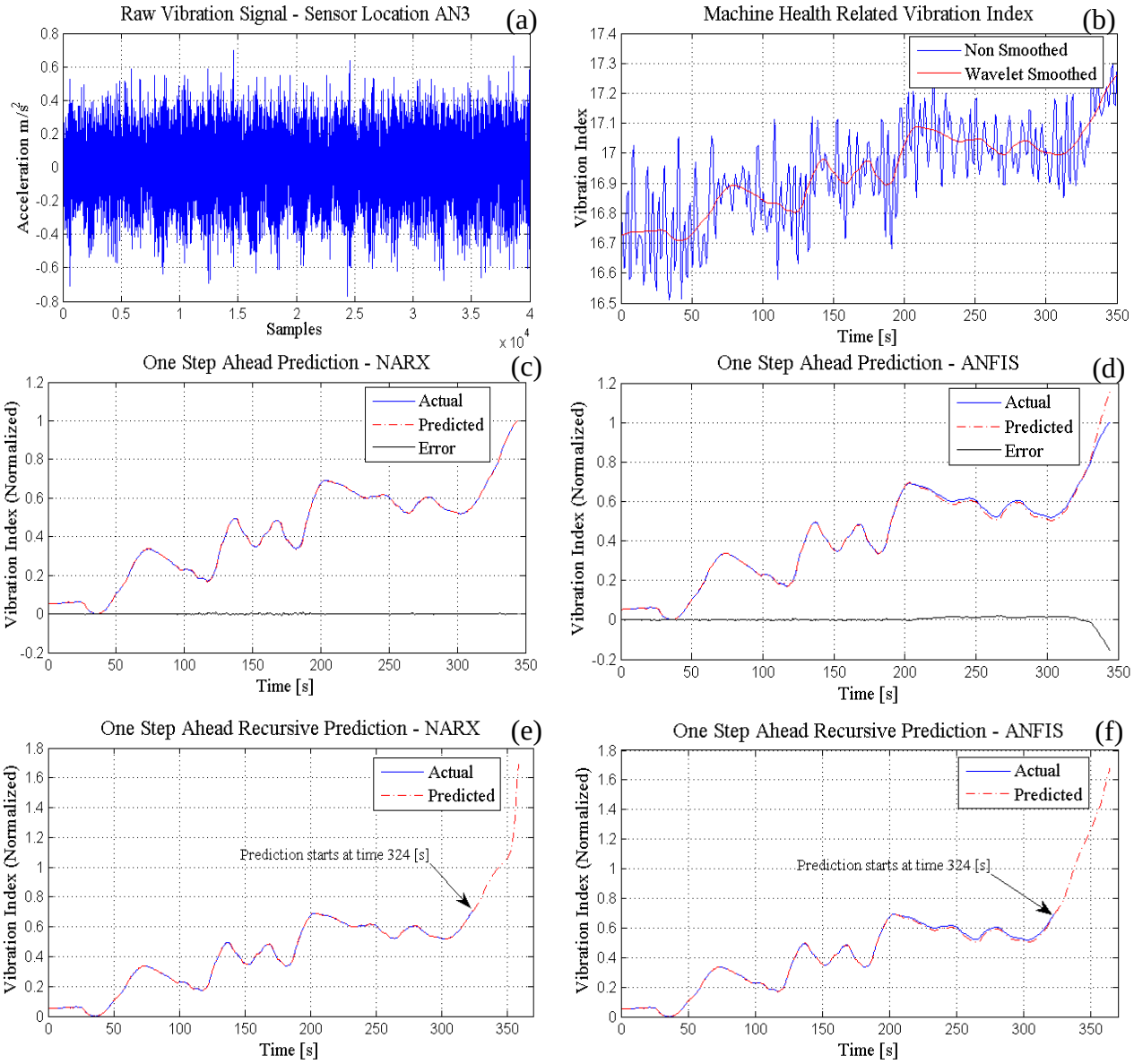


Figure 7.7: Vibration Data with Extracted Features, Wavelet Smoothing and Prediction

Table 7.1: MAE and MSE – NARX & ANFIS

Data Set	Model	MAE	MSE
Sunspot	NARX	0.0034	2.144×10^{-5}
	ANFIS	0.0035	2.421×10^{-5}
Vibration Index	NARX	0.0013	3.626×10^{-6}
	ANFIS	0.0086	4.079×10^{-4}

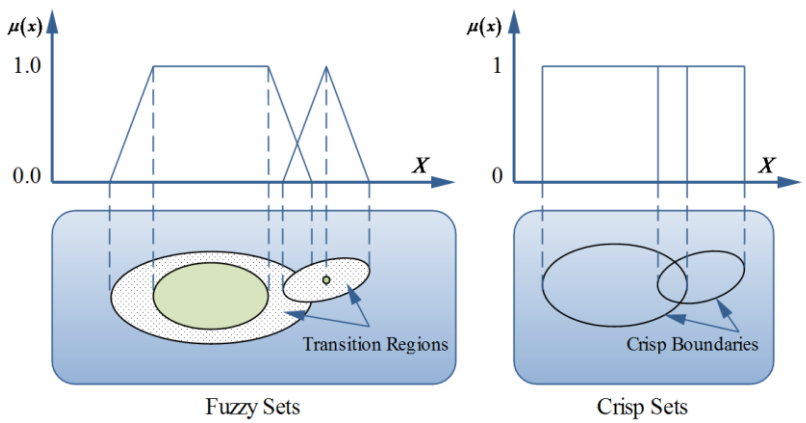


Figure 7.8: Difference between Fuzzy Sets and Crisp Sets

The main aim of the algorithm is to cluster those data points in a cluster that minimize a dissimilarity function of the form [117].

$$J_m\left(\mathbf{U}^*,\mathbf{v}\right)=\min\left[J_m\left(\mathbf{U},\mathbf{v}\right)\right]=\min\left[\sum_{k=1}^{n}\sum_{i=1}^{c}\left(\mu_{ik}\right)^{m'}\left(d_{ik}\right)^2\right]. \tag{7.11}$$

Where, μ_{ik} is the membership values of kth data point in the ith cluster space, m' is a weighting parameter in the range $\left[1\ \ \infty\right]$, $\mathbf{U}$ is fuzzy partition matrix, $\mathbf{v}$ is cluster centre matrix and $\mathbf{d}$ is similarity matrix describes as

$$d_{ik}=d\left(x_k-v_i\right)=\left\|x_k-v_i\right\|=\sqrt{\sum_{j=1}^{m}\left(x_{kj}-v_{ij}\right)^2}\quad\forall\quad i=1:c\quad\text{and}\quad k=1:n. \tag{7.12}$$

Where, m is the number of features, x_k is kth data point and v_i is the centroid of ith cluster and can be written as $v_i=\left\{v_{i1},v_{i2},...,v_{im}\right\}\ \ \forall i=1:c$. Cluster centres are calculated from the following formula [95]:

$$v_{ij}=\frac{\sum_{k=1}^{n}\mu_{ik}^{m'}x_{kj}}{\sum_{k=1}^{n}\mu_{ik}^{m'}}\quad\forall\quad i=1:c\quad\text{and}\quad j=1:m. \tag{7.13}$$

Where, x is a fuzzy variable representing the data point. The fuzzy partitioning is performed iteratively with the following criteria.

$$u_{ik}\left(s+1\right)=\frac{1}{\left[\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m'-1}}\right]}. \tag{7.14}$$

The sum of the membership values should be equal to 1 as

$$\sum_{i\in I_k} \mu_{ik} = 1, \tag{7.15}$$

and the stopping criteria is defines as $\left\| U^{(r+1)} - U^{(r)} \right\| \leq \varepsilon$. Where, ε varies in the range $[0\ \ 1]$. The major objective is to minimize the cost function J_m. The flowchart for the fuzzy c-means clustering is shown in Figure 7.9.

7.3.1.1 Simulations

The data we use in this section is from Figure 6.7, the second order transient features with natural frequency ω_n and damping ratio ζ. Figure 7.10 shows the clustered data with four clusters in 7.10(a) and five clusters in 7.10(b) along with corresponding cost functions' trend in 7.10(c). The following parameters are used:

No. of clusters = 4, Exponent for $\mathbf{U}$ = 2 and Error tolerance $\varepsilon = 1e-6$.

No. of clusters = 5, Exponent for $\mathbf{U}$ = 2 and Error tolerance $\varepsilon = 1e-6$.

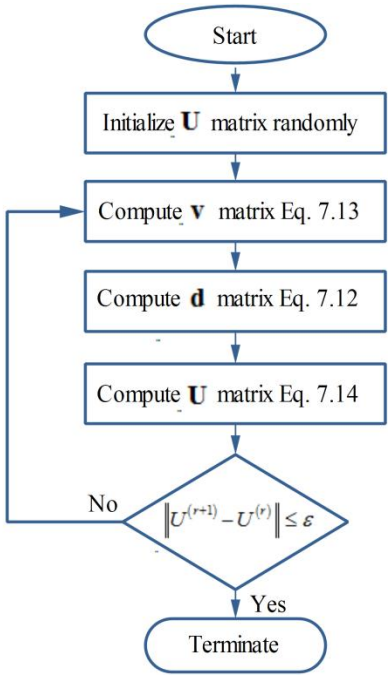


Figure 7.9: Fuzzy c-means Algorithm Flow Chart

7.3.2 Linear Least-Squares Regression Clustering

Linear least-squares regression fits a straight line $y = a_o + a_1 x + e$ through a set of paired observation $\{(x_1, y_1), (x_2, y_2), ...(x_n, y_n)\}$. Where a_o and a_1 are the intercept and slope

of the straight line and e is the error or residual between the model and the observations. The error e can be written as $e = y - a_o - a_1 x$. In linear least-square regression fitting, the aim is to minimize the sum of squares of the residual as

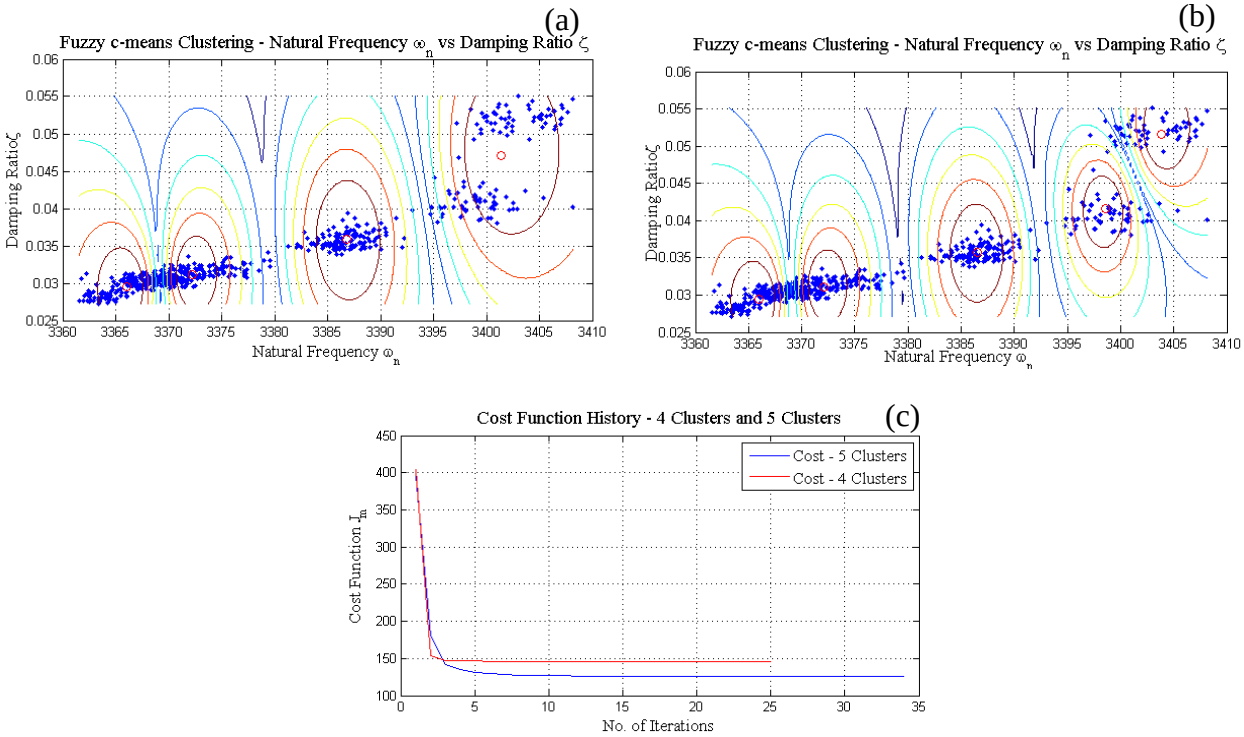


Figure 7.10: Fuzzy c-means Clustering

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y - a_o - a_1 x\right)^2. \tag{7.16}$$

The values of the coefficients a_o and a_1 can be found by [97]

$$a_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \tag{7.17}$$

and

$$a_o = \bar{y} - a_1 \bar{x}. \tag{7.18}$$

Where, $\bar{y}$ and $\bar{x}$ are the mean values of y and x respectively. The standard deviation for the regression line can be determined by

$$s_{y/x} = \sqrt{\frac{S_r}{n-2}}. \tag{7.19}$$

Where, $s_{y/x}$ is called the standard error of estimates for predicted values of y corresponding to particular values of x. The standard error of estimates quantifies the spread of the data around the regression line, similar to standard deviation that quantifies the spread of data around the mean value. If we calculate sum of the squares around the mean for dependent variable y, the quantity is quantified as $S_t = \sum_{i=1}^{n}\left(y - \bar{y}\right)^2$ and the coefficient of determination r^2 becomes

$$r^2 = \frac{S_t - S_r}{S_t}. \tag{7.20}$$

Another important quantity called correlation coefficient is quantified as r and is calculated as $r = \sqrt{r^2}$. The correlation coefficient can also be calculated as [97]

$$r = \frac{n\sum_{i=1}^{n}(x_i y_i) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}}. \tag{7.21}$$

7.3.2.1 Simulations

The data we use in this section is from Figure 6.4(d), which plots the faulty pulses' amplitudes versus slopes for four different states of the gearbox, "Spur1 - Good", "Spur2 - Chipped", "Spur4 - Broken" and "Spur5 – Chipped & Broken". Four distinct clusters are seen in Figure 6.4(d). Linear least-square regression analysis is performed for each cluster. Figure 7.11 shows the results.

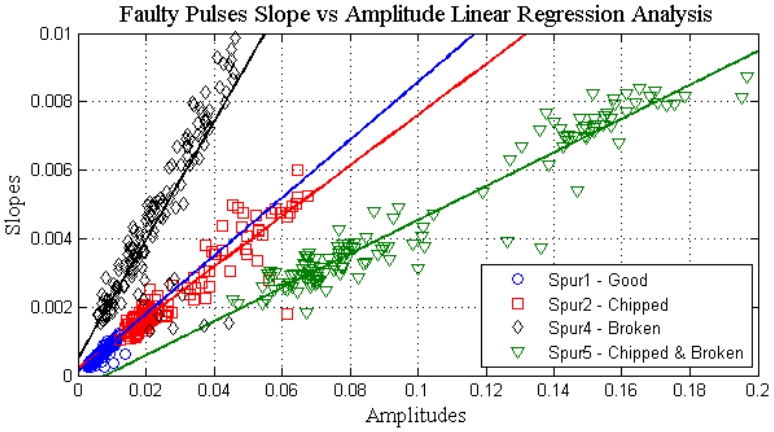


Figure 7.11: Linear Regression Analysis

Table 7.2 lists all the linear equations calculated through linear least-squares regression analysis along with respective coefficient of determination and correlation. The value of $r^2 = 1$ in Equation 7.20 determines a perfect line fit that explain 100% variability of the data. In Equation 7.20, $S_r = 0$ for a perfect fit and $S_r = S_t$ explains the fit with no improvement. In Table 7.2, the line slops of "Spur1-Good" and "Spur2-Chipped" cases are almost similar as evident from Figure 7.11 visually. But the "Spur1-Good" cluster points contain very small values of amplitudes and slopes. The values of amplitudes and slopes for faulty clusters are relatively higher than the good cluster. Any new, incoming data point will belong to a cluster with the fitted line closer to the data point in distance.

Table 7.2: Linear Regression Analysis

Data Set Case	Model	r^2	r
Spur1 - Good	$y = 0.0001 + 0.0847x$	0.8094	0.8996
Spur2 - Chipped	$y = 0.0002 + 0.0741x$	0.8604	0.9275
Spur4 - Broken	$y = 0.0005 + 0.1729x$	0.7367	0.8583
Spur5 – Chipped & Broken	$y = -0.0004 + 0.0494x$	0.9315	0.9651

7.4 Conclusion

This chapter discusses techniques for prognostics and clustering of machine health related features. Two different techniques are used for prognostics, NARX and ANFIS. The prognostics performance of the predictors is illustrated on two data-sets, sunspot activity and vibration index. Both NARX and ANFIS predictors perform quite well. The performance of the NARX predictor is found to be better than ANFIS in predicting vibration index time history. Results show the effectiveness of the predictors in estimating the variations of the monitoring indices. In this work, one-step-ahead prediction is considered and a recursive multi-step prediction is also shown for both the NARX and ANFIS cases. Both NARX and ANFIS have the potential to capture the dynamics of a nonlinear dynamic system but these models are not without problems. They have limitation in learning long time dependences due to the "vanishing gradient",

and like any dynamical system are affected by instability, and have lack of a procedure of optimizing embedded memory.

Fuzzy c-means clustering and linear regression analysis are used for features clustering. Both the methods are able to group the features into clusters correctly. After clustering and prognostics of features has been performed, decision support module uses the information to extracts the representative rules that describe the health of the gearboxes, related faults, causes, and consequences. In the following chapter, we demonstrate three such techniques that can be used in implementing decision support module.

CHAPTER 8

Cause and Consequence Analysis Methods

8.1 Motivation

The determination of what needs to be repaired, when it needs to be repaired and how it needs to be repaired is the main task of a maintenance function. Decisions are taken by interpreting available information and by acquiring or inferring necessary information that is not yet available. Hence, the decision making process becomes a critical task, especially in a complex environment. Clustering, prognostics, and diagnosis are inevitable to aid in the decision making process and a desired intelligent decision has to be made before the failure actually occurs. For example, in rotating machinery, the condition monitoring system may detect that the level of vibration on a bearing is too high and the intelligent decision process looks at the vibration pattern, performs a detailed diagnosis and suggests the solutions automatically. It can also determine whether the problem is due to bad bearings, misalignment or low oil level, to name a few.

The output of the intelligent decision support system is to suggest the cause of the problem and how to tackle it efficiently. This leads to considerable efficiency in the maintenance process but at the same time increases the complexity of the decision support system. Hence, there is a need of automated methods to support the maintenance personnel in taking efficient decisions while providing them with meaningful results and keeping the information overflow to a minimum. In this chapter, we propose three decision support techniques as shown in Figure 1.8 demonstrate their applicability on gearbox fault diagnosis.

8.2 Fault Semantic Network (FSN)

The concept of semantic network was first proposed by Richens in 1956 [118]. It is a network structure that represents relations between concepts. The concept of semantic network was further developed by Collins and Quillian in 1969 [119] where they introduced semantic network in a tree structure (directed or undirected graph) consisting of nodes and arcs. The nodes represent concepts and the connections show relations between nodes and as depicted in Figure 8.1.

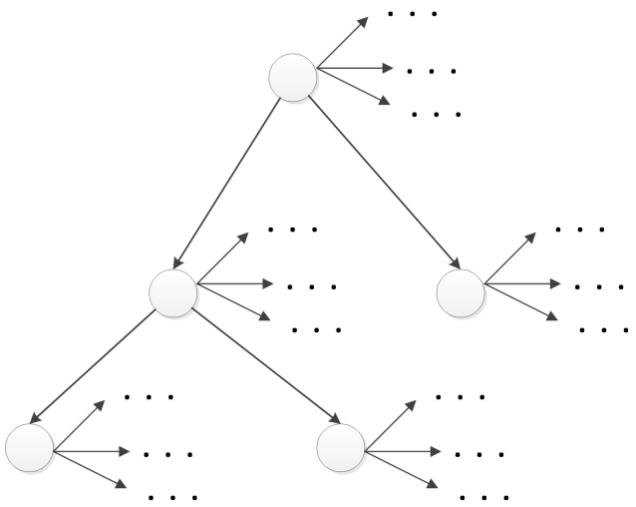


Figure 8.1: Tree Structure of a Semantic Network

Fault semantic network (FSN) originally realized by Gabbar in [120] and further elaborated by Gabbar in [121; 122], is a mean of representing fault knowledge based on relationships between objects. In FSN, the nodes correspond to different faults/causes/consequences and the links between them describe the dependencies. Initially, FSN is constructed based on ontology structure of fault models on the basis of process object oriented methodology (POOM) where failure mode (FM) is described using symptoms, enablers, variables, causes, consequences, and repair actions. Rules are associated with each transition of the causation model within FSN. The rules can be quantitative (probabilistic) or qualitative. For example, failures related to gear tooth breakage might be associated with a qualitative rule such as

IF Structure = Gears AND PV = Vibration AND Symptom = Mesh Frequency Sidebands AND Dev = Very-High THEN FM = Gear Tooth Breakage AND Consequence = Damage/Production Loss AND Repair = Replacement.

These rules are initially defined in generic form based on domain knowledge, i.e., regardless of plant specific knowledge and then further explained or trained for plant specific knowledge based on observations. As described above, the structure of FSN represents relations between variables quantitatively or qualitatively and this can be done through (1) probabilistic approach, (2) fuzzy approach and (3) mathematical model approach. Each approach is discussed as follows:

Probabilistic Approach: In probabilistic approach, a probability value is assigned to each node depending upon its hierarchy in the network as a parent node or a child node. We will further discuss the probabilistic approach in detail in this chapter when we will explore the Bayesian belief networks (BBN).

Fuzzy Approach: This type of reasoning is a pure qualitative reasoning. In fuzzy approach, rules are associated with each transition of the causation model within FSN. In other word, relations between variables are described by specifying *if-then* statements. We will further discuss this approach in detail in fuzzy expert system (FES) section of this chapter.

Model Formulation: This type of reasoning is a pure quantitative reasoning. Relationship between two variables is specified by mathematical equation such as $y = \left[\log\left(x + \cos x\right) - \cos^2 x \right]$, where x is an independent variable and y is a dependent variable, or NN or ANFIS models. The mathematical models can be derived from system identification theory using genetic programming (GP) or other statistical based methods or the methods discussed in Chapter 7 like NN and ANFIS.

8.2.1. Process Object Oriented Methodology (POOM)

As described above that the FSN is constructed based on ontology structure of fault models on the basis of POOM. The POOM is an object oriented approach to construct the

process model in its static, dynamic or functional paradigms. In static paradigm, the faults are related with structures of machines like gears, shafts or bearings. In dynamic paradigm, the faults are related with dynamic behaviour of machines like over-loading, saturation or overheating. In functional paradigm, the faults are related with operation of machines like start-up, shutdown or wrong operation. Figure 8.2 shows the basic architecture of the POOM. All three paradigms are explained. The static view describes facilities, materials and topologies. Dynamic view describes behaviour required to do necessary actions and operational view describes purpose of each structure and set of actions to achieve desired functions [120].

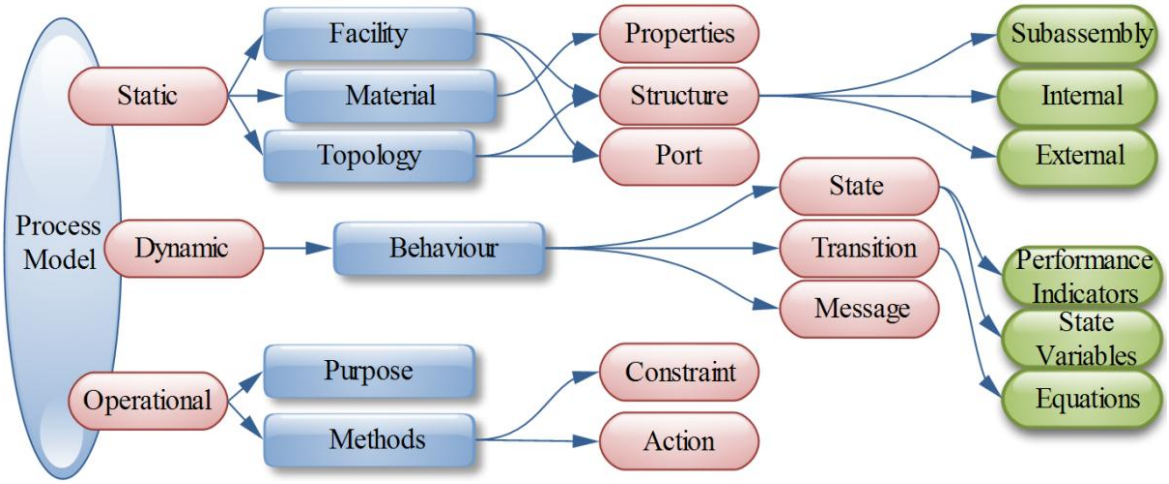


Figure 8.2: The POOM Methodology [120]

Figure 8.3 shows an instance of the FSN database using the POOM methodology. All the tables and the fields in the database are calculated either through historical data or expert opinion.

One elaborated case example for gearbox failure is shown in Figure 8.4. For a particular symptom, there is a corresponding semantic network either fuzzy expert system (FES) or Bayesian belief network (BBN). The symptoms also have related hypotheses with corresponding diagnoses and repairs.

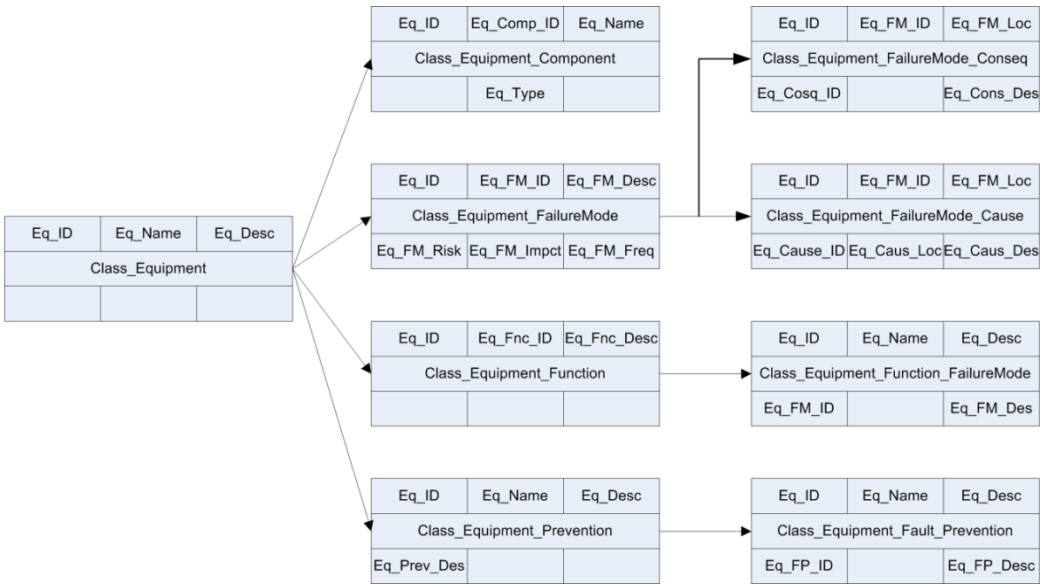


Figure 8.3: FSN Database using POOM Methodology

Case Example
.ID **G20**
.Component **Gears**
.Location **Planetary Gearbox**

.Symptom	S_ID	Symptom	Weight	Semantic Network (BBN / FES)	Hypotheses
	1	Vibration	0.35	FES 1	H2, H3
	2	Temperature	0.65	BBN 1	H1, H2
.Hypothesis	**H_ID**	**Hypothesis**	**Weight**		**Diagnosis**
	1	Overload	0.65		D1
	2	Oil Loss	0.20		D2
	3	Equipment Wear	0.10		D3
	4	Teeth Crack/Breakage	0.05		D4
.Diagnosis	**D_ID**	**(Sensor, Features) Pairs**	**Weight**		**Repair**
	1	D1: {(S1, F1), (S1, F2)}	0.69		R1
	2	D2: {(S2, F3)}	0.75		R2
	3	D3: {(S1, F2), (S1, F4)}	0.89		R3
	4	D4: {(S1, F2), (S1, F4), (S1, F5)}	0.50		R4
.Repair	**R_ID**	**Repair**	**Weight**		
	1	Reduce Load	0.33		
	2	Replenish Oil	0.33		
	3	Lubrication	0.33		
	4	Replacement	0.33		
.Version	**Last_Updated**	**Case_Quality**	**Success**	**Failure**	**Conditions**
	14 Feb, 2013	0.70	7	3	Full Load Windy

Sensor_ID	Name	Location	Description
S1	Acc_Bearing1	Generator Drive End	Accelerometer (100 kHz)
S2	Oil_Planetary	Gearbox Oil Level	Oil Level Sensor
S3	Temp_Bearing1	Generator Drive End Bearing	Temperature Sensor

Feature_ID	Feature Name	Description	Detection
F1	1_2_3: ShaftOrder	Amplitude measurement of 1x, 2x and 3x fundamental shaft frequency	Shaft load.
F2	SRS_Transient	Novel shock response spectrum feature. Detection of transient severity level in vibration.	Shaft load, crack, damage, scuffing.
F3	Oil_Level1	Oil level measurements	Oil level.
F4	Kurt_Transient	Time domain kurtosis, crest factor, P2P amplitude detection.	Equipment wear measurements.
F5	Gear_Mesh	Side bands measurememts around gears mesh frequency.	Gear teeth breakage.

Figure 8.4: FSN Case for Gearbox Failure

8.3. The Fuzzy Expert System (FES)

Lofti Zadeh introduced the fuzzy logic theory in 1965 [115]. Fuzzy logic deals with uncertain and imprecise information mathematically. In fuzzy theory, the linguistic constructs are handled by a mathematical mechanism. The linguistic constructs can be of the form “cold”, “warm” or “hot”. The fuzzy logic provides an inference mechanism for imprecise and partial information to incorporate human reasoning capabilities. Fuzzy systems use IF-THEN rules, where the IF part is called “antecedent” and the THEN part is called “consequent”. The basic configuration of a pure fuzzy system is shown in Figure 8.5. A fuzzy inference mechanism maps fuzzy IF-THEN rules from input space to output space by using fuzzy logic methods.

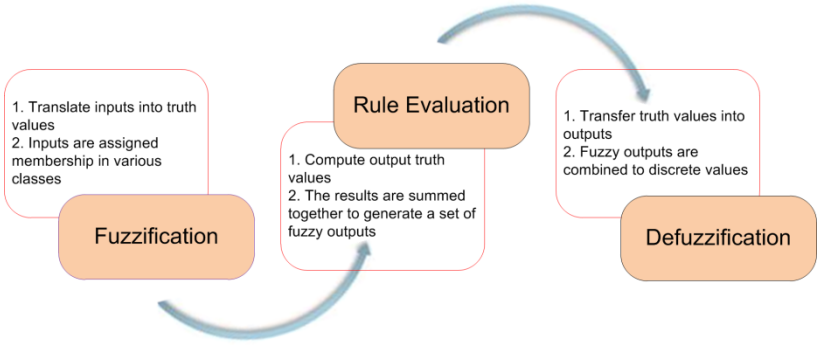


Figure 8.5: The Fuzzy Expert System

For example, fuzzy inference rules for a gearbox with a vibration sensor can be formulated as

IF 'Kurt' IS 'N' AND 'GMF' IS 'N' AND '1x' IS 'N' THEN 'GearFailure' IS 'Gear No Wear'
IF 'Kurt' IS 'H' THEN 'GearFailure' IS 'Gear Component Wear'
IF 'GMF' IS 'H' THEN 'GearFailure' IS 'Gear Misalignment'
IF 'GMF' IS 'VH' THEN 'GearFailure' IS 'Gear Misalignment'
IF 'GMF' IS 'VVH' THEN 'GearFailure' IS 'Gear Misalignment'
IF 'Kurt' IS 'VH' THEN 'GearFailure' IS 'Gear Cracked'
IF 'Kurt' IS 'VVH' THEN 'GearFailure' IS 'Gear Cracked'
IF '1x' IS 'H' THEN 'GearFailure' IS 'Gear Unbalance'
IF '1x' IS 'VH' THEN 'GearFailure' IS 'Gear Unbalance'
IF '1x' IS 'VVH' THEN 'GearFailure' IS 'Gear Unbalance'

Where, *N*=Normal, *H*=High, *VH*=Very High, *VVH*=Very Very High, *Kurt* = Kurtosis, *GMF*=Gear Mesh Frequency and *1x* = Shaft fundamental frequency. Different parameters

are used for fuzzy inference that includes vibration based time-domain and frequency-domain features like kurtosis, FFT meshing frequency and shaft fundamental frequency levels. The corresponding membership functions in the fuzzy inference systems are shown in Figure 8.6.

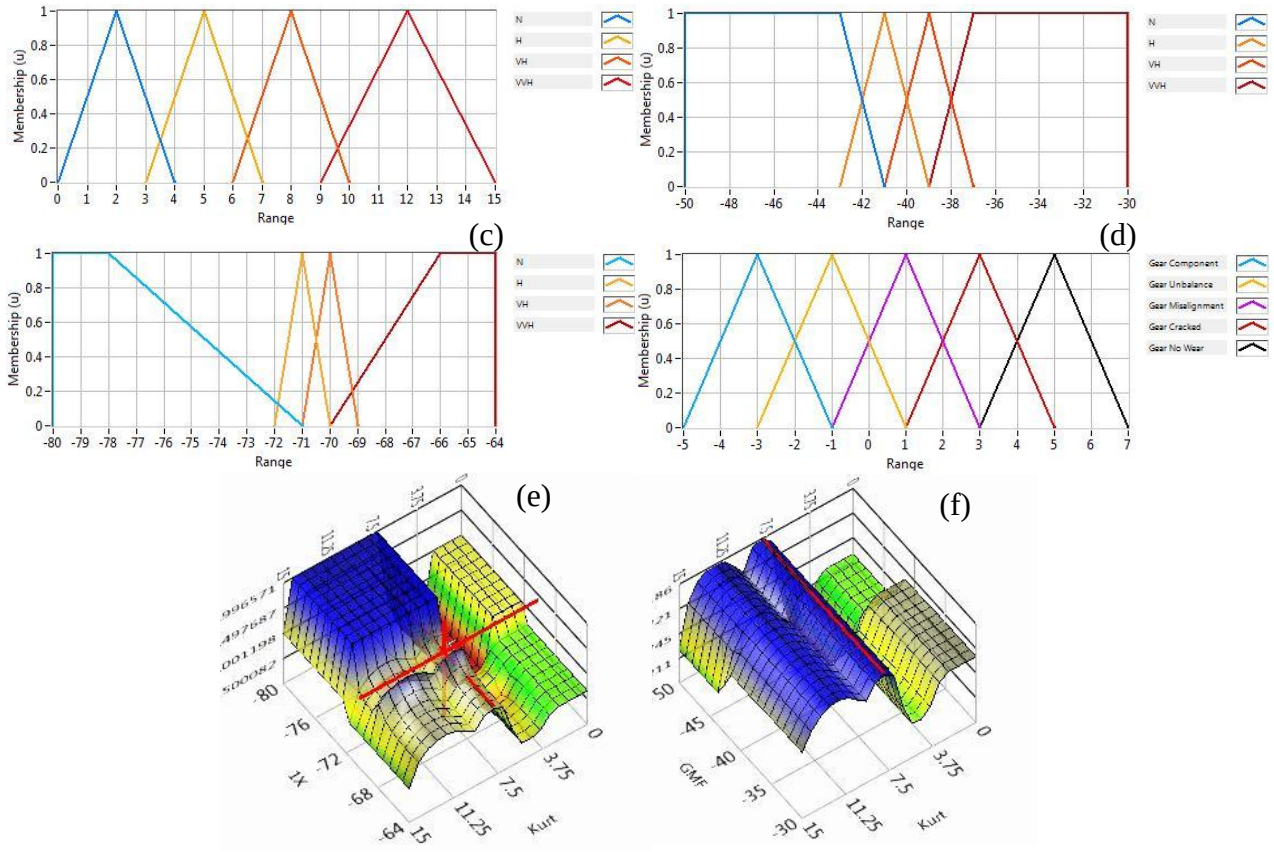


Figure 8.6. FES: Gearbox Failure (a) Kurtosis MFs (b) GMF MFs (c) 1x MFs (d) Decision MFs (e) Kurt vs 1x Fuzzy Surface (f) Kurt vs GMF Fuzzy Surface

8.4. Bayesian Belief Network (BBN)

Bayesian belief network (BBN) is widely used for fault diagnosis, root cause and consequence analysis. Details about some previous research in BBN for fault diagnosis are discussed in Chapter 2. Figure 8.7 shows the process of constructing BBN for gearbox fault diagnosis, root cause and consequence analysis. Data collected from gearbox, combined with maintenance history, maintenance expert opinion and evidences collected by field operators is used for BBN construction. The K2 algorithm [123] is used for learning the BBN structure and node probabilities. After the network has learned from

the training data, junction tree algorithm [124] is used to query the network. The inference process updates the network probabilities according to the evidences entered.

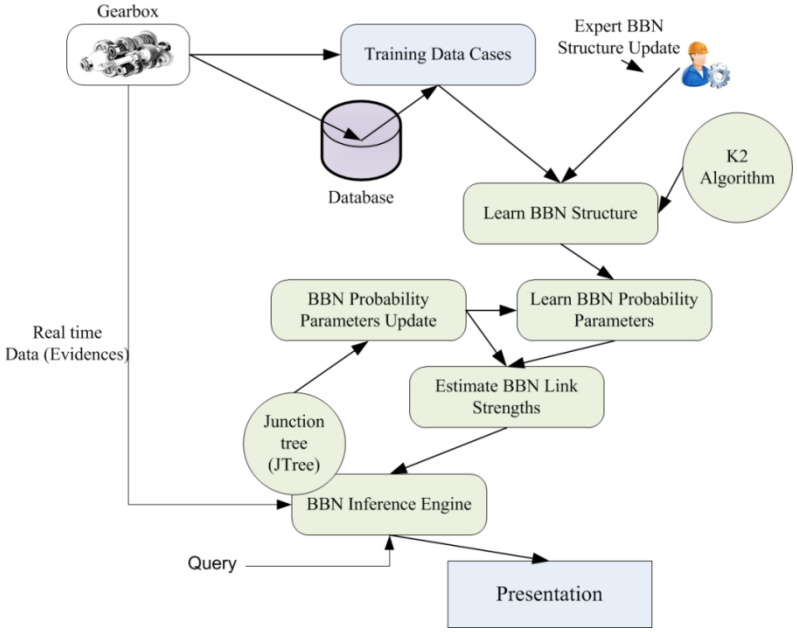


Figure 8.7: BBN: Learning and Inference Process

8.4.1 Querying the BBN

The BBN has the ability to change dynamically by incorporating new data and updating its internal structure, rules and interaction strengths. In order to get expected results from the BBN, it is necessary to query the BBN in a structured way by incorporating the observations in the BBN to update the BBN structure according to the observed nodes and to get an answer to the query. Below, we describe different types of queries as used in BBN.

Diagnostic Query: In this type, query starts from symptoms and results in causes so the direction is opposite to the arc direction. In an example of gearbox fault diagnosis, vibration test results can be considered as symptoms. The BBN can be queried to recognise the heavy load or tooth breakage as causes.

Predictive Query: This type of query predicts faults in advance even without assuming symptoms. For instance increasing load on gearbox can cause particular vibration index to go high.

Inter-causal Query: This type of query is used when multiple causes result in one symptom. For instance, both high load and tooth breakage can result in one symptom, which is high vibration index in gearbox.

Figure 8.8 shows an example of a BBN for gearbox failure. In Figure 8.8, either the node probabilities are assigned by an expert or learnt by historical data. The BBN is very flexible and powerful to query in any direction. In the query process, we enter an evidence of an occurrence in the BBN and the BBN returns the answer to the query and updates the probability tables.

Diagnosis Query:

The diagnosis query can be as follows:
What is the probability of the cause "Begrime (a)" provided the "Large Mag 1x (m)" has occurred?
Evidence: Large Mag 1x (m) = True
Query the BBN:

P (a | m)	=	81.09% False
		18.91% True

The BBN returns that there is 18.91% chance of "Begrime (a)".

Prediction Query:

The diagnosis query can be as follows:
What is the probability of observing "SRS index (q)" provided the "Lack of Lubrication (e)" has occurred?
Evidence: Lack of Lubrication (e) = True
Query the BBN:

P (q | e)	=	91.51% False
		08.49% True

The BBN returns that there is 8.49% chance of observing "SRS index (q)" above threshold.

Inter-causal Query:

The inter-causal query can be as follows:

What is the probability of observing "Fatigue (g)" and "Corrosion (h)" combined provided the "SRS Index (q)" has occurred?

Evidence: SRS Index (q) = True

Query the BBN:

P (g, h | q) = 76.03% False
23.97% True

The BBN returns that there is 23.97% chance of observing both the "Fatigue (g)" and "Corrosion (h)" together.

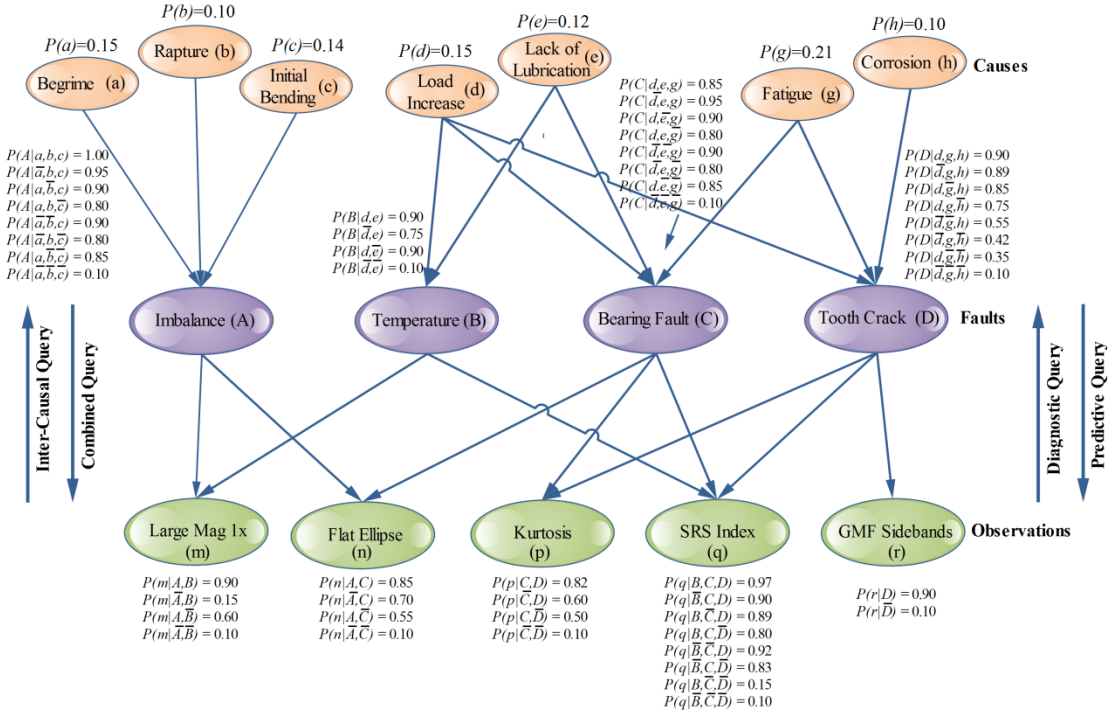


Figure 8.8: BBN for Gearbox Failure

8.5 Conclusion

In this chapter we discussed three methods for cause and consequence analysis, FSN, FES and BBN. FES is a qualitative method while BBN is a probabilistic and FSN is flexible to incorporate both FES and BBN as sub-methods. The POOM methodology

discussed in FSN has a structured approach in storing the data in the database. Based on the complexity of the application and decisions to be made, one or all of the above mentioned methods can be used.

In the next chapter, we will apply some of the developed analysis techniques on real world data emanating from gearboxes inside wind turbines. The main purpose would be to access the health status of the gearbox or mechanical components inside the mechanical drive train of a wind turbine.

CHAPTER 9

Condition Monitoring of Wind Turbines: Case Studies

9.1 Motivation

The advent of modern technologies and increased environmental awareness makes wind turbines a unique and important technical identity. For the last couple of decades the technology to harness the wind energy has flourished in its aerodynamic shape, structural dynamics, control system, and efficiency. Over the years, the annual energy produced by a wind turbine has increased massively and the weights of the turbines and noise levels they emit have been halved. As compared to other renewable energy resources and technologies to harness these energies, wind power generation is more technologically mature and relatively cost effective. Wind energy has a big share in the future of green energy of the world [125].

The importance of reliability cannot be ignored in any technological design, and formulation of novel strategies and technologies to mediate incipient faults and inherent shortcomings in the design has craved for years. The reliability, maintainability, and robustness of a design can be enhanced by developing such technologies. In wind turbines, this can be achieved by implementing an efficient, adaptable, and responsive system for condition monitoring and fault diagnosis. Such systems are already under constant development and use in the industry as described in Chapter 3. The main aim of this thesis work is to contribute in technological advancements for development of such systems.

This chapter uses some of the methods described in previous chapters along with some new techniques for condition monitoring and fault diagnosis of wind turbines. Two real world case studies are used where wind turbines developed faults in gears and bearings.

9.2 Case Study 1: Channel7 - Planetary Stage (Gearbox)

The data used in this section emanate from the planetary gearbox inside the wind turbine. The details about the data are described in Section 1.10.3. Adaptive band-pass filtering is performed to extract the faulty information present in the gearbox. Figure 9.1(a) shows a vibration signal sample of Channel7: Planetary Stage (Gears) and its kurtogram in Figure 9.1(b). The signal is adaptively band-pass filtered according to the information returned by the kurtogram and the conventional optimization techniques presented in Chapter 4. The filtered signal is shown in Figure 9.1(c) with filtered kurtogram in Figure 9.1(d). The filtered signal in Figure 9.1(c) clearly shows the presence of the impulses. Initial analysis on the filtered signal shows promising results. The duration between two consecutive pulses is changing because of the change in rotation peed (RPM). Since we only analyse the signal in two seconds window, the changing RPM does not have dramatic effects on the analysis. Performing pulse position analysis and taking the average of the changing time values in Figure 9.1(c) as follows:

Generator shaft RPM and corresponding frequency = 1676.5 → 1676.5/60 = 27.94 Hz
Time Duration Average = (0.312 + 0.367 + 0.400 + 0.316 + 0.286 + 0.255)/6 = 0.3226
Relative Frequency at 27.95 Hz of Generator Shaft = 1/0.3226 = 3.099 Hz
RPM of the damaged gear = 3.099 x 60 = 185.9

It is evident form the above calculations that the RPM of the faulty gear is 185.9 (3.099 Hz). In this particular case, we do not know the basic kinematics data of the planetary gearbox but a peak in FFT at 3Hz in Figure 9.1(e) indicates a component revolving at 3Hz. Figure 9.1(f) plots the kurtosis history for the planetary gearbox. The signals are adaptively filtered through the band-pass and wavelet filters presented in Chapter 4 and

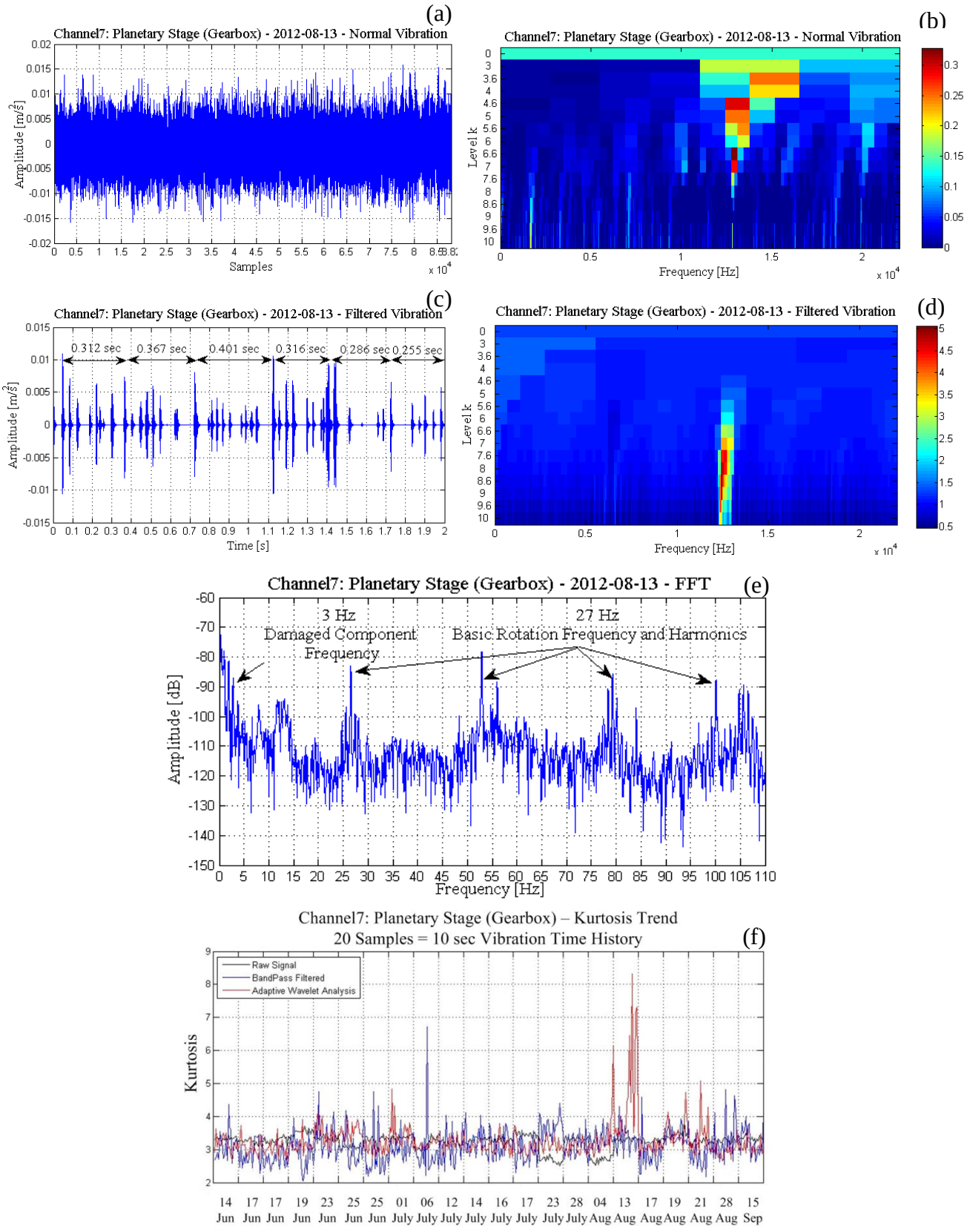


Figure 9.1: Adaptive Filtering: Channel7 Planetary Gearbox

the kurtosis of the filtered signal is calculated. In Figure 9.1(f), the kurtosis is around 3.0 in most cases, but the value of the kurtosis jumps to 8.32 on 13 August, 2012 and this is the time when the fault has been detected. The fault could have started earlier but there is no vibration signal present between 04 August, 2012 and 13 August, 2012.

9.3 Case Study 2: SCADA Data

The data used in this section is recorded from the SCADA system of a wind turbine. The details about the data are described in Section 1.10.3. The data are spanned over two years (May, 2010 – April, 2012). There are a total of 1198 chunks of the data with 5 outliers as listed in Table 9.1 and in each chunk there are 4096 vibration samples for Channel1: Generator Drive End (Bearing). Each chunk also contains the time history of the following parameters:

(a) Time (Date/Time)
(b) Speed (RPM)
(c) Wind speed (m/s)
(d) Output power (kW)
(e) Bearing temperature (^{o}C)
(f) Pitch angle (degrees)
(g) Torque (N.m)

9.3.1 Outliers

Some of the issues with the data are treated as outliers because of the following reasons (Table 9.1):

(a) Presence of RPM values is zero.

(b) Power output is negative.

(c) Torque is zero.

(d) Wind speed is large or negative.

Table 9.1: SCADA Data Instances with Out-of-Range Values

Date	Time	RPM	Power (kW)	Torque (N.m)	Wind Speed (m/s)
24/05/2010	02:10:20	0	-3	0	-91
24/05/2010	04:12:23	0	-3	0	-96
29/05/2010	10:15:56	0	-3	0	39
29/05/2010	16:33:47	0	-3	0	137
29/05/2010	17:11:24	0	-3	0	126

9.3.2 The Power Curve

The overall health of a wind turbine can be accessed by the shape of the power curve. An ideal power curve is a relationship between wind speed and the output power modelled as a sigmoid function. The power curve is plotted for the minimum and maximum wind speeds or cut-in and cut-out wind speeds for a particular wind turbine. A power curve based on SCADA data can deviate from the actual one because of the following reasons: (1) the values of the output powers are identical for the same wind speed; (2) some values of the power output are negative; (3) power errors due to malfunctions in turbine's systems, sub-systems and components. So, the output power and wind speed are the important parameters for detection of any malfunction in a wind turbine. Figure 9.2 plots the relationship between wind speed and output power or the power curve.

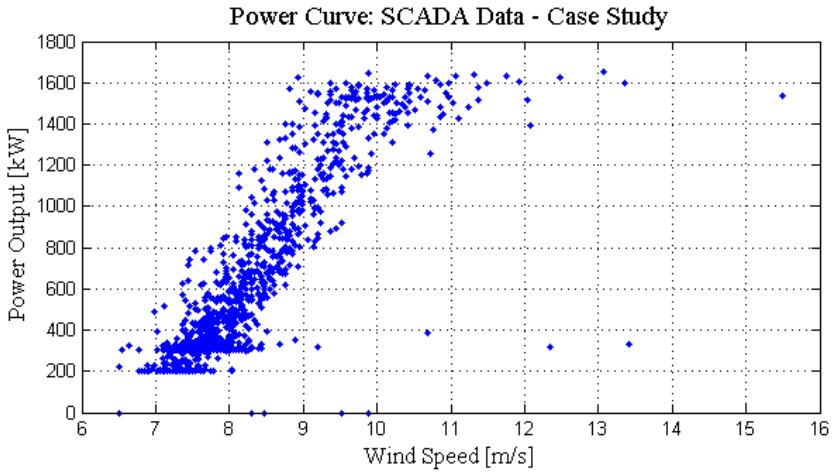


Figure 9.2: Wind Turbine's Power Curve

9.3.3 Neural Network Prediction of the Output Power

We use non-linear input-output NN for predicting the power values from three parameters wind speed, RPM, and torque. We modify the Equation 7.9 as

$$y(n+1) = f\left[x(n), x(n-1), ..., x(n-k)\right]. \tag{9.1}$$

Where, $\mathbf{x}(n)$ and $\mathbf{y}(n)$ are the input and output of the system at time step n, and $f(.)$ is a non-linear mapping function. We approximate $f(.)$ by a non-linear input-output feed-forward NN with multi-layer perception (MLP). The main purpose is to predict future value $y(n+1)$ of the series $\mathbf{y}(n)$ given k past values of $\mathbf{x}(n) = \left[x(n), x(n-1), ..., x(n-k)\right]$. The general structure of a feed forward MLP is shown in Figure 9.3, which involves the input layer, the output layer, and several hidden layers of nodes. It is the inclusion of one or more hidden layers that makes the NN capable of approximating nonlinear functions (mappings) or classifying patterns in nonlinear separable classes. The input–output equations of the kth neuron in Figure 9.3 are as follows [126]:

$$y_k = \Phi(u_k - \theta_k), \qquad u_k = \sum_{j=1}^{n} w_{kj} x_j \tag{9.2}$$

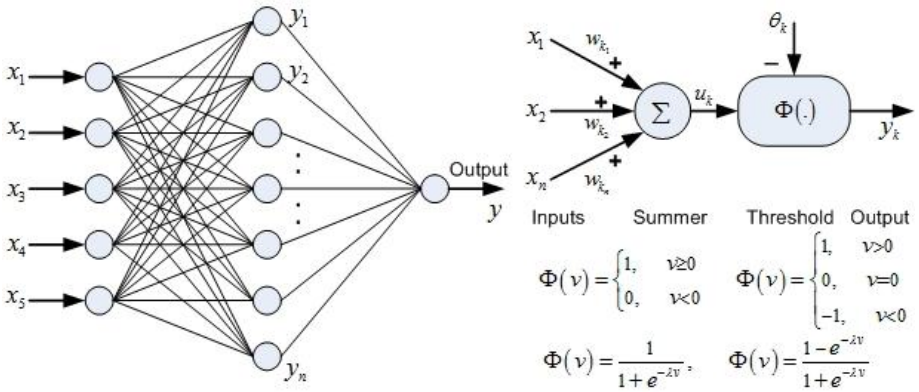


Figure 9.3: Multilayer Feed Forward NN

In Figure 9.3, $x_1, x_2, ...x_n$ are the input elements, u_k is the output of the summer, θ_k is a given threshold, $w_1, w_2, ...w_n$ are the neuron's synaptic weights and $\Phi(.)$ is the neuron's activation function. The most important characteristic of the neural network is the ability

to learn from the environment. The learning is the process of updating the synaptic weights of the neural network through an enforcement applied to it from the environment. The learning rule has the following form [126]:

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \tag{9.3}$$

Where $\Delta w_{kj}(n)$ is the adjustment of the weight $w_{kj}(n)$ at the nth time stamp. The computation of $\Delta w_{kj}(n)$ is performed through a well-known back propagation algorithm. Since we have a total of 1193 valid data samples over the span of two years we use 50% (596 samples) for training, 20% (239 samples) for validation and 30% (358 samples) for testing.

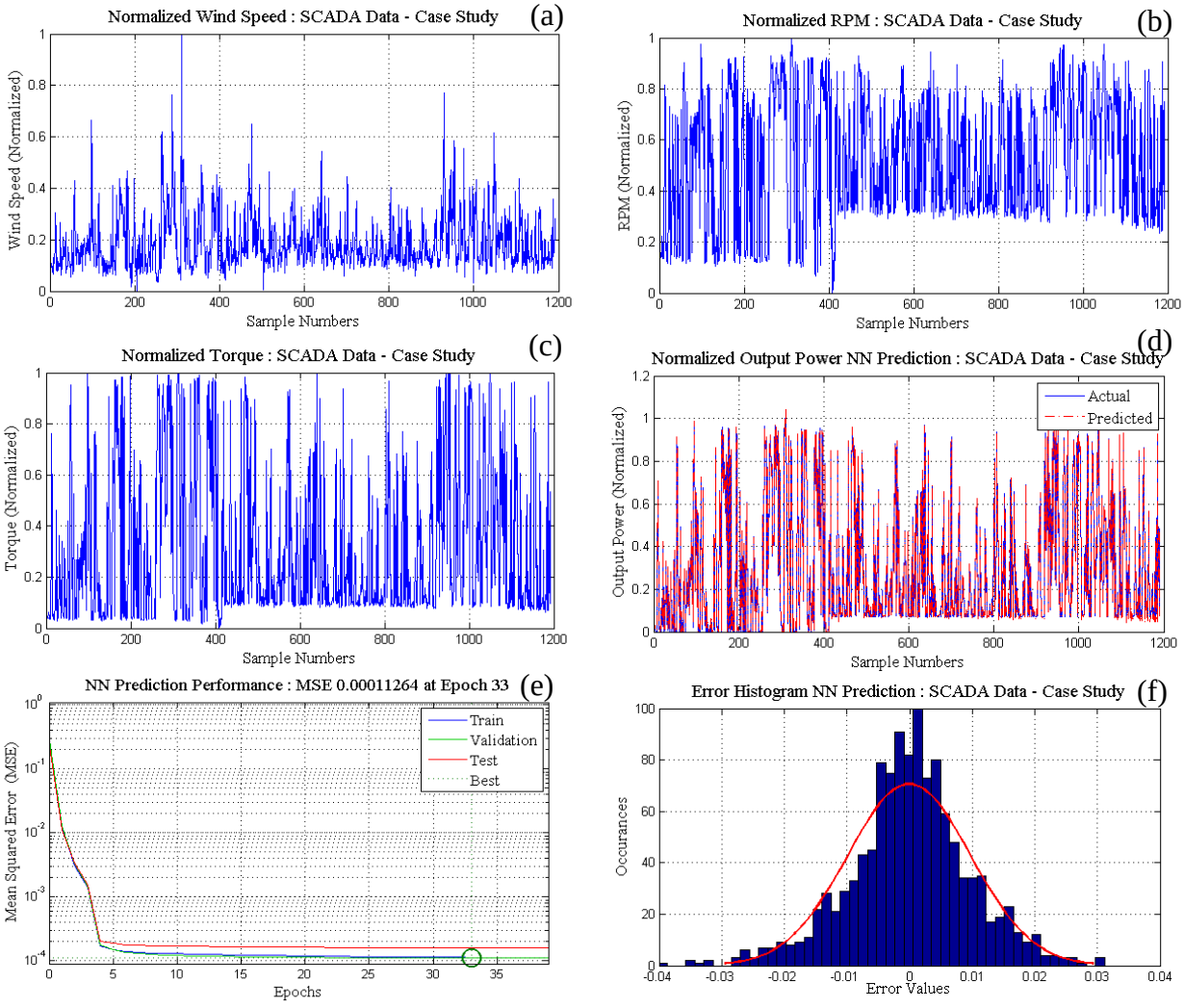


Figure 9.4: Feed Forward NN Prediction

Figures 9.4(a), 9.4(b) and 9.4(c) show the normalized wind speed, normalized RPM and normalized torque, respectively, used as inputs to the NN. Figure 9.4(d) shows the actual normalized output power and predicted output power through the NN. The prediction performance is shown in Figure 9.4(e) along with the histogram of the prediction error in Figure 9.4(f). The prediction error's histogram being centred at mean value of zero and distributed normally shows the accuracy of the prediction.

9.3.4 Condition Monitoring Channel1: Generator Drive End (Bearing)

The vibration data emanating from "Channel1: Generator Drive End (Bearing)" are used in this section. We have 1198 samples with 5 outliers and this gives us 1193 valid samples to analyse. Vibration feature extraction process in Figure 7.1(a) and Equation 6.7 is used along with principal component analysis (PCA) and fuzzy c-means clustering as shown in Figure 9.5.

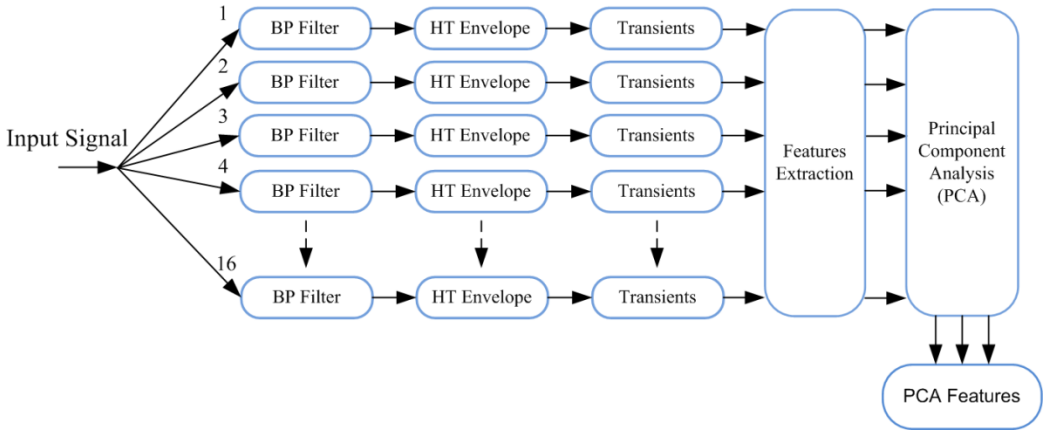


Figure 9.5: Principal Component Analysis (PCA) of Vibration Features

In the band-pass filter bank, we use 16 gammatone filters. Figure 9.6 plots all the 16 vibration features extracted from the 1193 samples. Boxplots are used to explore the distribution of the features. In descriptive statistics, box plot is a convenient way of explaining the statistical spread of the numerical data graphically. Figure 9.7 shows the box plots for 16 vibration features plotted in Figure 9.6. It becomes immediately apparent that features 1,2,15 and 16 dominate the set. Based on this information, we can suppose that these four features contribute mostly to the main dynamics of the features set.

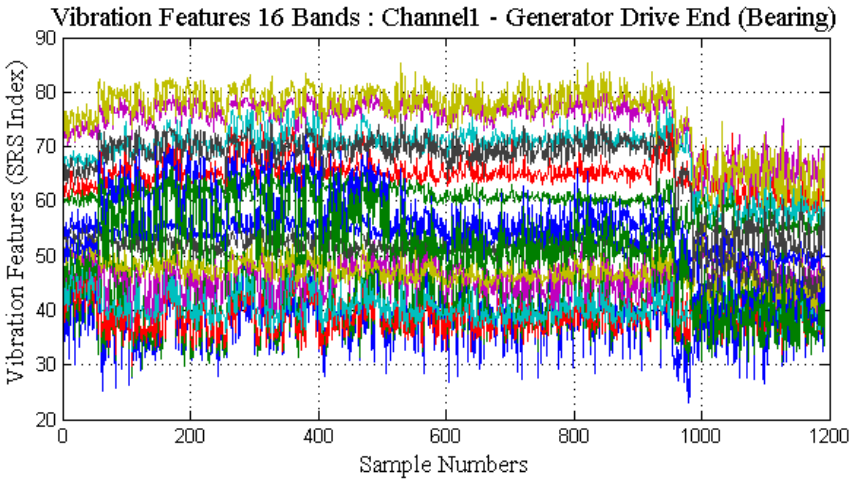


Figure 9.6: Vibration Features – Channel1: Generator Drive End (Bearing)

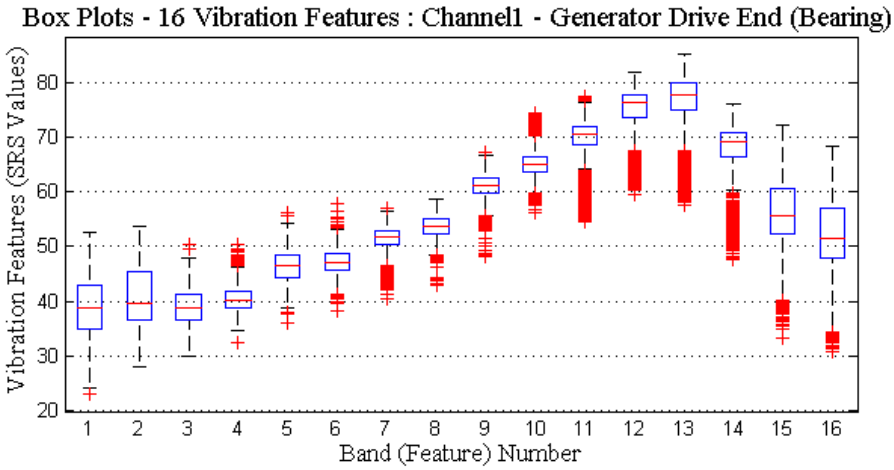


Figure 9.7: Box Plots – Channel1: Generator Drive End (Bearing)

However, this assumption is not always true. There can be cases where the contributions of dominant features are not that important as compared to some non-dominant features. So, box plots are not always helpful in choosing the right features that captures the dynamics of the system. One of such methods is PCA [99; 100] first proposed by Hotelling in 1936 [127]. The PCA is used extensively in features extraction, data dimensionality reduction and descriptive statistics. We perform PCA analysis on 16 vibration features and analyse the resultant principal components. The process also reduces the data dimensionality and at the same time captures the data dynamics. Also, the process discards the redundant information in the data set. This is useful in preventing the overflow of the information. In PCA, we

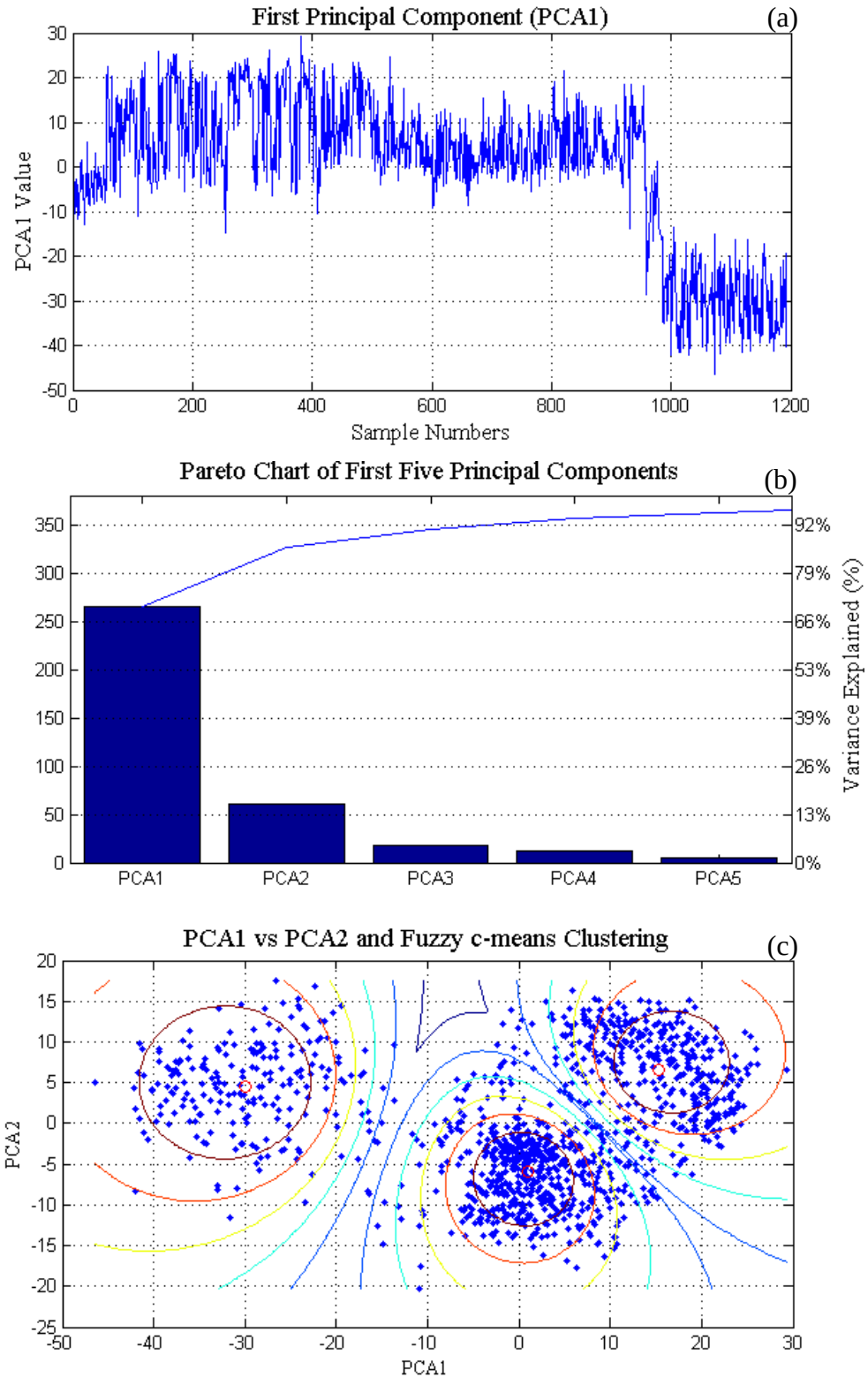


Figure 9.8: PCA Analysis and Fuzzy c-means Clustering

construct new variables in principal component space as linear combinations of original variables. Figure 9.8(a) plots the first principal component of the 16 features and the Pareto chart in Figure 9.8(b) shows that the first two principal components explain around 86% variance of the data. The first two principal components are plotted in Figure 9.8(c) where three distinct clusters are found when the fuzzy c-means clustering is performed. If we analyse the 16 extracted features one by one in Figure 9.6, we see that feature no. 15 or feature no. 16 can explain the system dynamics. Figure 9.9 plots vibration features no. 15 and 16.

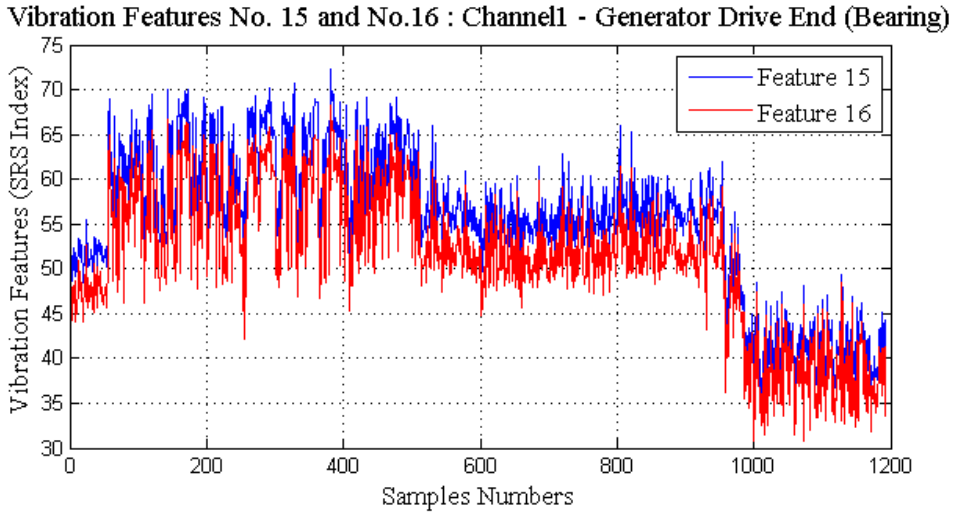


Figure 9.9: Vibration Features No. 15 and No. 16

The high values of features 15 and 16 in Figure 9.9 or for that matter the high values of PCA1 in Figure 9.8(a) indicate the tendency of the bearing towards roughness or faults. The 1193 vibration feature samples shown in Figures 9.6, 9.8(a) or 9.9 are actually not continuous. There are some pauses in the collection of the data over the span of two years. We insert these pauses in Figure 9.10 to get clearer picture about the wind turbine's status. Figure 9.10 shows that the turbine was down for 37 days in May-June 2011. This was due to the maintenance activities performed. After maintenance, the amplitude of the bearing's vibration signatures (SRS index) came down on y-axis and became smoother. Another downtime was in Jan-Feb 2012 where the bearing was replaced with the new one causing the vibration signatures amplitudes to come down on y-axis.

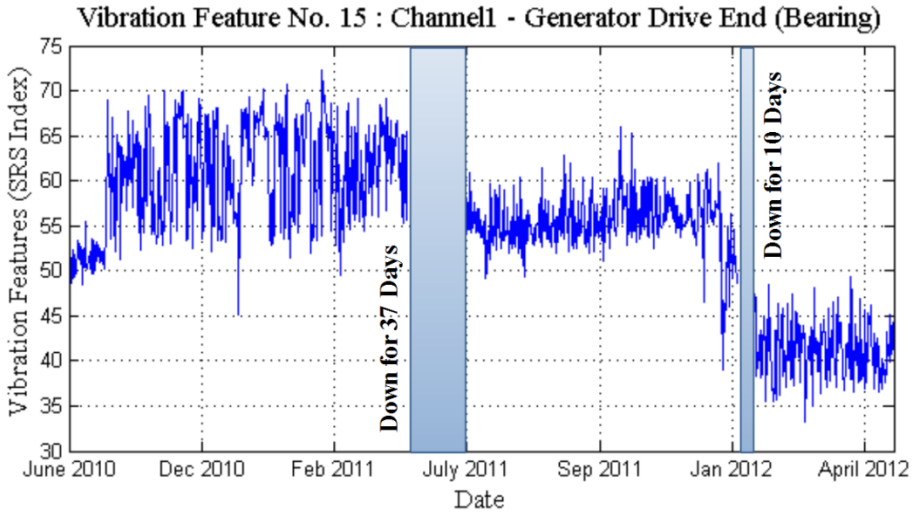


Figure 9.10: Vibration Feature No. 15 and Turbine Down Time

Let us analyse one vibration signature from Feb, 2011 and use the techniques developed in Chapter 4 or 5 to filter the signal for detection of any faulty pulses present in the signal. The results are shown in Figure 9.11, where sample no. 220 is analysed. The signal is band-pass filtered with filter parameters $F_c = 6524.3, Q = 0.70337, N = 4, R_p = 1.0048$. We can clearly see that the faulty pulses in the filtered signal and the kurtosis is maximized to 9.42.

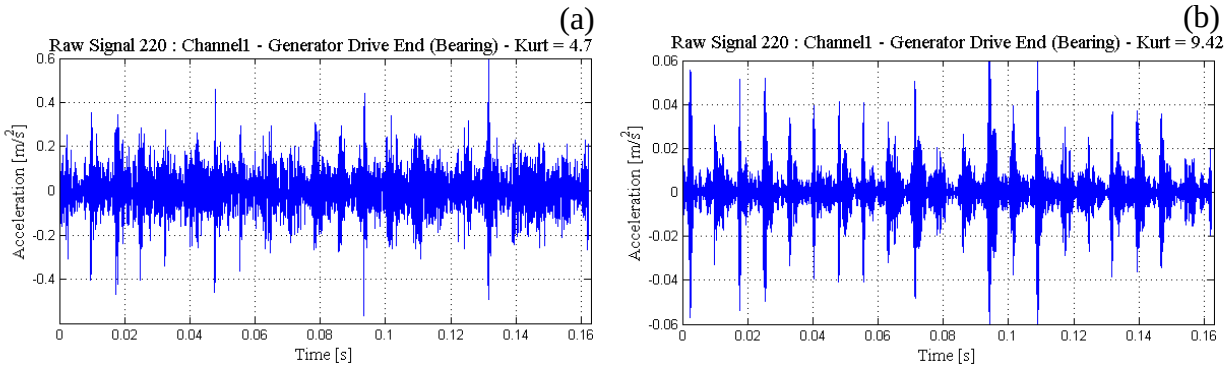


Figure 9.11: Adaptive Band Pass Filtering Fault Detection

Anomalies are found in the bearing and we see that the data points in the power curve in Figure 9.2 are deviating from their original path. This is because of the power fluctuations caused by bad bearing. Figure 9.9 shows the fluctuations in vibration indices 15 and 16 till sample no. 430 on the x-axis and Figure 9.4 (d) confirms the fluctuations in output power till sample no. 400 on x-axis.

9.4 Conclusion

In this chapter, two different case studies for monitoring wind turbines are presented. In the first case study, we detected fault in planetary stage (gearbox) of a wind turbine based on vibration analysis. We successfully detected the fault present in the signal and calculated the rollover frequency of the faulty component. In the second case study, we analysed wind turbine SCADA data for accessing the health status of the wind turbine. The power curve of the wind turbine deviates from its original position if there are anomalies in wind turbine's systems, sub-systems and components. We trained an NN prediction model for the output power of the wind turbine from wind speed, torque and RPM. The model is helpful in predicting the future behaviours of the wind turbine output status. We also performed vibration analysis on wind turbine's generator drive end bearing by using the methods proposed in previous chapters. The vibration features extracted are reduced in dimensionality through the PCA. Fuzzy c-means clustering is performed to determine the health status of the wind turbine. Overall, the wind turbine has performed according to the specifications. The generator's drive end bearing developed a fault with time and that caused fluctuation in the output power. The maintenance was first performed and the bearing was replaced with a new one later on. The bearing status is very clear by analysing the trend of the extracted vibration feature.

CHAPTER 10

Conclusions and Future Research

10.1 Motivation

The aim of this research work is to make advancements in the field of condition monitoring and fault diagnosis of gearboxes in wind turbines. The growing interest in renewable energy resources has increased the need for wind energy systems or wind turbines and therefore, the importance of condition monitoring and fault diagnosis systems (CM-FDS) for wind turbines has unveiled.

10.2 Summary and Conclusions

In this thesis, the challenges and practicalities of condition monitoring and fault diagnosis of wind turbine gearboxes are discussed and different commercially available CM-FDS systems are listed with their analysis capabilities. Also a detailed literature survey is performed to summarize some of the existing state of the art techniques in this domain. However, there are issues and limitations because of variable load and variable speed of wind turbines and severe environmental conditions. There is also a trade-off between the time complexity and the quality of results in fault detection and diagnosis methods. Various methods proposed in past research mainly focus on detection quality at the expense of computation time. The approach is not suitable in many real time applications including monitoring of wind turbines.

This thesis presents some novel techniques to enhance the capabilities of existing CM-FDS systems in variable load and variable speed conditions. Conventional and evolutionary optimization techniques are used to intelligently search for faults present in the signals recorded from wind turbine gearboxes. Efforts have been made to reduce the computational time complexity while preserving the quality of the detection results. Three main processes in fault diagnosis are discussed: (1) intelligent filtering; (2) features

extraction and (3) intelligent decision support. Features extraction also incorporates clustering and prediction modules to differentiate different types of faults and to project the state of the system into the future.

In intelligent filtering module, kurtogram is used as a front-end method for detection of transient faults [22]. Although kutogram is an efficient method for detecting the faulty bandwidth in non-stationary signals in real time, it suffers from low frequency resolution. Filters designed based on faulty bandwidth information returned by the kurtogram showed poor quality of results. In order to overcome this deficiency, and with less computational complexity, we proposed a hybrid filtering framework. We used two types of filtering techniques: (1) band-pass filtering; (2) wavelet analysis and optimized them through conventional and evolutionary optimization methods. The conventional optimization framework included one-dimensional golden section search and multi-dimensional Nelder-Mead methods. The evolutionary optimization framework included real coded genetic algorithm and particle swarm optimization. We saw considerable improvements in detection quality and computational complexity as compared to various filtering methods proposed in research [22; 45; 82; 83]. Another filtering operation used in this module was based on psychoacoustic phenomenon. Psychoacoustic filtering applies gammatone filter bank to mimics the filtering operation of human ear. Not much research has been done on psychoacoustic filtering for vibration based fault detection and especially its applications for gearboxes. All three filtering techniques successfully extracted the faulty information present in the vibration signals.

In features extraction module, we used transient based time-domain features extraction techniques. As described earlier in this thesis, transients are short duration pulses generated by interaction of faulty components in a machine. The shapes and amplitudes of these transients represent the severity of the fault. Therefore, transient based features extraction techniques can best represent the status of the machine. We have successfully demonstrated cases where transient based features are able to differentiate among different types of faults in a gearbox. However, transient based features extraction techniques presented in this work are at early stage of research and need to be investigated in more details as part of the future work.

The intelligent decision support module used three methods: (1) fault semantic networks (FSN); (2) fuzzy expert systems (FES) and (3) Bayesian belief networks (BBN). Due to a complex interaction patterns among the equipment variables, a construction of a mathematical model is very difficult and sometimes impossible. In this case, graphical semantic knowledge and fuzzy expert system is demonstrated to be very helpful. As described in Chapter 2, FSN, FES and BBN have been used for fault diagnosis in process environments and BBN is very mature in medical decision support systems. However, their application for fault diagnosis in machines is relatively new.

10.3 Why Time Domain?

While frequency domain fault detection and diagnosis techniques like fast Fourier transform (FFT) are well-established in the industry and in academia, they pose some disadvantages. The FFT based analysis techniques are not effective at light load conditions and time information is lost in the process [128]. Also an inherent averaging process present in FFT loses the important information present in the signal, especially, low energy faults.

On the other hand, joint time-frequency analysis techniques like wavelet analysys (WA) or short time Fourier transform (STFT) are gaining much attention in recent years. STFT analyses signals with fixed sized windows and gives poor frequency resolution. WA is more flexible in terms of time and frequency resolutions and is suitable in detecting low energy incipient faults in machines. Although, time-domain analysis cannot match well-established FFT diagnosis capabilities, it is good in detecting incipient faults in vibration signals that is otherwise not possible with FFT alone [129; 130].

10.4 The Innovative Contributions of the Thesis Work

The main aim of this research work is to develop techniques for intelligent CM-FDS systems to monitor gearboxes in wind turbines with the help of suitable signal processing and soft computing techniques. In order to demonstrate the viability of the developed

system, different experimental and real world vibration datasets are used. The main contributions of this thesis are summarized as follows:

10.4.1 Literature Review

(1) A literature review has been performed to study the existing methodologies in the research and to determine their shortcomings in order to bridge the gap.

(2) A small survey has been conducted in order to determine the existing CM-FDS systems used in industry, their analysis capabilities and shortcomings.

10.4.2 Hybrid Optimization Framework

(1) A hybrid conventional optimization framework has been developed to enhance the capabilities of the fault detection systems for gearboxes in wind turbines. The framework integrates different conventional techniques to speed up the process of vibration signal filtering and fault detection. It also support non-stationary signal processing suitable for detecting faults in wind turbines gearboxes.

(2) A hybrid evolutionary framework has been developed to harness the optimization capabilities of the evolutionary algorithms and at the same time making it suitable for real world applications such as gearbox fault diagnosis in wind turbines.

10.4.3 Time-domain Transient Based Features Extraction

Three novel time-domain transient based feature extraction techniques have been proposed for fault detection in gearboxes. The developed techniques have been successfully tested on different vibration signals emanating from different gearboxes operating under different conditions. The extracted features are representative of the gearbox health conditions.

10.4.4 Decision Support Systems

Three types of decision support systems have been proposed and their applicability on gearbox fault diagnosis, cause and consequence analysis have been demonstrated.

10.4.5 Real World Case Studies

The developed methods have been demonstrated by two real world case studies. Vibration data from real world wind turbines have been analysed and the fault detection capabilities of the proposed methods have been verified. Also, the applicability of the extracted features to determine the health of components in wind turbine mechanical drive train has been verified.

10.5 Future Research

In the framework of future research, we propose the followings:

10.5.1 Time-Domain Features for Bearing Fault Detection

In this research work, we have presented frequency-domain features extraction framework for fault detection of roller element bearings in Chapter 6. We have used the developed filtering mechanisms presented in Chapter 4 and 5, and analysed the filtered waveform with FFT to extract features related with characteristic defect frequencies for bearings. Therefore, the present research work lacks a detailed investigation for time-domain features extraction for bearing fault detection. Further investigation is needed to develop new methods for time-domain features extraction for detection of different faults in bearings as proposed in [130].

10.5.2 Condition Based Maintenance for Wind Turbines

Maintenance process has evolved over the years from breakdown to preventive and then predictive in tandem with condition based maintenance (CBM) as shown in Figure 10.1. The CBM approach, if implemented correctly can reduce the total operation and maintenance costs of wind turbines. In this research, our objective is to use real time vibration data and the vibration features extracted in Chapter 6 combined with prognostic approach in Chapter 7 to devise failure probability estimation at component level and system level as described in [131; 132] .

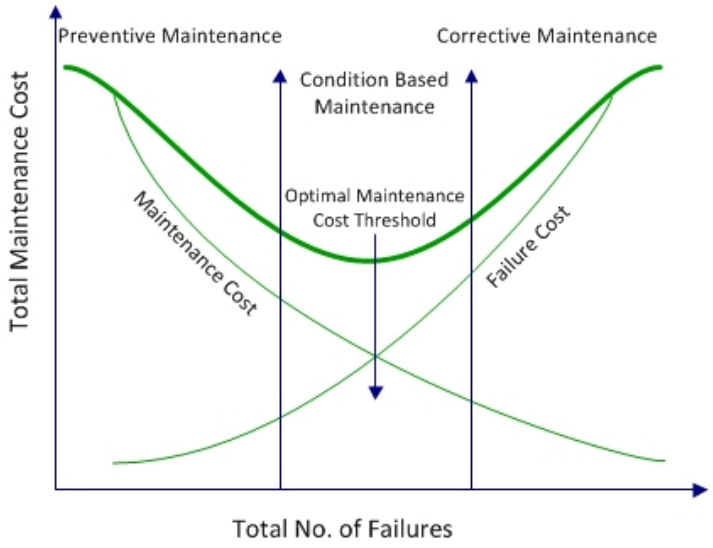


Figure 10.1: Condition Based Maintenance Approach

References

[1] C. C. James. “Condition Monitoring Techniques for Wind Turbines”, PhD thesis, Durham University, UK, 2011.

[2] “Greenhouse Gas Sources and Sinks in Canada”, National Inventory Report, Environment Canada, 1990-2010.

[3] “Electricity Generation Resources”, Canadian Electrical Association. [Online]. Available: [Cited: 2nd Feb, 2013]: http://www.electricity.ca/media/Electricity101/Electricity%20101.pdf.

[4] “WindVision 2025: Powering Canada's Future”, Canadian Wind Energy Association. [Online]. Available: [Cited: February 02, 2013]: http://www.canwea.ca/windvision_e.php

[5] Tavner, P. J., Xiang, J., Spinato, F., “Reliability Analysis for Wind Turbines”, Wind Energy, 10(1) (2007) 1-18.

[6] Faulstich, S. et al., “Windenergie Report Deutschland”, Institut für solare Energieversorgungstechnik (Hrsg.), Kassel, 2008.

[7] Mari Cruz Garcia, Miguel A. Sanz-Bobi, Javier del Pico, "SIMAP: Intelligent System for Predictive Maintenance Application to the Health Condition Monitoring of a Wind Turbine Gearbox", Computers in Industry, 57(6) (2006) 552–568.

[8] Thomas Braun, CEO, LGC kyrota Wind Energy Corp (2009). LGC Skyrota Wind Energy Corp. Selected by Eco Wind Power to Provide Gearbox Maintenance Services [Online]. Available: [Cited: 20th March, 2012]: http://www.newswire.ca/en/story/483435/lgc-skyrota-wind-energy-corp-selected-by-eco-wind-power-to-provide-gearbox-maintenance-services.

[9] Steve Barber, P. Golbeck. Case study 2: Wind Farm in Canada, Megawatt Class Turbines, Age: 3 years. [Online]. Available [Cited: 20th March, 2012]: http://www.wwindea.org/technology/ch03/en/3_4_3.html.

[10] Enayet B. Halim, M.A.A. Shoukat Choudhury, Sirish L. Shah, Ming J. Zuo, "Time Domain Averaging Across All Scales: A Novel Method for Detection of Gearbox Faults", Mechanical Systems and Signal Processing, 22(2) (2008) 261–278.

[11] F. Combet, L. Gelman, "An Automated Methodology for Performing Time Synchronous Averaging of A Gearbox Signal Without Speed Sensor", Mechanical Systems and Signal Processing, 21(6) (2007) 2590–2606.

[12] Yong Xiang Zhang, R.B.Randall, "Rolling Element Bearing Fault Diagnosis Based on the Combination of Genetic Algorithms and Fast Kurtogram", Mechanical Systems and Signal Processing, 23(5) (2009) 1509–1517.

[13] Amir Hosein Zamanian, Abdolreza Ohadi, "Gearbox Fault Detection through PSO Exact Wavelet Analysis and SVM Classifier", 18th Annual International Conference on Mechanical Engineering, (2010) 1-6.

[14] Jianping Ma, Jin Jiang, "Applications of Fault Detection and Diagnosis Methods in Nuclear Power Plants: A Review", Progress in Nuclear Energy, 53 (2011) 255-266.

[15] H. Link, W. LaCava, J. van Dam, B. McNiff, S. Sheng, R. Wallen, M. McDade, S. Lambert, S. Butterfield, F. Oyague, "Gearbox Reliability Collaborative Project Report: Findings from Phase 1 and Phase 2 Testing," NREL/TP-5000-51885, 2011.

[16] Oyague, F., S. Sheng, S. Butterfield. "NREL Gearbox Reliability Collaborative Analysis Round Robin," Wind Power Conference, 2009.

[17] "Data Analysis Competition - 2009", Prognostics and Health management Society. [Online]. Available [Cited: December, 2012]: http://www.phmsociety.org/competition/PHM/09/apparatus.

[18] Hui Li, YupingZhang, Haiqi Zheng, "Application of Hermitian Wavelet to Crack Fault Detection in Gearbox ", Mechanical Systems and Signal Processing, 25 (2011) 1353–1363.

[19] Cyril Hory, Nadine Martin, Alain Chehikian, "Spectrogram Segmentation by Means of Statistical Features for Non-stationary Signal Interpretation", IEEE Transactions on Signal Processing, 50 (12) (2002) 2915-2925.

[20] Hui Li, Haiqi Zheng, Liwei Tang, "Wigner Ville, Distribution based on EMD for Faults Diagnosis of Bearing", Lecture Notes in Computer Science, 4223 (2006) 803–812.

[21] Hui Li, Yuping Zhang, Haiqi Zheng, "Hilbert–Huang Transform and Marginal Spectrum for Detection and Diagnosis of Localized Defects in Roller Bearings", Journal of Mechanical Science and Technology, 23 (2) (2009) 291–301.

[22] Jerome Antoni, "Fast Computation of the Kurtogram for the Detection of Transient Faults", Journal of Mechanical systems and Signal Processing, 21(1) (2007) 108-124.

[23] N. Saravanan, V.N.S. Kumar Siddabattuni, K.I. Ramachandran, "A Comparative Study on Classification of Features by SVM and PSVM Extracted using Morlet Wavelet for Fault Diagnosis of Spur Bevel Gearbox", Expert Systems with Applications, 35(3) (2008) 1351–1366.

[24] N. Saravanan, K.I. Ramachandran, "A Case Study on Classification of Features by Fast Single-shot Multiclass PSVM using Morlet Wavelet for Fault Diagnosis of Spur Bevel Gearbox", Expert Systems with Applications, 36(8) (2009) 10854–10862.

[25] Halima, E. B., Shoukat Choudhury, M. A. A., Shah, S. L. and Zuo, M. J., "Time Domain Averaging Across all Scales: A Novel Method for Detection of Gearbox Faults", Mechanical Systems and Signal Processing, 22(2) (2008) 261-278.

[26] Lei, Y. G., He, Z. J., Zi, Y. Y., Chen, X. F., "New Clustering Algorithm Based Fault Diagnosis Using Compensation Distance Evaluation Technique", Mechanical Systems and Signal Processing, 22(2) (2008) 419-435.

[27] Lei, Y. G., He, Z. J., ZI, Y. Y., "A New Approach to Intelligent Fault Diagnosis of Rotating Machinery", Expert Systems with Applications, 35(4) (2008) 1593-1600.

[28] Jiawei Xiang, Yongteng Zhong, Xuefeng Chen, Zhengjia He, "Crack Detection in a Shaft by Combination of Wavelet Based Elements and Genetic Algorithm", International Journal of Solids and Structures, 45 (17) (2008) 4782-4795.

[29] Xinsheng Lou, K. A. Kenneth, A. Loparo, "Bearing Fault Diagnosis Based on Wavelet Transform and Fuzzy Inference", Mechanical Systems and Signal Processing, 18 (5) (2004) 1077-1095.

[30] J. Rafiee, M.A. Rafiee, P.W. Tse, "Application of Mother Wavelet Functions for Automatic Gear and Bearing Fault Diagnosis", Expert Systems with Applications, 37 (6) (2010) 4568-4579.

[31] Jing Yuan, Zhengjia He, Yanyang Zi, "Gear Fault Detection Using Customized Multiwavelet Lifting Schemes", Mechanical Systems and Signal Processing, 24 (5) (2010) 1509-1528.

[32] F. Combet, L. Gelman, "Optimal Filtering of Gear Signals for Early Damage Detection Based on the Spectral Kurtosis", Mechanical Systems and Signal Processing, 23 (3) (2009) 652-668.

[33] Tomasz Barszcz, Robert B. Randall, "Application of Spectral Kurtosis for Detection of a Tooth Crack in the Planetary Gear of a Wind Turbine", Mechanical Systems and Signal Processing, 23 (4) (2009) 1352-1365.

[34] N. Sawalhi, R.B. Randall, H. Endo, "The Enhancement of Fault Detection and Diagnosis in Rolling Element Bearings Using Minimum Entropy Deconvolution Combined with Spectral Kurtosis", Mechanical Systems and Signal Processing, 21 (6) (2007) 2616-2633.

[35] Hiram Firpi, George Vachtsevanos, “Genetically Programmed-based Artificial Feature-extraction Applied to Fault Detection”, Engineering Applications of Artificial Intelligence, 21 (4) (2008) 558-568.

[36] V. Indira, R. Vasanthakumari, V. Sugumaran, “Minimum Sample Size Determination of Vibration Signals in Machine Learning Approach to Fault Diagnosis Using Power Analysis”, Expert Systems with Applications, 37 (12) (2010) 8650-8658.

[37] Jordan McBain, Markus Timusk,” Fault Detection in Variable Speed Machinery: Statistical Parameterization”, Journal of Sound and Vibration, 327 (3-5) (2009) 623-646.

[38] Yaguo Lei, Ming J. Zuo, Zhengjia He, Yanyang Zi, “A Multidimensional Hybrid Intelligent Method for Gear Fault Diagnosis”, Expert Systems with Applications, 37 (2) (2010) 1419-1430.

[39] Zˇeljko Kanovic´, Milan R. Rapaic´, Zoran D. Jelicˇic, "Generalized Particle Swarm Optimization Algorithm - Theoretical and Empirical Analysis with Application in Fault Detection", Applied Mathematics and Computation, 217(24) (2011) 10175–10186.

[40] M. Wahde, "Biologically Inspired Optimization Methods – An Introduction", WIT Press, Southampton UK (2008).

[41] Hitoshi Iba, Topon Kumar Paul, Yoshihiko Hasegawa, "Applied Genetic Programming and Machine Learning", CRC Press, Taylor & Francis Group (2010).

[42] Qing Zhao, Zhihan Xu, "Design of a Novel Knowledge-Based Fault Detection and Isolation Scheme ", IEEE Trans. Syst., Man, Cybern. B, Cybern., 34 (2) (2004) 1089-1095.

[43] Hong Guo, Lindsay B. Jack, Asoke K. Nandi, "Feature Generation Using Genetic Programming With Application to Fault Classification", IEEE Trans. Syst., Man, Cybern. B, Cybern., 35 (1) (2005) 89-99.

[44] Feng Zhao, Xenofon Koutsoukos, Horst Haussecker, Jim Reich, Patrick Cheung, "Monitoring and Fault Diagnosis of Hybrid Systems", IEEE Trans. Syst., Man, Cybern. B, Cybern., 35 (6) (2005) 1225-1240.

[45] Yongxiang Zhang, R.B. Randall, "Rolling Element Bearing Fault Diagosis Based on the Combination of Genetic Algorithms and Fast Kurtogram", Mechanical Systems and Signal Processing, 23 (2009) 1509-1517.

[46] Bing Li, Pei-lin Zhang, Hao Tian, Shuang-shan Mi, Dong-sheng Liu, Guo-quan Ren, "A New Feature Extraction and Selection Scheme for Hybrid Fault Diagnosis of Gearbox", Expert Systems with Applications, 38(8) (2011) 10000–10009.

[47] De Z. Li, Wilson Wang, "An Enhanced GA Technique for System Training and Prognostics", Advances in Engineering Software, 42(7) (2011) 452–462.

[48] Irina Ciornei, Elias Kyriakides, "Hybrid Ant Colony-Genetic Algorithm (GAAPI) for Global Continuous Optimization", IEEE Trans. Syst., Man, Cybern. B, Cybern., 42 (1) (2012) 234-245.

[49] ZhangWei guo, LuTian yu, "The Research of Genetic Ant Colony Algorithm and Its Application", Procedia Engineering 37 (2012) 101 – 106.

[50] Kennedy, J., Eberhart, R.C., "Particle Swarm Optimization", In Proceedings of the IEEE International Conference on Neural Networks, IV (1995) 1942–1948.

[51] B. Samanta, C.Nataraj, "Use of Particle Swarm Optimization for Machinery Fault Detection", Engineering Applications of Artificial Intelligence, 22 (2009) 308–316.

[52] Poli, R., Kennedy, J., Blackwell, T., "Particle Swarm Optimization: An Overview", Swarm Intelligence 1 (2007) 33–57.

[53] Tao Cai, Feng Pan, Jie Chen, "Adaptive Particle Swarm Optimization Algorithm", In Proceedings of Fifth World Congress on Intelligent Control and Automation, 3 (2004) 2245‑2247.

[54] Yuelin Gao, Yuhong Duan, "An Adaptive Particle Swarm Optimization Algorithm with New Random Inertia Weight", In Advanced Intelligent Computing Theories and Applications with Aspects of Contemporary Intelligent Computing Techniques, 2 (2007) 342-350.

[55] Yuelin Gao, Zihui Ren, "Adaptive Particle Swarm Optimization Algorithm with Genetic Mutation Operation", In Proceedings of the Third International Conference on Natural Computation, 2 (2007) 211‑215.

[56] Chen DeBaoa, Zhao ChunXia, "Particle Swarm Optimization with Adaptive Population Size and Its Application", Applied Soft Computing, 9(1) (2009) 39–48.

[57] Pan Hongxia, Wei Xiuye, Xu Xin, "Research of Optimal Placement of Gearbox Sensor Based on Particle Swarm Optimization", 8th IEEE International Conference on Industrial Informatics (INDIN), (2010) 108-113.

[58] Cao Feng-cai, Wei Xiuye, "Research on Measuring Point Configuration Based on Particle Swarm Optimization Technology in Fault Diagnosis", International Conference on Computer Application and System Modeling, 15 (2010) 543 - 546.

[59] Pan Hongxia, Ma Qingfeng, Wei Xiuye, "Research on Fault Diagnosis of Gearbox Based on Particle Swarm Optimization Algorithm", IEEE International Conference Mechatronics, (2006) 32 - 37.

[60] Liu Ying, Liu Jie, Mao Hongwei, Pan Hongxia, "Wavelet Neural Network Based on Modified PSO and its Application in Fault Diagnosis of Gearbox", Fifth International Conference on Natural Computation, 3 (2009) 338 - 342.

[61] B. Paya, I. Esat, M. Badi, "Artificial Neural Network Based Fault Diagnostics of Rotating Machinery Using Wavelet Transforms as A Preprocessor," Mechanical Systems and Signal Processing, 11(5) (1997) 751-765.

[62] L. Kuncheva, "How Good are Fuzzy If-Then Classifiers," IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, 30(3) (2000) 501-509.

[63] T. Liu, J. Singonahalli, N. Iyer, "Detection of Roller Bearing Defects Using Expert System and Fuzzy Logic," Mechanical Systems and Signal Processing, 10(5) (1996) 595-614.

[64] X. Lou, K. Loparo, "Bearing Fault Diagnosis Based On Wavelet Transform and Fuzzy Inference," Mechanical Systems and Signal Processing, 18 (2004) 1077-1095.

[65] Jian-Da Wu, Chuang-Chin Hsu, "Fault Gear Identification Using Vibration Signal With Discrete Wavelet Transform Technique and Fuzzy–Logic Inference", Expert Systems with Application, 36(2) (2009) 3785–3794.

[66] A. Biedermann, F. Taroni, "Bayesian Networks For Evaluating Forensic DNA Profiling Evidence: A Review and Guide To Literature", Forensic Science International: Genetics, 6 (2012) 147‑157.

[67] Bin Gang Xu, "Intelligent Fault Inference for Rotating Flexible Rotors Using Bayesian Belief Network", Expert Systems with Applications, 39(1) (2012) 816–822.

[68] Yuan Kang, Chun-Chieh Wang, and Yeon-Pun Chang, "Gear Fault Diagnosis in Time Domains by Using Bayesian Networks", Analysis and Design of Intelligent Systems using Soft Computing Techniques, Springer-Verlag, Berlin, 41(2007) 618-627.

[69] Jigang Chen, Guowen Hao, "Research on the Fault Diagnosis of Wind Turbine Gearbox Based on Bayesian Networks", Practical Applications of Intelligent Systems, Springer-Verlag Berlin, 124 (2012) 217-223.

[70] B. Samanta, C. Nataraj, "Prognostics of Machine Condition using Soft Computing", Robotics and Computer-Integrated Manufacturing, 24(6) (2008) 816– 823.

[71] Andrew K.S. Jardine, Daming Lin, Dragan Banjevic, "A Review on Machinery Diagnostics and Prognostics Implementing Condition-based Maintenance, "Mechanical Systems and Signal Processing, 20 (7) (2006) 1483–1510.

[72] Wang WQ, Golnaraghi MF, Ismail F., “Prognosis of Machine Health Condition using Neuro-fuzzy Systems”, Mechanical Systems and Signal Processing, 18(4) (2004) 813–31.

[73] Wang W., “An Adaptive Predictor for Dynamic System Forecasting”, Mechanical Systems and Signal Processing, 21(2) (2007) 809–23.

[74] Wang W., Ismail F., Golnaraghi F., “A Neuro-fuzzy Approach to Gear System Monitoring”, IEEE Transactions on Fuzzy Systems, 12 (2004), 710–23.

[75] De Gooijer JG, Hyndman RJ., “25 Years of Time Series Forecasting”, Int J Forecasting, 22 (2006) 443–73.

[76] L. Gelman, I. Petrunin, I. K. Jennions, M. Walters., “Diagnostics of Local Tooth Damage in Gears by the Wavelet Technology”, International Journal of Prognostic and Health Management, 3(2) 2012.

[77] Tse P., Atherton D., “Prediction of Machine Deterioration using Vibration Based Fault Trends and Recurrent Neural Networks”, J Vib Acoust, 121 (1999) 355–62.

[78] Jang JSR., “ANFIS: Adaptive-network-based Fuzzy Inference System”, IEEE Trans Syst, Man Cybernet, 23 (1993) 665–85.

[79] Jamie Coble, J. Wesley Hines., “Applying the General Path Model to Estimation of Remaining Useful Life”, International Journal of Prognostic and Health Management: 2(1) 2011.

[80] Hai Qiu, Jay Lee, Jing Lin, Gang Yu, "Wavelet Filter-Based Weak Signature Detection Method and Its Application on Rolling Element Bearing Prognostics", Journal of Sound and Vibration, 289(4-5) (2006) 1066–1090.

[81] Jing Lin, M. J. Zuo, "Gearbox Fault Diagnosis using Adaptive Wavelet Filter", Journal of Mechanical Systems and Signal Processing, 17(6) (2003) 1259–1269.

[82] Jing Lin, "Feature Extraction Based on Morlet Wavelet and Its Application for Mechanical Fault Diagnosis", Journal of Sound and Vibration, 234(1) (2000) 135-148.

[83] Ruqiang Ya, Robert X. Gao, "Harmonic Wavelet-based Data Filtering for Enhanced Machine Defect Identification", Journal of Sound and Vibration, 329(15) (2010) 3203–3217.

[84] Tavner, P. J., Yang, W., “Survey of Commercially Available Wind Turbine Condition Monitoring Systems”, Supergen Wind Energy Technologies Consortium, under revision to April 2009.

[85] Feng, Y., Crabtree, C. J., Long, H., Tavner, P. J., "Use of SCADA and CMS Signals for Failure Detection and Diagnosis of a Wind Turbine Gearbox", Proceedings of the European Wind Energy Conference, Brussels, Belgium, March 2011.

[86] R. W. Hyers, J. G. McGowan, K. L. Sullivan, J. F. Manwell and B. C. Syrett, "Condition Monitoring and Prognosis of Utility Scale Wind Turbines", Energy Materials 1(3) (2006) 187-203.

[87] Becker E., Posta P., "Keeping the Blades Turning: Condition Monitoring of Wind Turbine Gears", Refocus 7(2) (2006) 26-32.

[88] Editorial, "Managing the Wind: Reducing Kilowatt Hour Costs with Condition Monitoring", Refocus 6(3) (2005) 48-51.

[89] Echavarria E, Tomiyama T, Van Bussel G., "Fault Diagnosis Approach Based on a Model-based Reasoner and a Functional Designer for a Wind Turbine: An Approach towards Self-maintenance", Journal of Physics Conference Series 75(1) (2007).

[90] Echavarria E., Tomiyama T., Huberts H., Bussel G., "Fault Diagnosis System for an Offshore Wind Turbine Using Qualitative Physics," In Proc. EWEC2008, Brussels, Belgium, 2008.

[91] Zaher A., McArthur S., "A Multi-agent Fault Detection System for Wind Turbine Defect Recognition and Diagnosis." In Proc. IEEE Lausanne Power Tech (2007) 22-27.

[92] Crabtree, C. J., "Survey of Commercially Available Condition Monitoring Systems for Wind Turbines" [Online]. Available: [Cited: 2nd Feb, 2013]:http://www.supergen-wind.org.uk.

[93] Ronald L. Rardin, "Optimization in Operations Research", Prentice Hall, New Jersey, (1998) 727- 735.

[94] Steven W. Smith, "The Scientist and Engineer's Guide to Digital Signal Processing", California Technical Publishing, (2002) 333-341.

[95] Sajid Hussain, Hossam A. Gabbar, "A Novel Method For Real Time Gear Fault Detection Based on Pulse Shape Analysis", Mechanical Systems and Signal Processing, 25(4) (2011) 1287–1298.

[96] Sajid Hussain, Hakan Grahn, "Fast *kd*-tree Construction for 3D-rendering Algorithms Like Ray Tracing", Proceedings of the 3rd International Conference on Advances in Visual Computing, (LNCS) (2007) 681-690.

[97] Chapra, "Applied Numerical Methods with MATLAB for Engineers and Scientists", McGraw-Hill Primis, (1996) 171- 178.

[98] J. H. Holland, "Adaptation in Natural and Artificial Systems", The University of Michegan Press, (1975).

[99] Biot, Maurice A., “Transient Oscillations in Elastic Systems”, PhD Thesis, California Institute of Technology, 1932.

[100] Sajid Hussain, Hossam A. Gabbar, "A Robust Method For Coupling Detection Among Process Variables", Int. J. Process Systems Engineering, 2 (1) (2012) 93–110.

[101] Sajid Hussain, Hossam A. Gabbar, "Fault Diagnosis In Gearbox Using Adaptive Wavelet Filtering and Shock Response Spectrum Features Extraction", Int. J. Structural Health Monitoring, 12 (1) (2013) 1–12.

[102] Seok Won Lee, Boo Hee Nam, “Peak Detection of ECG Signal using Wavelet Transform and Radial Bases Functions”, Proceedings of the 39th SICE Annual Conference, (1999) 1067-1070.

[103] David O. Smallwood, “An Improved Recursive Formula for Calculating Shock Response Spectra”, Shock and Vibration Bulletin, 51 (1981).

[104] S.J. Drew, B.J. Stone, "Torsional Damping Measurements for a Gearbox", Mechanical Systems and Signal Processing, 19 (2005) 1096–1106.

[105] Sajid Hussain, Hossam A. Gabbar, "Non-destructive Testing of Machines to Reduce Maintenance Time and Cost", Reliability and Maintainability Symposium (RAMS), (2012) 1–6.

[106] Randall, R. B., “Frequency Analysis”, Copenhagen: Bruel & Kjaer, 1987Jafar.

[107] Jafar Zarei, “Induction Motors Bearing Fault Detection Using Pattern Recognition Techniques”, Expert Systems with Applications 39 (2012) 68–73.

[108] E. Zwicker, H. Fastl., “Psychoacoustics - Facts and Models”, Springer 2nd edition, 1999.

[109] S. Seneff, “A Joint Synchrony/mean-rate Model of Auditory Speech Processing”, Journal of Phonetics 16(1) (1988) 55-76.

[110] R. D. Patterson, K. Robinson, J. Holdsworth, D. McKeown, C.Zhang, and M. H. Allerhand, “Complex Sounds and Auditory Images”, Proc. 9th International Symposium on Hearing, (1992) 123-177.

[111] Graps, A., “An Introduction to Wavelets”, IEEE Computational Science and Engineering, 2(2) (1995), 50–61.

[112] Edmundo G. de Souza e Silva, Luiz F.L. Legey and Edmundo A. de Souza e Silva, "Forecasting Oil Price Trends Using Wavelets and Hidden Markov Models", Energy Economics 32 (2010) 1507–1519.

[113] T. Lin, B. G. Horne, P. Tino, C. L. Giles, "A Delay Damage Model Selection Algorithm for NARX Neural Networks", IEEE Transactions on Signal Processing, 45(11) (1997) 2719–2730.

[114] RWC Belgium World Data Center, Online sunspot data archive, SIDC, [Online]. Available: [Cited: 20th Dec, 2012]: http://sidc.oma.be/index.php3.

[115] Zadeh, L.A., "Fuzzy Sets", Information and Control, 8(3) (1965) 338–353.

[116] Bezdek JC., "Pattern Recognition with Fuzzy Objective Function Algorithms", New York, Plenum, 1981.

[117] Wang XL., "A Course in Fuzzy Systems and Control", New Jersey, Prentice Hall, Inc., 1997.

[118] Richens, R.H., "Preprogramming for Mechanical Translation", Mechanical Translation 3(1) (1956) 20-25.

[119] Collins, A. M., Quillian, M. R., "Retrieval Time from Semantic Memory", Journal of Verbal Learning and Verbal Behavior 8(2) (1969) 240–247.

[120] Hossam A. Gabbar, "Improved Qualitative Fault Propagation Analysis", Journal of Loss Prevention in the Process Industries 20 (2007) 260–270.

[121] Hossam A. Gabbar, "Fault Semantic Network for Accident Forecasting of LNG Plants", KES, Part II, LNAI 6277, 2010, 427–437.

[122] Hossam A. Gabbar, Faisal I. Khan, "Design of Fault Semantic Networks to Integrate Fault, Failure, Hazard, and Accident Models for LNG Plants", International Workshop on Real Time Measurement, Instrumentation & Control [RTMIC]. 2010, 7(1) – 7(8).

[123] Gregory F. Cooper Edward Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data", Machine Learning 9 (1992) 309-347.

[124] Anders L. Madsen , FinnV. Jensen, "LAZY Propagation: A Junction Tree Inference Algorithm Based on Lazy Evaluation", Artificial Intelligence 113 (1999) 203–245.

[125] Ezio S, Claudio C., "Exploitation of Wind as an Energy Source to Meet the World's Electricity Demand", Journal of Wind Engineering and Industrial Aerodynamics 74 (1998) 375-387.

[126] Haykin, S., “Neural Networks: A Comprehensive Foundation”, MacMillan, New York 1994.

[127] I.T. Jolliffe, “Principal Component Analysis”, 2nd Edition, Springer, New York, 2002.

[128] Neelam Mehala, “Condition Monitoring and Fault Diagnosis of Induction Motor Using Motor Current Signature Analysis”, PhD thesis, National Institute of Technology, Kurukshetra, India, 2010.

[129] Wu, C.Y., Pollock, C., “Time Domain Analysis of Vibration and Acoustic Noise in the Switched Reluctance Drive”, Sixth International Conference on Electrical Machines and Drives, (1993) 558 – 563.

[130] Tahsin Doguer, Jens Strackeljan,"Vibration Analysis using Time Domain Methods for the Detection of Small Roller Bearing Defects", 8th International Conference on Vibrations in Rotating Machines, Vienna, Austria, (2009) 1 – 12.

[131] Zhigang Tian, Tongdan Jin, Bairong Wu, Fangfang Ding, “Condition Based Maintenance Optimization for Wind Power Generation Systems under Continuous Monitoring”, Renewable Energy 36 (2011) 1502–1509.

[132] Fangfang Ding, Zhigang Tian, “Opportunistic Maintenance for Wind Farms Considering Multi-level Imperfect Maintenance Thresholds”, Renewable Energy (2012) 1–8.